

Article

Impact of Sampling Technique on the Performance of Surrogate Models Generated with Artificial Neural Network (ANN): A Case Study for a Natural Gas Stabilization Unit

Mohamed Ibrahim ¹, Saad Al-Sobhi ^{1,*}, Rajib Mukherjee ² and Ahmed AlNouss ¹ 

¹ Chemical Engineering Department, College of Engineering, Qatar University, Doha 2713, Qatar; m.ibrahim@qu.edu.qa (M.I.); a.alnouss@qu.edu.qa (A.A.)

² Gas and Fuels Research Center, Texas A&M Engineering Experiment Station, College Station, TX 77843, USA; rajib@tamu.edu

* Correspondence: saad.al-sobhi@qu.edu.qa

Received: 2 March 2019; Accepted: 15 May 2019; Published: 18 May 2019



Abstract: Data-driven models are essential tools for the development of surrogate models that can be used for the design, operation, and optimization of industrial processes. One approach of developing surrogate models is through the use of input–output data obtained from a process simulator. To enhance the model robustness, proper sampling techniques are required to cover the entire domain of the process variables uniformly. In the present work, Monte Carlo with pseudo-random samples as well as Latin hypercube samples and quasi-Monte Carlo samples with Hammersley Sequence Sampling (HSS) are generated. The sampled data obtained from the process simulator are fitted to neural networks for generating a surrogate model. An illustrative case study is solved to predict the gas stabilization unit performance. From the developed surrogate models to predict process data, it can be concluded that of the different sampling methods, Latin hypercube sampling and HSS have better performance than the pseudo-random sampling method for designing the surrogate model. This argument is based on the maximum absolute value, standard deviation, and the confidence interval for the relative average error as obtained from different sampling techniques.

Keywords: surrogate model; sampling technique; stabilization unit; process simulation; process systems engineering (PSE)

1. Introduction

The integration of simulation tools with rigorous mathematical programming techniques is a challenging process due to the inherent complexity of calculations with a process simulator [1,2]. To overcome this limitation, a surrogate model for the process is derived based on the simulation results and included within an optimization framework. The surrogate models of a process flowsheet are generated based on the input parameters and output results of the process. The entire process or system is treated as a black box with input and output data, and the surrogate models are generated based on the sampled data. To obtain the sampled data, data of the input variables are exported from the process flowsheet, and the corresponding output data are calculated.

Several methods may be used to develop surrogate model generation including Kriging models, artificial neural networks (ANN), and simple table lookup method [3]. In this work, ANN fitting is used to predict the output variables (Y), from a set of input variables (X). The number of variables (dimension of X) corresponds to the dimension of the sampling space. In an artificial neural network, a specific input to a specified output is mapped. The network achieves this by a self-organizing process.

ANN is capable of learning complicated functional relations by creating a general trend from a limited number of input data; hence, it works as a black-box model for different systems such as nonlinear, multivariable static, and dynamic systems. Based on the input and output data set obtained from a process, the approach of artificial neural networks is implemented to generate surrogate models. The ability of ANN to learn from observed data is the main advantage of the process [4]. Therefore, it was applied to the present case study for surrogate model generation. Several important applications have been reported in the literature. For example, a new method for optimizing crude oil distillation systems which consider an ANN model for representing the distillation process has been proposed by Ochoa-Estopier et al. (2013) [5].

The distillation column and the accompanied heat exchangers network have been considered in an optimization study to thoroughly define the parameters that enhance the economic aspects of the process. One of the important issues in that field is optimizing the process parameters to favor a higher yield from the high-value products and at the same time decreases the yield of the lower value products. The feasibility of these parameters is represented in terms of utility cost and equipment limitations. In this specific study, the input variables of the distillation column are represented in the crude oil properties and the distillation column specifications. On the other hand, the output variables are the distillation products specifications. Hence, the ANN model is built based on these input and output parameters. Liao et al. (2004) [6] designed a similar system to examine the effect of input parameters on the end product of the distillation column. The results of the system were used as a database for further operational optimizations. The built system can provide on-line optimal operating information of the distillation process to the operators corresponding to the change of crude oil properties.

Motlaghi et al. (2008) [7] optimized a crude oil distillation column using a genetic algorithm (GA) with the aim of the system's output minimization and product rate requirement maximization for a system of distillation columns for crude oil. The input data for the system was a set of operational parameters such as crude oil flow rate, feed temperature to optimize the crude oil properties based on desired specifications. Osulale and Zhang (2015) [8] proposed a new approach for energy consumption optimization for the same topic. The proposed model incorporates ANN that is used to improve system performance and predictability. An economic analysis of the system revealed possible energy saving opportunities by using more efficient process parameters. Similarly, Arce-Medina and Paz-Paredes (2009) [9] developed ANN model to predict the performance of the hydro-desulfurization process for the prediction of sulfur removal from naphtha.

In all the above methods process data as obtained for the ANN model generation was gathered via Monte Carlo sampling. In the present work, the focus is given to the accuracy of data generation and the coupling with input-output data generators such as process simulators. The surrogate model must be generated using the equivalent distribution of sampling points for each variable such that the whole design space is covered. In this work, the developed surrogate model is robust enough to represent the original process. The case study will investigate how the selection of sampling techniques influences the performance of the surrogate model. An investigation on how to sample the design space for a surrogate model fitting is done. Three different sampling techniques are considered: Pseudo-random Sampling, Latin hypercube sampling (LHS), and quasi-monte Carlo using Hammersley Sequence Sampling (HSS) are used in our analysis.

Eason and Cremaschi (2014) [10] proposed new adaptive sequential sampling algorithms: the first is only adaptive and the second combines adaptive and space-filling properties for determining and minimizing the number of samples required to generate an efficient surrogate model. The mixed adaptive sampling algorithm was found to be more efficient; reducing the size of the sample needed by up to 40% compared to a space-filling design in optimizing the CO₂ capture process using aqueous amines.

A computational algorithm for LHS was developed by McKay et al. [11]. It was further developed by Iman and Shortencarier [12] and then by Stein [13]. LHS can improve efficiency compared to the Monte Carlo. HSS is also an efficient sampling technique developed by Diwekar and co-workers

(Kalagnanam and Diwekar, 1997) [14]. It uses Hammersley points to sample a unit n -dimensional hypercube uniformly. It causes the sampling process distributed uniformly over n -dimensional space. Simulation with HSS results in faster conversion. It is possible because HSS has n -dimensional uniformity. For example, to find the morphology of polymers in dilute solution n -dimensional uniformity of sampling is required to obtain the optimal configuration. It is shown that HSS has led to an optimized system that cannot be achieved by random sampling [15].

An accurate model of chemical processes comes with increasing complexity which leads to a substantial computational effort. When a clear relationship between input variables and their responses is not accessible to the designer, it is challenging to perform optimization and study sensitivity and reliability analysis. A surrogate model, on the other hand, can replace these complex, underlying models with simple models which has several advantages including much faster simulation [16]. In recent times we see increasing use of surrogate model within the chemical process engineering particularly for model predictive control (MPC) [17], model-based optimization [18], sensitivity analysis [19], reliability assessment [20] among others. In this work, we developed a surrogate model using ANN method for natural gas (NG) stabilizing unit. Natural gas stabilization unit is the primary unit of any gas processing facility. Optimal separation of residual gas and condensate is required for maximum profit. The model is intended to perform model-base optimization of the operating conditions to meet the specifications with the minimum operating cost. For the case of the stabilization unit, we want to optimize the stabilizer reboiler temperature to find the proper Reid Vapor Pressure (RVP) value.

ANNs are reliable and flexible surrogate models that can be used to simulate challenging processes such as natural gas stabilization units by identifying the nonlinear relationships between independent and dependent variables. Overall, the numerical performance of ANNs models is diverse, and it depends on the specific simulated process. Some of the available surrogate models could show limitations based on the nature of the process. Through the ANN-based model, we can predict the dependent variable after learning from a given sample set. Hence, it is essential to fine-tune the ANN by optimizing the sampling method. In the case of NG stabilization unit, finding an appropriate sampling technique is a paramount step that can ascertain how successful the surrogate model can be. This is the first paper in the series, and the prime of this work is to demonstrate that different sampling techniques can affect the model accuracy and identify the appropriate sampling technique and thus the model. The work will then be extended to apply the framework and optimize the operating conditions of a fully integrated facility (more than one processing unit) such as Liquefied natural gas (LNG) facility with more specifications i.e., Wobbe index, H_2S , CO_2 content, and so on.

The paper assesses the impact of selecting the sampling technique on the model robustness and introduces a computational approach to the coupling of surrogate models with optimization platforms. A case study on natural gas stabilization is used to demonstrate the proposed approach. The remainder of the paper is structured as follows: The relevant previous works on natural gas stabilization are discussed in Section 2. The overall methodology is mentioned in Section 3. Section 4 discusses the case study findings. Finally, concluding remarks are stated in Section 5.

2. Process Description: Gas Stabilization Unit

The natural gas stabilization unit is used to recover the heavy components from the feed stream to sell them as products. Sale of optimal liquids products maximizes profits by optimizing the fraction of liquids recovered while achieving the requirement for the residual gas sent to subsequent units. The stabilized liquid has two important characteristics that must be met to achieve safe and environmental handling, processing, and exportation. The stabilization process is performed in the chemical industry through fractionation. Flash vaporization is an old technique and is not being used in the gas plant [21]. It is replaced by gas stabilization by fractionation as an accepted approach in the industry. The gas stabilization is typically carried out in an absorber with a reboiler and internal trays. Figure 1 shows the flow diagram of the gas stabilization unit. The used of a refluxed distillation tower

is enhances the separation. The condensate product is benchmarked based on its RVP. The RVP can be adjusted by the bottom reboiler temperature. The surrogate model will only be fitted to one output; the RVP value of the outlet gas of the stabilization unit. A detailed discussion can be found in GPSA (2004) [22].

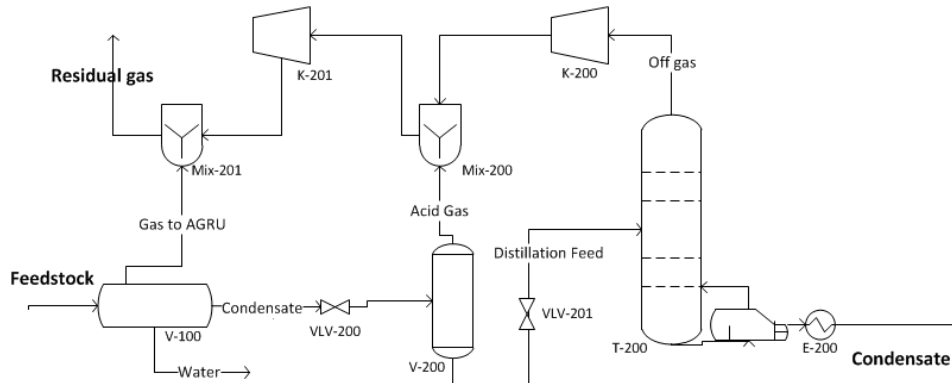


Figure 1. Process flow diagram of the gas stabilization unit.

3. Methodology

The overall methodology of this work consists of six steps that include simulation, variable, and operating range identification, data generation using sampling techniques, dimension reduction, space fitting, and deviation estimation. Figure 2 shows the sequential framework adapted to develop a surrogate model for stabilization unit.

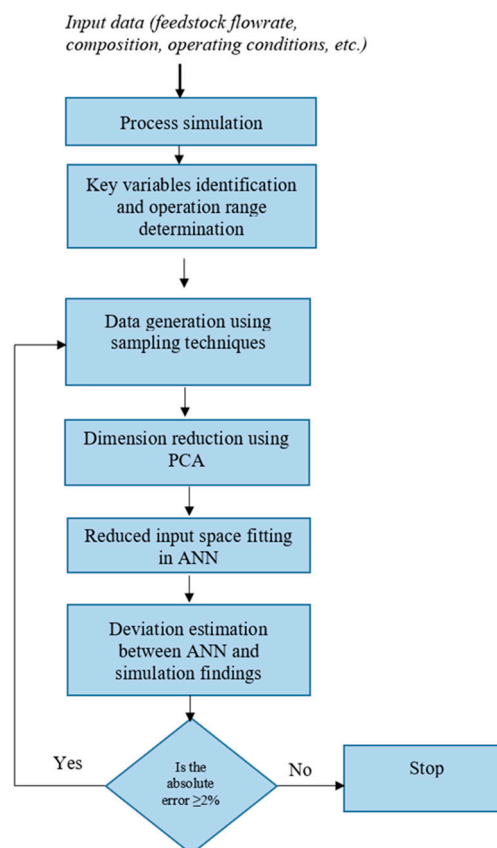


Figure 2. Adapted overall methodology.

3.1. Step One: Process Simulation

Steady state simulation of gas stabilization unit is performed using ASPEN HYSYS under different operating conditions to estimate the material and energy requirements.

3.2. Step Two: Key Variables and Range Identification

Variable selection is a crucial step and requires an understanding of the process. The usefulness of the pseudo-model and its applicability depends on the choice of an optimal set of variables that are essential and sufficient. Initial process variable selection can be made using the process knowledge. With many variables, a reduced order model can be accomplished through a data-driven approach using input–output data using partial least squares-variable importance in the projection (PLS-VIP) method [23].

3.3. Step Three: Data Generation Using Different Sample Techniques

In this step, different sampling techniques used for our analysis are presented. An overall approach using different sampling techniques is also presented. It is followed by an explanation of the dimension reduction of the input space to have less computational cost.

3.3.1. Pseudo-Random Sampling

The pseudo-random sampling method is the most basic form of the sampling technique. In this sampling technique, samples randomly over the entire space. It does not take into account the previous sampling points when it generates new points. Therefore, one can run into the risk of only a sample in one region. Homogeneous sampling over the design space is not certain, but with a sufficient amount of sampling, it may converge to a homogeneous distribution of points.

3.3.2. LHS

LHS ensures that a sample does not appear more than once in the design space. In the LHS method, for each input variable M , N_p sampling points are used. Each variable M is divided into N_p equal probable intervals. The sampling space is designed in the way that only one of every sample is the only one in its axis-aligned hyperplane containing it. The advantage of this method is that it ensures a more orderly random sampling space. Another advantage is that the sampling space does not increase when additional input variables are chosen since each variable is chosen from N_p with equal likelihood.

3.3.3. HSS

The pseudo-random number generator creates samples grouped in a specific region of the unit hypercube and is which will lead to non-uniform samples. To avoid this, pseudo-random sampling with the Monte Carlo technique requires many samples. To increase the efficiency of pseudo-random number-based Monte Carlo sampling, variance reduction techniques are used. LHS, and HSS use such techniques. HSS generates quasi-random numbers based on Hammersley sequences. Typically, the samples generated by HSS are more uniformly distributed than other samples generated by different techniques (Kalagnanam, 1997) [14]. HSS has been used for several implementations to improve the efficiency of simulations (Mukherjee and Diwekar, 2016 [15], Dige & Diwekar, 2018 [24]). In this study, we are exploring the use of LHS and quasi-random technique to demonstrate the process which cannot be done through the pseudo-random number.

3.4. Step Four: Dimension Reduction Using Principle Component Analysis (PCA) Score

PCA transforms the original input space using a linear transformation technique. It is an unsupervised method of dimension reduction. The transformed dataset with a smaller dimension (depending on the number of principal components used) contains most of the variability of the

original information. In PCA, the normalized eigenvectors (Q) of the covariance matrix of the original input space (X) are in order of reduction in variance (eigenvalues, Δ). The first eigenvector with the highest eigenvalue corresponds to a direction of maximum variance. The second eigenvector with the next highest eigenvalue corresponds to the direction of the next variance and so on. Only a small number of the eigenvectors will be chosen based on the sum of the variance or information of the input space. Generally, surrogate optimization is useful if applied to the entire data space. In cases where the input space is high dimension cases applying PCA scores seem reasonable [25]. The original variables in the lower dimensional space are represented by PCA scores (T) given as:

$$XL = T \quad (1)$$

where $L = \sqrt{\Delta}Q$ is the loading matrix.

3.5. Step Five: Surrogate Model Creation Using ANN

In this step, a brief theoretical background of ANN and how they are used in the present problem. It starts with the explanation of the ANN model structure and how surrogate models are generated using ANN in the present case.

A popular way to model a nonlinear system is through ANN-based modeling. ANNs have many highly interconnected nodes, also called artificial neurons that provide a nonlinear “black-box” system [26]. Model development using ANNs involves adapting or learning in response to variation in the environment through a training mechanism. Once a model is developed, ANNs can be re-trained to deal with minor changes in the environment. An ANN model can have a single input single output or multiple inputs and multiple output variables or a combination of both. For the present problem, multiple inputs single output model is chosen.

ANNs perform an input-output mapping of the data that can be used for multivariate pattern recognition. All possible patterns that need to be recognized should be present in the training data. The result may not be reliable if the training data is small. For that purpose, a large amount of data and rich in variation are used for training. It should be noted that due to the limitations of training data that may not be well distributed in the entire space, ANNs might be accurate in some zone and not in other. In the present case, data is generated for training purposes using both pseudo-random sampling, LHS, and quasi-Monte Carlo HSS. The HSS is expected to cover the entire range of the input variables. This is to ensure that output data is generated mainly by interpolation.

ANNs typically have multi-layered interconnected neurons that relate the input-output data in a nonlinear way. The nonlinear model has three basic attributes:

A set of connections (synaptic weights) that describes the influence of a node on nodes in the next layer. It can be positive or negative to excite or inhibit, respectively.

A summation operator of input signals weighted by the synopsis of the neuron

An activation function $\varphi(\cdot)$, limited by the amplitude of the output of the neuron. The range of the amplitude is in closed interval $[0,1]$ or $[-1,1]$. The activation function defines the output in terms of the activation potential. Typical activation function includes step function or sigmoid function.

In NN, the neuron k is related to the input and outputs through the following equations:

$$u_k = \sum_{j=0}^m w_{kj} x_j \quad (2)$$

$$v_k = u_k + b_k; b_k = w_{k0} \text{ (bias)} \quad (3)$$

$$y_k = \varphi(v_k) \quad (4)$$

where x_1, x_2, \dots, x_m are the input signals, $w_{k1}, w_{k2}, \dots, w_{km}$ are the synoptic weights of the neuron, u_k is the linear combined output of the input signals, v_k is the activation potential and y_k is the output signal of the neuron. The bias is an external parameter, which provides a transformation to the output u_k .

The activation function depends on the type of problem to be addressed and can be a threshold function, piecewise linear function, a sigmoid function, hyperbolic tangent function, or a Wavelet function. In the present work, a nonlinear function is used expecting the nonlinearity of the problem. The Levenberg-Marquardt training algorithm is used to fit the function. This algorithm generally uses more memory, but computes in less time. Training stops when generalization stops improving, as marked by an elevation in the mean square error of the validation samples. Multiple neurons can be linked with each other to form a network through a single-layer feedforward network or a multi-layer feedforward network. To keep the problem simple, we consider a single-layered feedforward network.

The neural fitting tool is used to fit the input–output data from the simulated process into a network. The data collected is being divided into three main sets; training, validation, and test set. The training set is used to fit the parameters (i.e., weights) of the classifier. The validation set is used to adjust the parameters of a classifier. Finally, the test set is used to assess the performance of a fully specified classifier. The performance of fitting is benchmarked based on the mean squared error and the regression R-value. Mean squared error (MSE) is the average squared difference between outputs and targets. The optimized ANN model is generated by minimizing the MSE between the output from the process simulator and that predicted by the ANN model.

4. Case Study: Results and Discussion

The HYSYS simulated process diagram of a gas stabilizing unit shown in Figure 3 is considered. The goal of the gas stabilization unit is to produce a condensate product with a RVP value between 10 to 12 psi. This range ensures that the light component of the stream will not be separated into a light gas phase while transporting the liquid stream. The separation of the feed stream is performed in a 3-phase separator (V-100) to remove the water and light components. The condensate is then expanded through (VLV-200) and sent to the flash drum (V-200). Acid gas is then removed in the flash drum, and liquid hydrocarbon which C_6 represents most of it with a mass fraction of approximately 62% is expanded in VLV-201. The expanded stream is sent to the column (T-200) without a reflux unit that operates at a pressure of 250 kPa. C_6 and heavy hydrocarbons are condensed in the column with a certain RVP, and light components are mixed with the acid gas outlet of the flash drum.

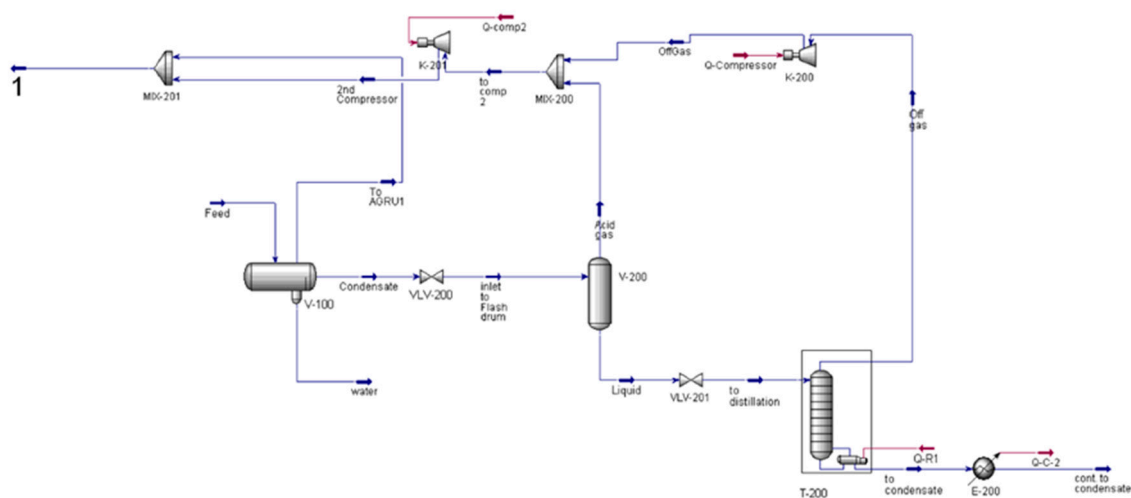


Figure 3. ASPEN HYSYS simulation of gas stabilizing unit.

The main defined variables in this flowsheet are the feed temperature, pressure, and flow rate. These variables will be altered and varied to see their effect on the RVP of the condensate. The surrogate model is generated based on the temperature and pressure of the feed. The sampling range of these

variables is shown in Table 1. These data represent different interactions between the three main variables and the resulted RVP due to the different data combinations.

Table 1. Input parameters range.

Temperature (°C)	15–30
Pressure (kPa)	1600–2500

It is worth mentioning that this range does not cover all the possible operational conditions. However, this range is taken to examine the effect of the sampling domain on the performance of this specific ANN.

4.1. Sampling

Random samples from the probability distribution of the inputs when running through the model system will create a distribution of the output. In the simplest form, the input distribution is approximated by the random samples (Monte Carlo method) with a uniform distribution $U(0,1)$ with n samples on a k dimensional unit hypercube. The numbers are random as they are generated by a specific algorithm. The uniformity properties of a sampling technique are important since the sample must approximate the entire domain by finite samples. Otherwise, larger sample sizes are needed.

For increasing the efficiency of random simulations, variance reduction techniques have been developed using LHS and quasi-Monte Carlo methods. The error of approximating the entire domain by a finite sample depends on the equidistributional properties of the sample used for $U(0,1)$ rather than on its randomness. The LHS and HSS are used in the present work for constructing uniform sequences that are typically referred to as low-discrepancy sequences. Figure 1 illustrates the procedure for evaluating each sampling technique. The main criterion is the error between ANN prediction and the actual simulation value. In the present work, only two input variables are used for model generation. Thus, dimension reduction is not considered. An error of 2% or lower is considered to be acceptable.

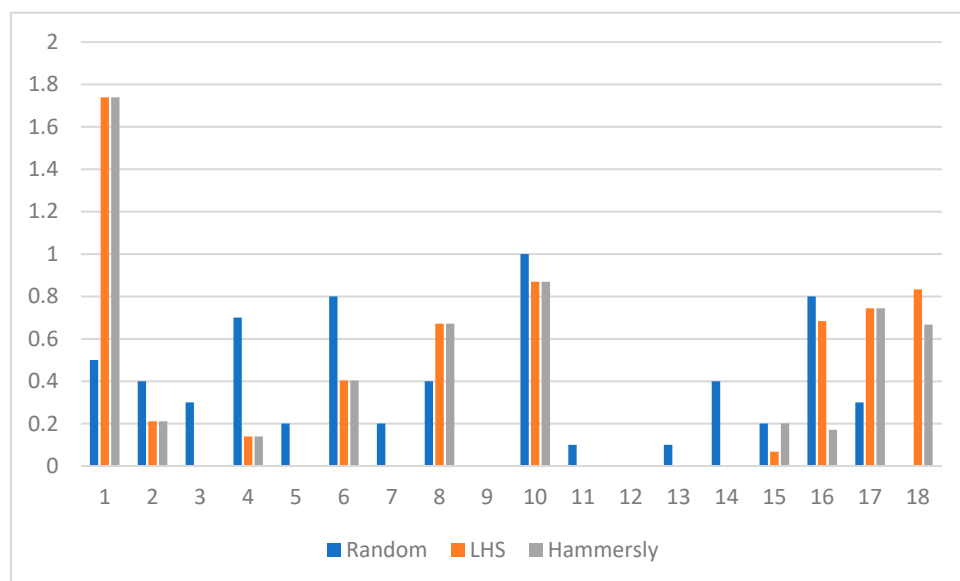
The results from different sampling techniques are illustrated in Table 2. It shows that the constructed ANN has excellent predictability when the range of the operational condition is within the same limit of the HYSYS operation condition with a maximum relative error value around 1%. However, choosing operating conditions outside the selected HYSYS range will result in a deviation of RVP values. The deviation magnitude depends on how far the selected operating conditions is from the HYSYS operation range. For example, in test number 1, the selected pressure is outside the HYSYS range, but the temperature is within the range. Therefore, the error value is acceptable.

Table 2. Results obtained from different sampling techniques.

Temperature °C	Pressure kPa	HYSYS® RVP psi	ANN RVP psi	Random Error %	LHS Error	Hammersley Error
16.1	1728.6	11.5	11.5	0.5	1.74	1.74
24.4	1737.5	14.2	14.2	0.4	0.21	0.21
24.0	1630.9	13.6	13.6	0.3	0	0
24.7	1617.6	14.4	14.5	0.7	0.14	0.14
22.0	1684.2	12.0	12.0	0.2	0	0
22.9	1622.1	12.4	12.3	0.8	0.40	0.40
19.7	1662.0	11.8	11.8	0.2	0	0
24.9	1804.1	14.9	14.8	0.4	0.67	0.67
24.0	1693.1	13.6	13.6	0	0	0
17.7	1781.9	11.5	11.6	1	0.87	0.87
17.9	1693.1	11.7	11.7	0.1	0	0
20.4	1688.7	11.9	11.9	0	0	0
20.8	1688.7	11.9	11.9	0.1	0	0
19.0	1746.4	11.7	11.7	0.4	0	0
24.9	1768.6	14.9	14.8	0.2	0.07	0.20
16.8	1630.9	11.7	11.6	0.8	0.68	0.17
22.4	1741.9	12.1	12.1	0.3	0.74	0.74
22.0	1715.3	12.0	12.0	0	0.83	0.67

4.2. Maximum Absolute Error

The maximum error on the performance for different sampling methods, sizes, and various combinations of the independent variables was calculated. Figure 4 shows a different maximum error in percentage for the different neural network fittings. In 4 cases (test numbers 1, 8, 17, and 18) we found that the sampling using pseudo-random number has a lower maximum error when compared to LHS and HSS. We found that LHS and HSS in 12 other cases (test numbers 2, 3, 4, 5, 6, 7, 10, 11, 13, 14, 15, and 16) have better performance. In 2 other cases (test numbers 9 and 12), the three sampling methods performed equally well with no error. Thus, in most of the cases, the random sampling method showed significantly higher maximum error than LHS and HSS. The maximum absolute error did not show any sign of decreasing when the sampling space increased.

**Figure 4.** The differences in max. Absolute error and the standard deviation for sampling.

5. Conclusions

LHS and HSS is a better alternative than pseudo-random sampling for the performance of the surrogate model. When compared between LHS and HSS, HSS is found to be the best. This can be seen in the differences in maximum absolute error and the standard deviation. HSS ensures that the whole sampling space is covered and therefore has the best performance of the three different methods that were compared.

The surrogate model developed in the present case is found to be robust enough to predict process output with error as low as 0% even when the input temperature parameter is out of range. The pseudo-random sampling has relatively high standard deviations and maximum absolute error as found from most of the test cases. The sampling size will become too big for validation and the computational time for generating the surrogate model will be time-consuming. HSS, because of n -dimensional uniformity, shows a lower absolute error. Future work is to study how the performance of the surrogate model change when the sampling space in the most important variables for the outlet pressure increases. Different combinations of the number of neurons in each hidden layer for the neural network can give a different performance on the surrogate model and should be investigated.

Author Contributions: The following statements should be used “Conceptualization, S.A. and R.M.; Methodology, S.A., M.I. and R.M.; Software, M.I. and A.A.; Validation, M.I. and A.A.; Formal Analysis, S.A., M.I., R.M. and A.A.; Investigation, S.A., M.I., and R.M.; Resources, M.I.; Data Curation, M.I. and S.A.; Writing-Original Draft Preparation, S.A., M.I. and R.M.; Writing-Review & Editing, S.A. and R.M.; Visualization, M.I. and R.M.; Supervision, S.A.; Project Administration, S.A.; Funding Acquisition, M.I.”

Funding: The authors would like to acknowledge the financial support received from Qatar University in the form of graduate assistance fellowship.

Acknowledgments: The authors would like to thank Prof Mahmoud El-Halwagi from Texas A&M University, college station, the USA for sharing his valuable comments on the work. The publication of this article was funded by the Qatar National Library.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

ANN	Artificial Neural Network
GA	Genetic Algorithm
AGRU	Acid Gas Recovery Unit
RVP	Reid Vapor Pressure
LHS	Latin Hypercube Sampling
RND	Random Sampling
x_j	input signals
w_{kj}	synoptic weights of neuron
u_k	linear combined output of the input signals
y_k	output signal of the neuron
φ	activation function
v_k	activation potential
VLV	Valve
MIX	Mixer
E	Heat Exchanger
T	Distillation Column
K	Compressor
V	Flash Drum

References

1. Straus, J.; Skogestad, S. Minimizing the Complexity of Surrogate Models for Optimization. *Comput. Aided Chem. Eng.* **2016**, *38*, 289–294.
2. Biegler, L.T.; Grossmann, I.E.; Westerberg, A.W. *Systematic Methods for Chemical Process Design*; Prentice Hall: Upper Saddle River, NJ, USA, 1997.
3. Shokry, A.; Bojarski, A.D.; Espuna, A. Using surrogate models for process design and optimization. In *Uncertainty Modeling in Knowledge Engineering and Decision Making*; 2012; pp. 483–488.
4. Basheer, I.A.; Hajmeer, M. Artificial Neural Networks: Fundamentals, Computing, Design, and Application. *J. Microbiol. Methods* **2000**, *43*, 3–31. [[CrossRef](#)]
5. Ochoa-Estopier, L.M.; Jobson, M.; Smith, R. Operational Optimization of Crude Oil Distillation Systems using Artificial Neural Networks. *Comput. Chem. Eng.* **2013**, *59*, 178–185. [[CrossRef](#)]
6. Liao, L.C.-K.; Yang, T.C.-K.; Tsai, M.-T. Expert System of a Crude Oil Distillation Unit for Process Optimization using Neural Networks. *Expert Syst. Appl.* **2004**, *26*, 247–255. [[CrossRef](#)]
7. Motlaghi, S.; Jalali, F.; Ahmadabadi, M.N. An Expert System Design for a Crude Oil Distillation Column with the Neural Networks Model and the Process Optimization using Genetic Algorithm Framework. *Expert Syst. Appl.* **2008**, *35*, 1540–1545. [[CrossRef](#)]
8. Osuolale, F.N.; Zhang, J. Multi-Objective Optimisation of Atmospheric Crude Distillation System Operations Based on Bootstrap Aggregated Neural Network Models. *Comput. Aided Chem. Eng.* **2015**, *37*, 671–676.
9. Arce-Medina, E.; Paz-Paredes, J.I. Artificial Neural Network Modeling Techniques Applied to the Hydrodesulfurization Process. *Math. Comput. Model.* **2009**, *49*, 207–214. [[CrossRef](#)]
10. Eason, J.; Cremaschi, S. Adaptive Sequential Sampling for Surrogate Model Generation with Artificial Neural Networks. *Comput. Chem. Eng.* **2014**, *68*, 220–232. [[CrossRef](#)]
11. McKay, M.D.; Beckman, R.J.; Conover, W.J. Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics* **1979**, *21*, 239–245.
12. Iman, R.L.; Shortencarier, M.J. *A Fortran-77 Program and User's Guide for Generation of Latin Hypercube and Random Samples for Use with Computer Models*; Sandia National Laboratories: Albuquerque, NM, USA, 1984.
13. Stein, M. Large Sample Properties of Simulations using Latin Hypercube Sampling. *Technometrics* **1987**, *29*, 143–151. [[CrossRef](#)]
14. Kalagnanam, J.R.; Diwekar, U.M. An Efficient Sampling Technique for Off-Line Quality Control. *Technometrics* **1997**, *39*, 308–319. [[CrossRef](#)]
15. Mukherjee, R.; Diwekar, U.M. Comparison of Monte Carlo and Quasi-Monte Carlo Technique in Structure and Relaxing Dynamics of Polymer in Dilute Solution. *Comput. Chem. Eng.* **2016**, *84*, 28–35. [[CrossRef](#)]
16. McBride, K.; Sundmacher, K. Overview of Surrogate Modeling in Chemical Process Engineering. *Chem. Ing. Tech.* **2019**, *91*, 228–239. [[CrossRef](#)]
17. Bian, S.; Henson, M.A.; Bélanger, P.; Megan, L. Nonlinear State Estimation and Model Predictive Control of Nitrogen Purification Columns. *Ind. Eng. Chem. Res.* **2005**, *44*, 153–167. [[CrossRef](#)]
18. Henao, C.A.; Maravelias, C.T. Surrogate-Based Superstructure Optimization Framework. *AIChE J.* **2011**, *57*, 1216–1232. [[CrossRef](#)]
19. Carrero, E.; Queipo, N.V.; Pintos, S.; Zerpa, L.E. Global Sensitivity Analysis of Alkali-Surfactant-Polymer Enhanced Oil Recovery Processes. *J. Pet. Sci. Eng.* **2007**, *58*, 30–42. [[CrossRef](#)]
20. Mukherjee, R.; El-Halwagi, M.M. Reliability of C-H-O Symbiosis Networks under Source Streams Uncertainty. *Smart Sustain. Manuf. Syst.* **2018**, *2*, 132–153. [[CrossRef](#)]
21. Mokhatab, S.; Poe, W.A. *Handbook of Natural Gas Transmission and Processing*; Gulf Professional Publishing: Houston, TX, USA, 2012.
22. Gas Processors Suppliers Association. *Engineering Data Book*; Gas Processors Suppliers Association: Tulsa, Oklahoma, 2004; Volume 2, pp. 16–24.
23. Mukherjee, R.; Sengupta, D.; Sikdar, S.K. Parsimonious Use of Indicators for Evaluating Sustainability Systems with Multivariate Statistical Analyses. *Clean Technol. Environ. Policy* **2013**, *15*, 699–706. [[CrossRef](#)]
24. Dige, N.; Diwekar, U. Efficient Sampling Algorithm for Large-Scale Optimization under Uncertainty Problems. *Comput. Chem. Eng.* **2018**, *115*, 431–454. [[CrossRef](#)]

25. Kamali, M.; Ponnambalam, K.; Soulis, E.D. Integration of Surrogate Optimization and PCA for Calibration of Hydrologic Models, A WATCLASS Case Study. In Proceedings of the 2007 IEEE International Conference on Systems, Man and Cybernetics, Montreal, QC, Canada, 7–10 October 2007; pp. 2733–2737.
26. Kayihan, F.; Cinar, A.; Palazoglu, A. *Chemical Process Performance Evaluation*; CRC Press: Boca Raton, FL, USA, 2007; Volume 20070111, pp. 7–36.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).