NTNU

Norwegian University of
Science and Technology

Surrogate Model Development for an Ammonia Synthesis Process

Kun Wang

Chemical Engineering
Submission date: June 2017
Supervisor: Sigurd Skogestad, IKP
Co-supervisor: Julian Straus, IKP

Norwegian University of Science and Technology
Department of Chemical Engineering

NTNU – Trondheim
Norwegian University of
Science and Technology

Surrogate Model Development for an Ammonia Synthesis Process

Kun Wang

June 2017

MASTER THESIS

Department of Chemical Engineering

Norwegian University of Science and Technology

Supervisor: Sigurd Skogestad

Co-supervisor: Julian Straus

Abstract

The existing ammonia synthesis process modeled in HYSYS is aimed to be optimized. But it is not possible to implement general optimization solvers to directly optimize the HYSYS model, since the derivative information is not available in HYSYS and the estimation of derivatives is difficult to achieve. Hence the surrogate model is introduced to replace the original HYSYS model to make the optimization feasible.

The ammonia process consists of four interconnected sections, the makeup section, the reaction section, the separation section and the refrigeration section. We aim to construct a surrogate model for each section and then combine all the models to obtain the complete model for the whole process. Then the process can be optimized based on the new surrogate model. The objective of this thesis is to construct the surrogate model for the separation-refrigeration section.

First, variable analysis including variable classification, variable reduction and variable identification is performed to determine the input and output variables for surrogate model generation. The input variables are then sampled by adaptive sampling to obtain the input sample space with minimized number of sample points. Next the input sample data are imported in the HYSYS model so that the output sample data can be calculated. Then based on the input and output sample data, the surrogate model is generated using artificial neural network. Finally, The resultant surrogate model is validated by comparing its output prediction with HYSYS' output results.

Acknowledgment

This thesis has been written as the conclusion of the M.Sc degree in Chemical Engineering from the Norwegian University of Science and Technology. The work has been completed at the group of Process Systems Engineering.

I would like to thank my main supervisor, Sigurd Skogestad, for the possibility for me to write this thesis on an interesting and challenging subject at his research group. I am grateful for the knowledge that has been instilled in my mind through working with the professors, PhD candidates and fellow students on the 2nd floor of Chemistry Block IV.

This whole project would not have been possible without the guidance and motivation from my co-supervisor, the PhD candidate Julian Straus. I hope that this work will prove to be of benefit for his doctoral work and I cannot express my gratitude enough for the effort he has spent on me.

Statement of Compliance

I declare that this is an independent work according to the exam regulations of the Norwegian University of Science and Technology.

Trondheim, 30th June 2017

Kun Wang

Summary and Conclusions

In this thesis, a surrogate model is aimed to be constructed for the separation-refrigeration (S-R) section in the ammonia plant. First, the variables defining the process are classified by input variables and output variables. Then the variables are reduced as many as possible using the dependency relationships. After the input and output variables are determined by variable identification, adaptive sampling is implemented to sample the input variables to obtain the input sample space with minimized number of sample points. The resultant input sample data is imported to HYSYS to obtain the corresponding output sample data. However, the HYSYS model was not able to calculate the corresponding output samples due to convergence issues. In order to address this issue, we divided the separation-refrigeration section furthermore into the HEx part, the separator part and the refrigeration section. Then we use the same approach to construct a surrogate model for the HEx part. We used two different variable identifications to define the output variables, of which one used the absolute output variables and the other one used variables differences. Sample spaces of both cases are obtained and surrogate models are generated using artificial neural network based on the sample spaces. The resultant surrogate models are validated by comparing their output predictions with the HYSYS' output results and the relative deviations are calculated.

It can be concluded that the surrogate models can be efficiently constructed based on the approach used in this thesis. Using the adaptive sampling, the number of sample points required to generate surrogate models can be successfully minimized, which saved computational expense of model construction effectively. In the approach of surrogate model construction, the variable identification is extremely crucial. It can affect both the computational expense

of surrogate model generation and the accuracy of resultant surrogate models. Generally, the variable differences can save the computational expense but results in less accurate surrogate models than absolute variables. The approach used in this thesis can also be implemented to construct the surrogate models for other simulators.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

An ammonia synthesis process modeled in HYSYS is aimed to be optimized. However, it is difficult to directly optimize the HYSYS model since the HYSYS simulator is essentially a black-box model, where the derivatives are difficult to obtain or estimated, leading to infeasibility of implementation of general optimization solvers [6]. To address this issue, the idea is firstly dividing the HYSYS model into four sections, the makeup section, the reaction section, the refrigeration section and the separation section. Then We build surrogate models for different sections and combine all the surrogate models together to form the complete model of the whole process. Hence the resultant complete model of the process can be optimized using general optimization solvers. The objective of this thesis is to construct a surrogate model for the separation section.

1.1 Ammonia Synthesis Process

The ammonia synthesis process modeled in HYSYS is an integrated plant which can be viewed as four interconnected sections, the makeup section, the reaction section, the separation section and the refrigeration section. The raw material

in the feed to the plant is typical syngas consisting of N_2 and H_2 with a molar ratio of 3:1, and other impurities including He, Ar, H_2O, CH_4 and CO_2. The feed stream to the plant firstly enters the makeup section, where the fresh syngas is compressed by multiple compressors and washed with liquid ammonia to remove the water in the feed. Then the syngas enters the reaction section where it is heated by integrated heat exchanger network and feed to the reactor beds for reactions. Next, the product stream from the reaction section passes to the separation section where it is cooled down and separated. The liquid stream leaving the separation section, mainly consisting of ammonia, then enters the refrigeration section, where the remaining impurities of water and some dissolved gases in the liquid stream are removed. Finally, the liquid ammonia leaving the refrigeration section is the main product. Other by-products are stored or used as fuel gas. A detailed process description is illustrated in the Chapter 2.

1.2 Optimization of HYSYS with Surrogate Models

Most commercial steady-state simulators, like Aspen HYSYS, SimSci PRO/II, or UniSim Design Suite utilize the sequential-modular approach to solve the flowsheet, where each unit operation is solved sequentially [7]. The sequential methods are based on the natural hierarchy of the chemical process flowsheet [2], of which the main advantages are the ability to solve a problem in a intuitive way and achieve straightforward process design [8].

However, problems arise when the process modeled in these simulators are aimed to be optimized. The optimization of general models usually utilize derivative-based or algebraic solvers. But in these simulators, the input and output are calculated by black-box models where the model functions can not be available directly in the desired algebraic form, and hence the derivative information is not

possible to obtain or it is difficult to estimate the derivatives. This leads to infeasibility of implementation of optimization solvers to optimize the processes in these simulators. In order to overcome this issue, the surrogate models have been constructed to substitute the original black-box models, making it feasible to implement the general solvers for optimization [9, 10].

1.3 Scope of Work

The objective of this thesis is to construct a surrogate model for the separation sections in the ammonia process modeled in HYSYS. The resultant surrogate model will be validated by comparing their output prediction with the original HYSYS' output.

1.4 Structure of the Thesis

This thesis consists of 6 chapters and 3 appendices.

Chapter 1. Gives a brief introduction to the ammonia process and the motivation for constructing surrogate models of the HYSYS model for optimization purpose.

Chapter 2. Describes the whole ammonia process in details and the process flow diagrams for all the sections are present .

Chapter 3. Discusses about general surrogate modeling techniques, the variable analysis and sampling methods.

Chapter 4. Constructs the surrogate models for two sections in the process.

Chapter 5. Validates the resulting surrogate models and discusses about the results.

Chapter 6. Summarizes the work and gives recommendations for future work.

Appendix A. Shows the nominal conditions and variation ranges of corresponding variables.

Appendix B. Shows the process flow diagrams of the separation-refrigeration section.

Appendix C. Shows the various variable defined in different sections.

Chapter 2

Process Description

This chapter gives descriptions of the ammonia synthesis process modeled in HYSYS. An overview of the process is illustrated first, and then each individual section in the process is described.

2.1 Overview of the Ammonia Plant

The objective process is an ammonia plant modeled in HYSYS. The whole plant is an integrated plant that can be viewed as four interconnected sections, the makeup section, the reaction section, the separation section and the refrigeration section. The configuration of these sections are visualized in the block diagram in Figure 2.1, where the blocks represent different sections, the solid lines represent the mass flows and the dash lines represent the energy flows. The raw material feed to the plant is typical syngas consisting of N_2 and H_2 with a molar ratio of 3:1 and other impurities including He, Ar, H_2O, CH_4 and CO_2. The main product of the plant is NH_3 and other by-products are stored, purged or used as fuel gas.

The ammonia process is described briefly by individual sections with corre-

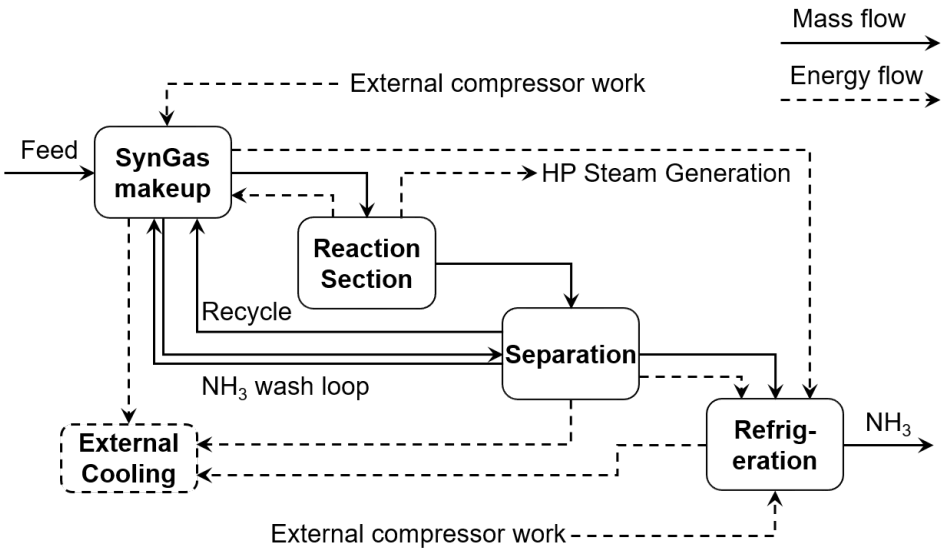


Figure 2.1: Block diagram of main processes in the ammonia plant

sponding process flow diagrams (PFD) in the following sections. In the PFD figures, the red streams note the inlet streams to the corresponding sections while the blue streams note the outlet streams of the corresponding sections. Besides, the green streams represent the flows leaving the plant.

2.2 Makeup Section

The syngas feed firstly enters the makeup section in the plant. The incoming syngas is compressed by the compressor C1-BP to make the pressure match the requirement. After the pressure is increased, the syngas is cooled down by both cooling water through CW-1 and the ammonia flow from the refrigeration section through HEx-1. Because of the cooling, some water condenses in the flow and it is then removed by the separator S-1. Next, the remaining H_2O and CO_2 in the gas flow out of S-3, which could damage the catalyst within the reactor, are removed through washing with liquid NH_3. Then the gaseous flow with syn-

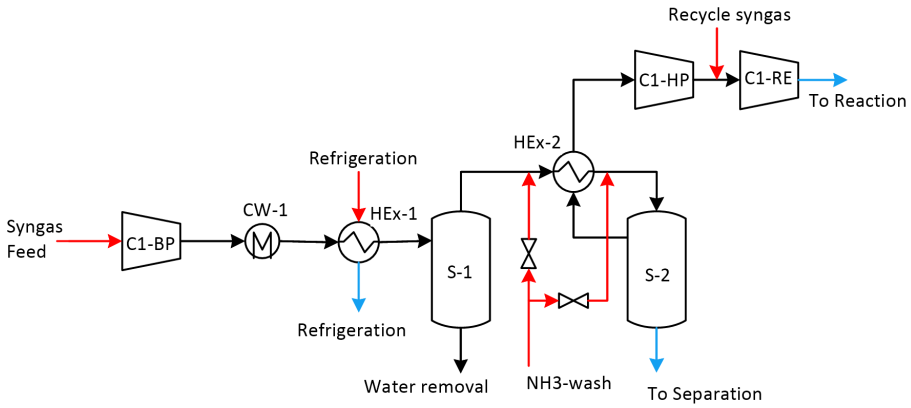


Figure 2.2: The PFD for the makeup section of the ammonia plant

gas and NH_3 is cooled down by HEx-2 and becomes a two-phase stream due to addition of liquid NH_3. The flow is then separated by the separator S-2 where H_2O, CO_2 and some other gases are dissolved in the liquid ammonia. The liquid stream leaves the separator S-2 with most of NH_3 and passes to the separation section. The gas flow out of the separator S-2 is then provided as cooling media in HEx-2 and afterwards compressed by compressor C1-HP to match the pressure of the recycle syngas from the separation section. After mixed with the recycle syngas, the flow is compressed again by C1-RE and then feed to the reaction section. The PFD for the syngas makeup section is presented in Figure 2.2.

2.3 Reaction Section

The syngas stream out of the makeup section enters the reaction section and is firstly heated through the heat exchanger HEx-3 as shown in the PFD in Figure 2.3. Then the syngas stream is split into three substreams, of which two are heated by heat exchangers IEX and HEx-4. After heated, the three substreams

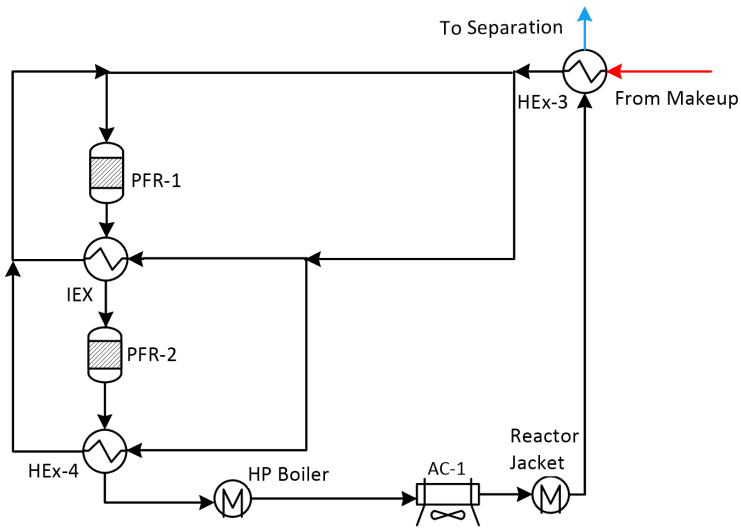


Figure 2.3: The PFD for the reaction section of the ammonia plant

are mixed and feed to the reactor, which is a jacked two-bed tubular reactor. The feed enters the reactor bed PFR-1 and the outlet flow of PFR-1 is then cooled down by quenching through IEX. Next, the gas flow enters the reactor bed PFR-2 which produces the product stream. The product stream leaving the reactor is cooled down through HEx-4 with quenching and then cooled down again by providing heat for the HP-steam boiler. Afterwards, the product stream is further cooled by the air cooler AC-1. The cooled product stream passes to the reactor jacket to cool the reactor beds before it heats the syngas feed stream through HEx-3. After HEx-3, the product stream leaves the reaction section and is then feed to the separation section.

2.4 Separation Section

The stream from the reaction section entering the separation section is firstly cooled down with cooling water through CW-1. Then the stream is split into two

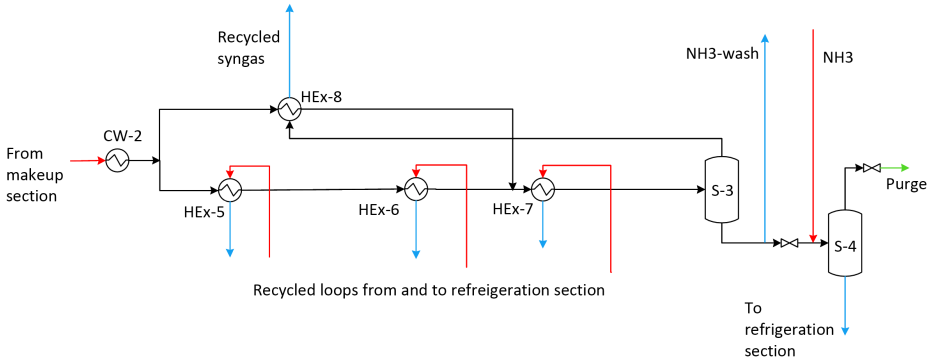


Figure 2.4: The PFD for the separation section of the ammonia plant

substreams as shown in the PFD in Figure 2.4. The upper substream is cooled by the recycle syngas stream, which returns to the makeup section, through the heat exchanger HEx-8. The lower substream is cooled by two heat exchangers HEx-5 and HEx-6, then mixed with the upper substream and further cooled by HEx-7. All the cooling media is provided by the streams from the refrigeration section containing mainly liquid ammonia. After cooling, the stream becomes a two-phase flow which is then separated in the separator S-3. The gas phase from S-3 passes to HEx-8 to cool down the upper substream and then returns to the makeup section as recycle syngas flow. A small part of the liquid stream leaving S-3 passes to the separation section for water washing. The majority of the liquid stream is mixed with the liquid ammonia flow from the makeup section and then passes to the separator S-4. The gas flow out of S-4 is purged and the liquid flow is then feed to the refrigeration section.

2.5 Refrigeration Section

The liquid stream, mainly consisting of ammonia, from the separation section is feed to the refrigeration section and firstly enters the separator S-5, as shown

in the PFD in Figure 2.5. Part of the liquid stream out of S-5 passes to the heat exchanger HEx-7, provided as the cooling media, and the other part passes on to storage. The pattern of the cooling cycle involving S-5/HEx-7 are repeated for S-6/HEx-6 and S-7/HEx-5/HEx-1. The gas flow from the top of S-5 is compressed by the compressor C2-1 at the first compression stage, then mixed with the gas flow from S-6 and compressed by C2-2 at the second compression stage. After, the resulting gas stream is mixed with the gas flow from S-7 and the mixed flow is compressed by the compressor C2-3 at the third compression stage. To be noticeable, the three compressors C2-1, C2-2 and C2-3 are driven by the same shaft so that they have the same revolution speed. The three-stage compression leads to a significant temperature increase which is partly reduced by inter-stage cooling through the cooler CW-3. After the three-stage compression, the gas flow is further cooled by the air cooler AC-2 and cooling water through CW-4, leading to the majority of flow condensing in the stream. The two-phase flow then passes to the separator S-8. The uncondensed ammonia out of S-8 is furthermore cooled by the HEx-9 and separated again by S-9 where the gas flow leaves the plant as fuel gas. The liquid stream out of S-8 is partly recycled to S-7 and the remaining ammonia leaves the plant as the main product.

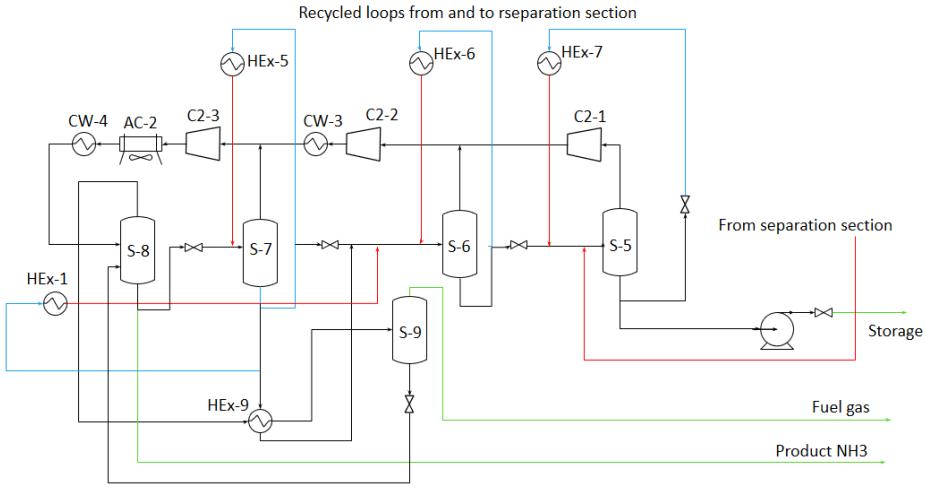


Figure 2.5: The PFD for the refrigeration section of the ammonia plant

Chapter 3

Surrogate Model Techniques and Sampling Methods

Since it is difficult to optimize the objective ammonia process directly in the HYSYS model, the surrogate modeling is introduced to address the infeasibility problem of optimizing the HYSYS model. Some typical surrogate modeling techniques are introduced and compared in this chapter. In order to generate the surrogate model, the input and output variables need to be determined primarily. The variables are determined by variable analysis including variable classification, variable reduction and variable identification. After the variables are determined, sample spaces are designed for the variables using adaptive sampling. The adaptive sampling is explained in this chapter and other general sampling methods are introduced and discussed briefly.

3.1 Surrogate Modeling and Optimization

A general optimization problem can be expressed as [6].

$$\min f(x)$$
$$\text{s.t. } g(x) \leq 0$$
$$x \in \mathbf{A} \subset \mathbb{R}^n$$

where the cost function $f(x)$ is aimed to be minimized, with respect to the variables x. At the same time, the variables x are required to fulfill the constrains $g(x) \leq 0$ as well as the box constraint $\mathbf{A}$ which contains the upper and lower bounds.

The optimization of this kind of problems usually utilize derivative-based or algebraic solvers (e.g., CONOPT [11], IPOPT [12], and SNOPT [13]) where the information of derivatives and the algebraic form of objective functions are required. But for some simulators like HYSYS, where for any given x, the corresponding $f(x)$ and $g(x)$ are calculated by an input-output black-box model, the functions $f(x)$ and $g(x)$ can not be available directly in the desired algebraic form, and hence the derivative information is not possible to obtain or it may take costly computational expense to estimate the derivatives. This issue leads to infeasibility of implementing the optimization solvers to optimize the simulators like HYSYS.

One alternative approach to overcome this issue is using derivative-free optimization algorithms, which are utilized to solve the optimization problems where derivatives are unavailable or expensive to estimate [14, 15]. These solvers can use a minimal number of black-box model calls to locate an optimal feasible point [6]. Even though the derivative-free optimization methods have the ability to deal with the black-box models, they have difficulties to find the optimal solutions when the number of the variables x is more than 10 [14].

Therefore, in order to make the optimization of the models with numerous variables x feasible, the black-box models as the objective functions are approximated before optimization. Significant work has been done to construct the

surrogate model to substitute the original black-box models for optimization purpose [9, 10].

In this project, for the purpose of optimizing the ammonia plant modeled in HYSYS, we firstly decompose the whole process into different sections and then construct surrogate models for individual sections. Then the surrogate models can be combined together to form the complete surrogate model of the whole process and hence the whole process can be optimized using general optimization solvers. The surrogate models of individual sections are connected in the pattern as the output variables $\mathbf{Y}_{s,n-1}$ of the upstream sections are the input variables $\mathbf{U}_{s,n}$ of the downstream sections. An example of a possible connection configuration of surrogate models is illustrated in the block diagram in Figure 3.1.

As mentioned in Chapter 2, the separation section and the refrigeration section are coupled with each other by three recycle loops so we consider the combined separation-refrigeration (S-R) section as one single section. Hence we divide the whole process into three sections as the makeup section, the reaction section and the separation-refrigeration section. The PFD of the S-R section is shown in the Figure B.1, where the sections are represented simply by grey shadow blocks.

3.2 Introduction to Surrogate Modeling

Nowadays, there exists much engineering analysis based on specialized mathematical models consisting of running complex computer codes or other simulators [16, 1]. However, these computer codes or simulators are usually computationally expensive, making it difficult to handle for analysis or further optimization [17, 16]. The surrogate model is proposed to address this issue.

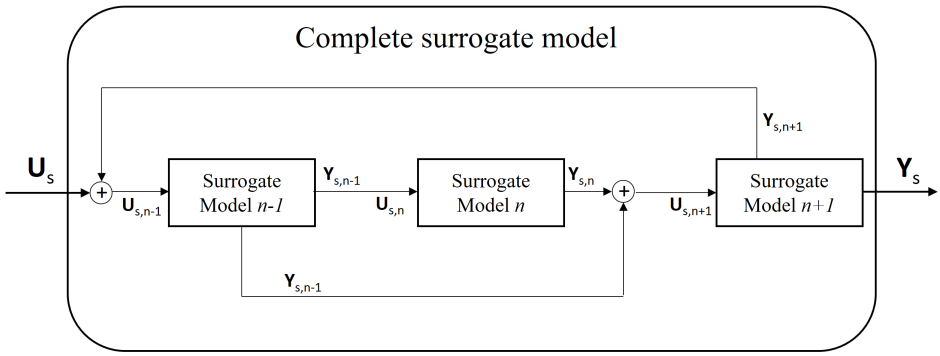


Figure 3.1: A possible connection configuration of surrogate models

The surrogate model, also know as metamodel [18], is an approximation model of the original specialized models. The surrogate model is built using some statistical techniques based on the available data, so that the surrogate model has the ability to replace the actual computer codes or simulators and predict the results without the original models [19]. The resulting surrogate model is also aimed to be many orders of magnitude faster than the original model and meanwhile it has a satisfied accuracy when predicting results.

Denoting $\mathbf{x}$ and $\mathbf{y}$ as the vectors of the input variables, also known as design variables, and the corresponding output variables also known as response variables, of the original model respectively [17]. If the true nature of the original model is [1]

$$\mathbf{y} = f(\mathbf{x}) \tag{3.1}$$

then a surrogate model can be denoted as

$$\hat{\mathbf{y}} = g(\mathbf{x}) \tag{3.2}$$

hence

$$\mathbf{y} = \hat{\mathbf{y}} + \epsilon \tag{3.3}$$

where ϵ represents the errors due to approximation.

In essence, surrogate modeling is an approach which aims to determine a function g of a set of input variables $\mathbf{x}$ from a limited amount of sample data obtained from the original model $\mathbf{y} = f(\mathbf{x})$ [17]. Hence the resulting surrogate model can replace the existing computer codes or simulators, providing a good understanding of the relationship between $\mathbf{x}$ and $\mathbf{y}$ and faster analysis tools for optimization and other explorations of the original models [1].

In order to design the sample data space that is used to construct the surrogate models, the design of experiments (DOE) [1] is applied to identify a efficient set of runs $(\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_n})$ for computer codes or simulators to obtain the corresponding output data $(\mathbf{y_1}, \mathbf{y_2}, ..., \mathbf{y_n})$. The design of experiment mainly concerns identification of the variable design space and the sampling plan to obtain the sample points in the variable design space. Hence the key issue in this part is what variables are required to determine the design space and how to get the sample data of design space as well as how to assess the goodness of such designs since the number of samples is severely limited due to expensive computation cost of the original models [17]. As the computation of each sample is expensive, the variables are aimed to be reduced as many as possible to obtain a reduced design space. Based on the reduced design space, multiple sampling methods can be utilized to design the sample space, such as complete sampling [20], simple random sampling [21, 22], Latin hypercube sampling [22, 23] and orthogonal Latin hypercube sampling [24, 25, 26, 27]. Based on specific sampling methods, the sample space used to construct the surrogate model can be determined. More detailed discussions are presented in Section 3.4.1 and Sec-

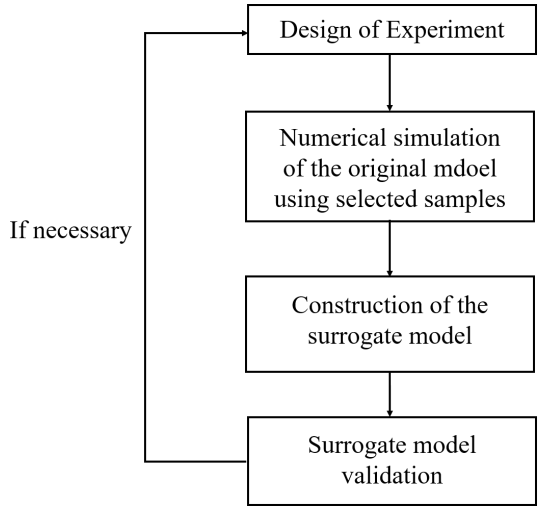


Figure 3.2: Key stages of the surrogate modeling approach [1]

tion 3.4.2.

Based on the determined sample space, a variety of approximation modeling methods can be implemented to construct the surrogate models, of which three widely used ones are Polynomial Regression [28, 29], Kriging modeling [30, 16, 31] and Artificial Neural network [32, 2, 33]. An illustration of these methods is presented in Section 3.3.

After constructing the surrogate model, it should be validated by evaluating its predictive capabilities using new data space, which is different from the sample data used to build the surrogate model. If the errors are acceptable, the surrogate model is considered accurate enough to replace the original model for further optimization or other exploration usage. The key procedures of the surrogate modeling is illustrated in Figure 3.2.

3.3 Surrogate Modeling Techniques

There is a variety of techniques to build the surrogate models by fitting data. These techniques can be categorized in two classes as parametric and non-parametric approaches for constructing surrogate models [17]. The parametric approaches specify the global functional form of the relationship between the output variables and the input variables, such as Polynomial Regression [28, 29] and Kriging modeling [30, 16, 31]. The non-parametric approaches use different types of simple, local models in different regions of the sample data to build up the overall model [17], such as Radial Basis Functions [34, 31] and Artificial Neural network [32, 2, 33]. Other statistical techniques such as Multivariate Adaptive Regression Splines [35] can be implemented using either parametric or non-parametric modeling approach. In this section, we illustrates three widely used surrogate modeling techniques, known as Polynomial Regression, Kriging method and Artificial Neural Network.

3.3.1 Polynomial Regression

The polynomial regression has been widely applied to design the models for statistics or engineering systems by a number of researchers [3, 36, 37]. The first-order polynominal used for low curvature can be expressed as in Equation (3.4a); the second-order polynominal including all two-factor interactions is expressed as in Equation (3.4b) [1, 3, 17].

$$\hat{\mathbf{y}} = \beta_0 \sum_{i=1}^{m} \beta_i x_i \tag{3.4a}$$

$$\hat{\mathbf{y}} = \beta_0 + \sum_{i=1}^{m} \beta_i x_i + \sum_{i=1}^{m} \beta_{ii} x_i^2 + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} \beta_{ij} x_i x_j \tag{3.4b}$$

Where $\hat{\mathbf{y}}$ is the set of output variables and x_i is the set of input variables; β_i is the

linear affect parameter and β_{ii} is the quadratic affect parameter. The regression parameters β are usually calculated using least squares regression analysis by fitting the output data to input data [1].

The application of polynomial regression model has an advantage of fast convergence because of its smoothing capability [38]. However, the drawback of instabilities could arise when using this method to model highly nonlinear behaviours though high-order polynomials can be utilized [39]. Besides, in high-dimension system, it may be too difficult to calculate all the parameters in the polynomial equation because it requires too much sample data for fitting the model [3].

3.3.2 Kriging Method

The Kriging model is a combination of a polynomial function and departures of the form [3]

$$y(\mathbf{x}) = \sum_{j=1}^{k} \beta_j f_j(\mathbf{x}) + Z(\mathbf{x}) \tag{3.5}$$

where $y(\mathbf{x})$ is the unknown function of interest, $f_j(\mathbf{x})$ is a known fixed function and β_j is the corresponding parameter of the function. And $Z(\mathbf{x})$ is assumed to be the realization of a normally distributed Gaussian random process with mean zero variance σ^2, and non-zero covariance. The term $\sum_{j=1}^{k} \beta_j f_j(\mathbf{x})$ approximates the design space globally while $Z(\mathbf{x})$ makes up the local deviations so that the kriging model is able to interpolate the whole design space. The covariance matrix of $Z(\mathbf{x})$ is given by [1]

$$cov[Z(\mathbf{x}^i), Z(\mathbf{x}^j)] = \sigma^2 R(\mathbf{x}^i, \mathbf{x}^j) \tag{3.6}$$

where $R(\mathbf{x}^i, \mathbf{x}^j)$ is the correlation function between any two of the n_s sample points, $\mathbf{x}^i$ and $\mathbf{x}^j$, in the data space. The function of $R(\mathbf{x}^i, \mathbf{x}^j)$ can be chosen

from a variety of correlation functions. A most widely used one is the Gaussian correlation function such as [1]

$$R(\mathbf{x}^i, \mathbf{x}^j) = exp[\sum_{k=1}^{n_s} \theta_k \mid \mathbf{x}_k^i - \mathbf{x}_k^j \mid^2] \tag{3.7}$$

where θ_k are the unknown correlation parameters that are used to fit the model, one θ_k for each of the k dimensions in the design space, and the x_k^i and x_k^j are the kth components of the two sample points $\mathbf{x}^i$ and $\mathbf{x}^j$. The specifics for fitting a kriging model is elaborated by Simpson *et al.* [30].

The Kriging model has an advantage of extreme flexibility due to the wide range of the correlation functions [3]. It can also provide a good basis for a stepwise algorithm which is able to determine the important factors, and the same data can be used to build the predictor model [40]. The main disadvantage of the kriging method is that it may be computationally expensive to fit a surrogate model in the case of a large sample space, as the $theta$ parameters used for model fitting are determined by solving a k-dimensional optimization problem which requires significant computational time when the sample data space is too large [3]. In addition, the correlation matrix **R** could be singular due to particular distribution of the sample points [3].

3.3.3 Artificial Neural Network (ANN)

The artificial neural network, also called neural network, is a mathematical modeling structure inspired by biological neural network [41]. The ANN is used to model complex relationships between large sets of input data and output data and find out the latent patterns inside the data sets. The ANN sets up a number of nodes and connections which connects the input data with corresponding output data through some weights or so-called hidden neurons [42].

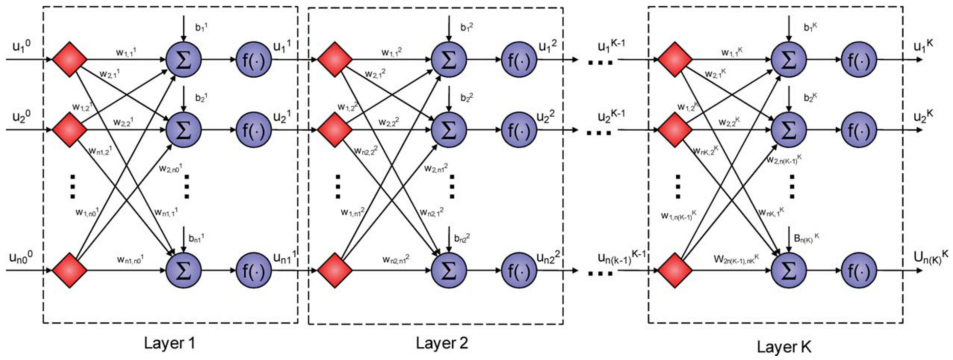


Figure 3.3: Graphic representation of a multilayer perceptron (MLP) [2]

One of the most widely used ANN configurations is the so-called multilayer perceptron (MLP) [43]. A typical MLP is a highly parallel computational structure composed by a series of layers [2]. In this structure, the output of each neuron is connected to every neuron in the subsequent layers, and the connection structure is in a cascade form with no connections between neurons in the same layer [43]. A representation of a typical MLP structure is presented in Figure 3.3. In essence, each layer k in the MLP structure acts as a transformation from an input vector $\mathbf{u}^{k-1}$ to an output vector $\mathbf{u}^k$, which can be expressed mathematically as [2]

$$\mathbf{u}^k = f(\mathbf{W}^k \cdot \mathbf{u}^{k-1} + \mathbf{b}^k), \quad k \in \mathbf{K} = \{1, ..., K\} \tag{3.8}$$

The relations defined in Equation (3.8) represents the mapping between the input $\mathbf{u}^0$ to the network and its corresponding output $\mathbf{u}^k$. Each layer k consists of n_k neurons with different parameters $\mathbf{b}^k$ and $\mathbf{W}^k$. The $\mathbf{b}^k$ is a layer bias vector with n_k dimensions and the $\mathbf{W}^k$ is a weight matrix with $n_k \times n_{k-1}$ dimensions. The function f is known as the so-called activation function, and generates the output vector $\mathbf{u}^k$ by acting on every component of its argument vector. The last

layer of this configuration is known as the output layer, whereas the rest of the layers are known as hidden layers. [2]

The fitting of ANN is known as training. Generally, the training mainly involves the minimization of total squared model deviations with respect to a subset of the sample points in the design space which is known as the training set [2]. There always exists a criterion to terminate the training. This termination criterion involves monitoring and evaluating the model deviations with respect to another subset of the sample points in the design space which is known as the validation set [2].The trained network features small deviations with respect to both the training set and the validation set can guarantee a good network with the capacity to reproduce data away from the training set. Hence the resulting network ensures a good model to replace the original model and predict the output data with satisfied accuracy.

3.3.4 Comparison of Different Surrogate Modeling Techniques

The three surrogate modeling approaches discussed above are summarized and compared in Table 3.1. In this project, the objective process is an ammonia plant modeled in HYSYS, indicating a deterministic process with no random errors present, and it involves highly nonlinear behaviors. In this case, ANN is well-suited for modeling this HYSYS process due to its ability to handle highly nonlinear behaviors as well as flexibility and parsimony.

There has been some works that utilized ANN to build surrogate models for process synthesis, optimization and control. Henao and Maravelias [2] constructed the main equipment models used in chemical plants with ANN-based surrogate modeling and applied the approach to process synthesis and optimization. Meert and Rijckaert [44] applied ANN to modeling a polymerisation reactor in a chemical process using two types of networks. Bloch and Denoeux

Table 3.1: Comparison of three surrogate model techniques [1, 3, 4]

Surrogate model techniques	Characteristics
Polynomial Regression	• Well-established and easy to use • Difficult to construct for high-dimension system • Fast convergence • Instabilities may arise when modeling highly nonlinear behavior
Kriging method	• Extremely flexible • Complex to construct • May be infeasible due to possible singular correlation matrix • Suitable for deterministic applications
Artificial Neural network	• Suitable for highly nonlinear or very large problems • Suitable for deterministic applications • Best for repeated application • Universal approximation capabilities • Parsimonious and flexible

[4] used ANN-based modeling technique in the control of alloying process in a hot dipped galvanizing line in steel industry and the control of a coagulation process in a drinking water treatment plant in water treatment, to address the complexity and non-linearity issues in the processes. Fernandes [45] constructed a neural network to substitute the complex reaction mechanism for optimization of diesel and gasoline production based on limited experimental data of the reaction. Mujtaba *et al.* [46] utilized ANN-based modeling technique in developing and implementing three different types of nonlinear control strategies in batch reactors. Therefore, in this project, ANN is chosen as the modeling technique to construct the surrogate models.

3.4 Design of Experiment

The design of experiment is performed to collect the sample data that can be used to generate the surrogate model. In this thesis, the term "experiment" essentially means the HYSYS model. The design of experiment consists of two steps. First the input variables and output variables are determine so that the variable design space can be determined. Next, the input design space is sampled using specific sampling method to obtain the input sample space. Then the input sample data can be imported to HYSYS to calculate the output sample space. The sample spaces are aimed to have the ability to provide enough information for the generation of surrogate models. More samples can certainly increase the accuracy of generated surrogate models. But the number of samples is severely limited due to the expensive computational cost of the HYSYS model, so the variables are desired to be reduced to a limited number. At the same time, the sampling method is aimed to be highly efficient so that we can use as few samples as possible to represent the variable design spaces. Based on the two aspects, variable analysis is performed to determine the variable design space and the proper sampling method is proposed in the following sections.

3.4.1 Variable Analysis

The variable analysis is crucial for surrogate modeling since the model generation fully depends on the sample data of input and output variables. It is desired to use as few variables as possible to save the computational expense of HYSYS and decrease the complexity of surrogate model generation. The variable analysis is performed by the variable classification, the variable reduction and the variable identification.

Variable Classification

In essence, a process modeled in HYSYS is a synthesis of individual unit models, which contains equations establishing relationships among different variables [2]. Therefore, a complete HYSYS model can be decomposed into different submodels which consist of various unit models. In this thesis, the submodels of HYSYS corresponds to the different sections of the ammonia process as mentioned in the Section 2. A modeling-oriented classification of variable sets are defined as follows:

$\mathbf{U}_h$: Set of input variables of a submodel in HYSYS;

$\mathbf{Y}_h$: Set of output variables of a submodel in HYSYS;

$\mathbf{I}_h$: Set of independent variables of a submodel in HYSYS;

$\mathbf{D}_h$: Set of dependent variables of a submodel in HYSYS;

$\mathbf{U}_s$: Set of input variables of a surrogate model;

$\mathbf{Y}_s$: Set of output variables of a surrogate model;

All the sets defined above is in the context of an individual submodel in HYSYS, corresponding to a single section of the ammonia process, which is aimed to be further re-modeled by surrogate modeling. For a particular submodel, the input variables ($\mathbf{U}_h$) are used to close the degrees of freedom, and they are usually not unique since there is normally not only one way to specify the input variables for HYSYS model definition. The input variables used to close the degrees of freedom are subset of independent variables ($\mathbf{I}_h$), and once some input variables are specified to close the degrees of freedom meanwhile the other input variables ($\mathbf{U}_h \setminus \mathbf{I}_h$) then become dependent variables ($\mathbf{U}_h \cap \mathbf{D}_h$). The selection of variables for $\mathbf{I}_h$ is important because it will affect the model fitting and also influence the performance of fitted surrogate models, which is discussed in details in Section 3.4.1. Certainly, all the output variables ($\mathbf{Y}_h$) are subsets of dependent variables ($\mathbf{D}_h$).

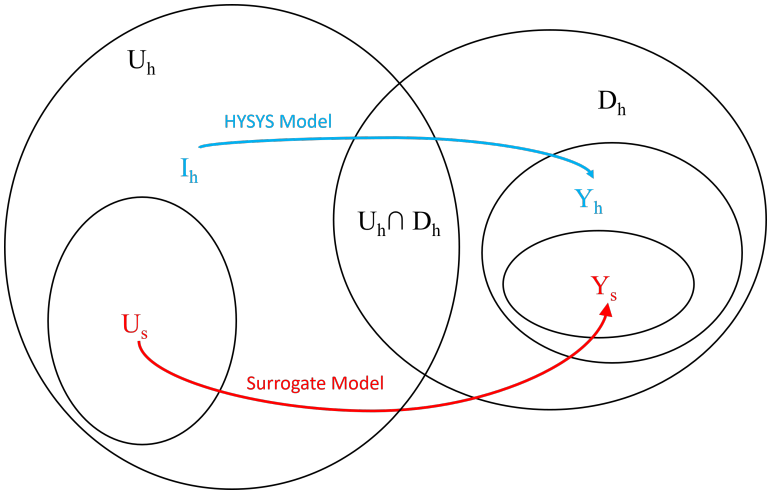


Figure 3.4: Venn diagrams representing relations among different variable sets for an individual model

The input variables ($\mathbf{U}_s$) and output variables ($\mathbf{Y}_s$) for model fitting are generally identical to or subsets of $\mathbf{U}_h$ and $\mathbf{Y}_h$. Since usually the surrogate models are desired to be easily handled, $\mathbf{U}_s$ and $\mathbf{Y}_s$ are higher likelihood to be subsets of $\mathbf{U}_h$ and $\mathbf{Y}_h$. Selecting the variables for $\mathbf{U}_s$ and $\mathbf{Y}_s$ are achieved by reducing the number of variables using internal relationships among the variables and\or other statistic methods. Figure 3.4 shows a Venn diagram for the relationships among different variable sets, where red arrow and sets represent the mapping and corresponding sets for surrogate models while blue arrow and sets represent the mapping and corresponding sets for the original HYSYS model.

Variable Reduction

When generating surrogate models, a certain level of accuracy is necessarily required. To fulfill the accuracy requirements, the variables used to construct the surrogate model should be identical to or at least contain the major information of the HYSYS variables. After the input and output variables for surrogate model

generation are determined, the resulting variable design spaces are sampled to obtain the sample spaces which are used to generate the surrogate models.

A key issue could arise when obtaining the sample spaces, which is the so-called *curse of dimensionality* [19], where the dimension indicates the number of variables. To demonstrate this issue, assuming a simple case where there are $k = 2$ input variables. Within the variation ranges, we need $n = 9$ samples of variables to generate a surrogate model satisfying the accuracy requirement. Hence, the total number of sample points for input sample space is 2^9, which means we need to call the HYSYS model 2^9 times to get the corresponding output sample data. This seems to be not too complex and we can still handle it. When it comes to a more complicated case where we have $k = 9$ variables. In order to obtain the sample space with the same sample density as the case only containing two input variables, the sample space would contain 9^9 points which is an incredibly large number. Calling the HYSYS for 9^9 times take a lot of computational expense. The situation could be worse if the number of variables is more or we need more sample points. Hence if the HYSYS model has many variables (in this thesis we got more than 9 variables for both input and output variables), the number of points needed to give reasonably uniform coverage of the variable design spaces rises exponentially [19]. This could cause massive trouble in obtaining sample spaces from the HYSYS model. Furthermore, the massive sample space could lead to incredible complexity for surrogate model generation. For example, if there are k_1 input variables and k_2 output variables of which n samples are required for all the variables in order to fulfill the accuracy requirement. Then the surrogate model need to be fitted based on the input sample space with n^{k_1} points and the output sample space with n^{k_2} points, which could spend a lot of computation expense on model generations and may cause unexpected troubles in further model manipulation. To tackle this issue,

one alternative approach is to reduce the number of variables used to generate surrogate models by utilizing some dependency relationships among the variables.

Dependency relationships among the variables are based on the physical properties and the specific features of the process itself. For example, the mass is always conserved throughout the process which implies that the mass flow of all inlet streams corresponds to the mass flow of all outlet streams. Other relations can be properties like input energy plus energy generation are equal to output energy plus energy consumption, or extent of reaction which indicates the mass conservation of the reactions [47]. Generally, relationships like these are basically based on mass balance or energy balances. Besides, other dependency relationships can be utilized such as the relations of temperatures and pressures which are often highly dependent on each other. For example, if there is a gas stream where the gas is assumed as classical thermodynamic ideal gas [48]. Then the properties of the ideal gas follow the ideal gas law noting as $PV = nRT$ so that the pressures can be determined by temperatures with the equation of ideal gas law or vice versa. Hence either the variable of pressure or temperature can be reduced using the ideal gas law equation. For describing the relation between vapor pressure and temperature for pure components, thermodynamic equations such as Antoine equations [49, 50] can be utilized to define the dependency of temperatures and pressures.

In addition, variable reduction can be also done by simply neglecting trivial variables. In some particular cases, some variables which have very limited influence on the process compared with the other ones can be neglected directly. For example, if there are two streams containing the same compositions mixed up with each other and the molar flow of one of them is 1000 times larger than the other one, the smaller one obviously has no significant influence on this

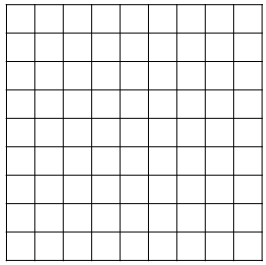
mixing process so that it can be kept as constants, meaning their variations are neglected, making the problem much simpler.

Variable Identification

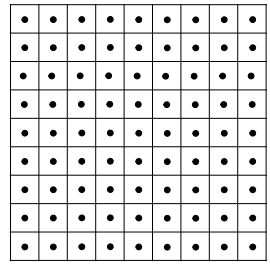
After variable reduction, the initial variables are reduced to design variables, which are able to identify the process with a minimized number of variables. These design variables are then used to generate the surrogate model. But the design variables can be determined in different ways. For example, we need the pressures of both inlet streams and outlet streams to identify the process. Meanwhile, we can also use the pressure of the inlet stream and the pressure difference between the inlet and outlet streams to obtain the same information. These two options are essentially based on the same information but the design variables are different in these two cases, which could have a significant influence on surrogate model generation.

3.4.2 Sampling Methods

After variable reduction and identification, the design variables are determined with corresponding variation ranges. Hence the variable design space is determined by discretizing the continuous input variables using the so-called space-filling design which treat all regions of the design space equally [51]. This is done by dividing the variation ranges into m equal intervals and hence a design space for the input variables is obtained. For example, if there are k input variables with m intervals for each, a design space with k dimensions and m values in each dimension is obtained. Consider a simple case where $k = 2$ and $m = 9$, the design space can be demonstrated as a rectangular square with 81 cells as shown in Figure 3.5a, where each cell contains a data point with two variable values.



(a) The variable design space, where each cell represents one data point



(b) The sample space obtained by complete sampling, where the black point representing the sample points selected from the data points

Figure 3.5: The design space and the sample space designed by complete sampling

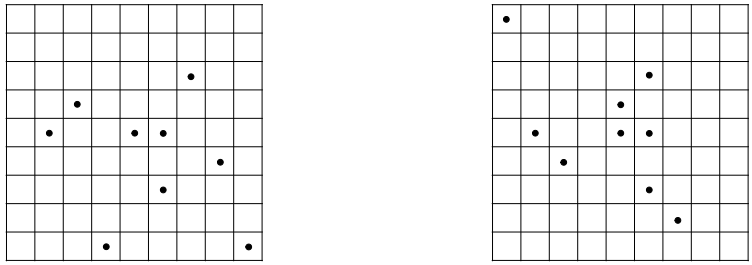
Based on the design space, a sample space needs to be determined using specific sampling methods. The most straightforward way to sample the design space is using the full design space as the sample space, meaning all the data points in the cells in the design space are selected as samples, the resulting sample space can be illustrated as Figure 3.5b. This way to design a sample space is called complete sampling [20], also known as full factorial sampling [19], where the sample space is in a uniform matter containing a full rectangular grid of points. However, the so-called problem *curse of dimensionality* [19] mentioned in Section 3.4.1 would be present.

To deal with the curse of dimensionality, variable reduction is used as mentioned in Section 3.4.1. By using dependency relationships, the variables can be reduced in some extent. Hence the computation expense of sample space design as well as the complexity of the resulting surrogate model are decreased. Furthermore, the manipulation of the surrogate model would be more efficient and easily handled. However, the reduction of the variables have limited improvements for solving curse of dimensionality because we need at least a certain number of variables to generate the surrogate model in order to fulfill the

desired accuracy. Although the variables are reduced, the sample points still increase exponentially with the minimized number of variables using complete sampling. The issue of curse of dimensionality is not essentially solved by just reducing the variables when using complete sampling. Hence instead of complete sampling, the sampling strategy used for sampling the variables should be more efficient, so that with a certain number of samples it can make the sample space contain as much information of the design space as possible. This ability depends on the wide spread of sample points throughout the design space, known as the property of space-filling [19]. In order to result in a certain level of surrogate model accuracy, a good space-filling design is required, in which the sample points throughout the design space can yield minimal unsampled regions [26], and hence the resultant sample space can represent the design space efficiently.

Therefore, instead of the complete sampling, other sampling strategies which could take less sample points but still can represent the design space efficiently are desired to design the sample space, such as simple random sampling [21, 22], Latin hypercube sampling [22, 23] and orthogonal Latin hypercube sampling [24, 25, 26, 27].

The simple random sampling (SRS) is a sampling design, where n distinct items are selected from the N items of the population in such a way that every possible combination of n items is equally likely to be the sample selected [52]. For example, we still consider the case where there are $k = 2$ variables and each variable has $m = 9$ intervals in between its variation range. If $n = 9$ samples are desired in the design space with a population of $N = 81$ cells in total, two possible sample spaces have equal probability of being selected, which are depicted in Figure 3.6. However, there is a significant disadvantage making this method less effective for this project. Because the samples existing in the same rows

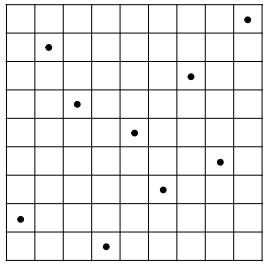


(a) (b)

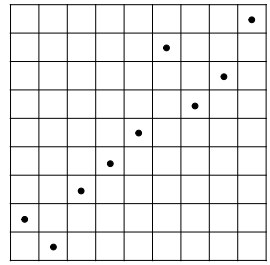
Figure 3.6: Two sample spaces designed by simple random sampling

or columns as shown in Figure 3.6 contain the repetitive information, so that the samples cannot represent the design space efficiently enough with a certain sample number, leading to waste of computational expense. In this project, the objective of sampling is to use limited number of samples to represent the whole design space efficiently. Hence a sampling method which can make samples containing the data information as much as possible is desired. Obviously, SRS is not satisfactory in this case.

The latin hypercube sampling (LHS) is a sampling method for generating a near-random sample space from a multidimensional design space, which can be viewed as a multidimensional extension of Latin square sampling [22]. In contrast with the complete sampling of the two-dimension design space, a Latin square is a square grid containing samples where there is only one sample in each row and each column. We still consider a case with two variables, indicating design space with two dimensions. Assuming $n = 9$ samples are required from the design space with a population of $N = 81$ cells, then the sampling method of LHS can be depicted as shown in Figure 3.7. Compared with Figure 3.6, it can be seen that for LHS there is only one point in each row and column. Hence with the same number of sample points, LHS can make samples contain more information of the design space so that it can represent the variables more



(a) A typical sample space designed by Latin hypercube sampling

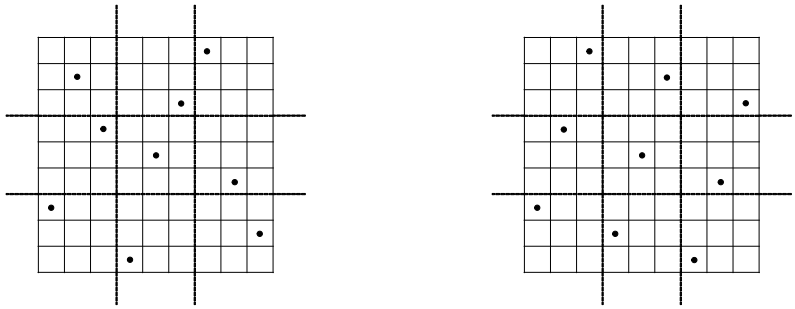


(b) A sample space designed by Latin hypercube sampling where high correlations occur among the sample points

Figure 3.7: Two sample spaces designed by Latin hypercube sampling

efficiently.

However, the conventional construction of LHS by mating samples near-randomly is possible to have potential high correlations among the samples [27] as depicted in Figure 3.7b. The sample space with high correlations may result in incorrect statistic features when representing the design space. Hence some optimal improvements have been introduced to LHS, of which the sampling method are known as orthogonal Latin hypercube sampling [26, 27], also known as orthogonal-array-based latin hypercubes [53]. The basic idea of this sampling method is dividing the whole design space into subspaces and using LHS in each subspace. Then the sample points are selected in the whole design space simultaneously, ensuring that the samples in both subspaces and the whole space are LHS samples. Hence it can result in a well-distributed sample space with good space-filling property, which has a good ability to represent the whole design space as we desire. Two examples with $k = 2$ variables and $m = 9$ intervals of each variable are shown in Figure 3.8. The uniformity can be designed as the desired form using optimal orthogonal-array-based latin hypercubes [53] as shown in Figure 3.8b but with a cost of higher computational expense.



(a) A sample space designed by orthogonal Latin hypercube sampling

(b) A sample space designed by orthogonal Latin hypercube sampling with desired uniformity

Figure 3.8: Two sample spaces designed by orthogonal Latin hypercube sampling

In this project, we utilize a nearly orthogonal Latin hypercube sampling method to design the sample space as shown in Figure 3.7a, which can ensure the resultant sample space contains as much information as possible with a certain number of sample points. Hence it can be a good representative of the variable design space.

Adaptive Sampling

A main problem of sample space design is how many sample points are required. Obviously more sample points can increase the accuracy of the generated surrogate model but at the same too many sample points require unaffordable computation expense. Hence an approach is aimed to be implemented to decide the number of sample points, so that the sample points can be as few as possible meanwhile the sample space can still ensure the accuracy of the surrogate model generation.

In this purpose, a sampling method called adaptive sampling proposed by Straus and Skogestad [54] was implemented to sample the input design space. The adaptive sampling is based on the LHS sampling method mentioned in the

Section 3.4.2. Since LHS requires sample points more than the number of input variables k, the adaptive sampling start with sample points equal to k at the initial sampling step $s = 0$. At each following sampling step s, the number of samples increase by a interval t. The interval t could be 1, 10 or more depending on the implementation situations. At the sampling step s, the number of samples n can be calculated as

$$n = k + s \cdot t \tag{3.9}$$

At each sampling step, the input samples of $\mathbf{U}_s$ are imported to HYSYS to obtain the corresponding output samples of $\mathbf{Y}_s$. We store the input sample space as a matrix **X** and the output sample space as a matrix **Y**. In the matrices, each row contains the data of one sample point and each column corresponds to different variables. Hence **X** is a n-by-k matrix and **Y** is a n-by-m matrix where m is the number of output variables. Then we compute a partial least square regression of **Y** on **X** with $ncomp$ partial least square regression components.

In this thesis, the PLS regression is calculated using MATLAB. The calculation results in the predictor loadings of input sample space denoted as **XL**. The **XL** is a k-by-$ncomp$ matrix where each row contains coefficients which define a linear combination of the PLS components that approximate the original input variables. Hence **XL** contains the latent information and properties of the input sample space.

In MATLAB, the PLS regression is calculated using the SIMPLS algorithm [55]. First, **X** and **Y** are centered by subtracting off the column means to get centered variables **X0** and **Y0**. Since it does not rescale the columns, normalization of **X** and **Y** are performed before implementing PLS regression. The predictor scores **XS** can be also obtained, which contains the PLS components that are

linear combinations of the variables in **X**. The matrix **XS** is an n-by-$ncomp$ orthonormal matrix with rows corresponding to observations and columns to components. The relationships between the scores **XS**, loadings **XL** and centered variables **X0** are shown in the Equation 3.10.

$$\mathbf{XL} = (\mathbf{XS} \backslash \mathbf{X0})' = \mathbf{X0}' \cdot \mathbf{XS} \tag{3.10}$$

As the number of sampling points is increased at each sampling step s, the corresponding $\mathbf{XL}_s$ is recalculated. By comparing the current loadings $\mathbf{XL}_s$ and the loadings $\mathbf{XL}_{s-1}$ of the last step, we can know whether the new sample space contains the same information. If the deviations of the loadings are very small and below a desired threshold, we think the current sample space have the ability to explain the input variables with the desired accuracy. Therefore the input sample space and the corresponding output sample space can be used to generate the surrogate model.

However, not all the PLS components have significant contributions to the input sample space. Hence we also compute the percentage of variance explained in **X** by each component, which is stored in a 1-by-$ncomp$ vector denoted as **XV**. If the element XV_{ncomp} is less than a threshold of σ, we think the corresponding component does not make a significant contribution to explain the input variables and hence they are not necessary to be compared at each sampling step.

The comparison of the significant components is achieved by calculating the 2-norm of the difference between the current components and the last components. If the number of the significant components are not the same, we use the more components for comparison. If the norm is below some threshold θ, we think the sample space is good enough to represent the input variables.

Then the input sample space and the corresponding output sample space can be used to generate the surrogate model.

The adaptive sampling has the advantage of saving the computation expense. Since the HYSYS simulation takes time and so does the ANN fitting of surrogate models, this sampling method is able to minimize the number of required sample points with the defined criterion. The termination condition and other parameters such as the sample step size are adjustable so that it is flexible enough to fulfill the desired accuracy requirement.

Chapter 4

Constructing the Surrogate Model

In this chapter, a surrogate model is aimed to be constructed for the S-R section. The variable analysis is firstly performed to determine the variables which can define the process in the HYSYS model. Then the variables are reduced using dependency relationships in order to decrease the computational expense of both sampling and subsequent surrogate model generation. After reduction, the resultant variables are identified differently based on absolute outlet variables and variable differences. Then different variable design spaces are obtained. The adaptive sampling is implemented to sample these design spaces to obtain sample spaces. Based on the resulting sample spaces, two surrogate models are aimed to be generated using ANN. However, since the convergence issue occurred in the HYSYS model, the surrogate model is not able to be generated for the S-R section. To address this issue, we decompose the S-R section furthermore into three sections as the HEx part, the separator part and the refrigeration section. The surrogate model has been generated for the HEx part

using the same approach. The surrogate models are then saved as functions for validation.

4.1 Model Construction for the S-R Section

As mentioned in Section 3.1, the separation section and refrigeration section are coupled with each other by three recycle loops hence they are considered as one single S-R section. The PFD of the S-R section is shown in the Figure B.1, where the shadow grey blocks represents the connected reaction section and the makeup section, the red flows note the inlet streams to the S-R section, the blue flows note the outlet streams from the S-R section to other sections and the green flows note the streams leaving the plant. We now consider the S-R section as the objective process and the aim is to construct a surrogate model for this section. Regarding the HYSYS model as the "experiment", the construction of the surrogate model follows the approach as shown in the Figure 3.2.

4.1.1 Variable Analysis

As mentioned in the Section 3.4.1, in order to construct a surrogate model for the HYSYS model process, the variables defining the HYSYS model are needed to be classified primarily as the set of input variables $\mathbf{I}_h$, which are used to close the degree of freedom of the HYSYS model defining the process, and the set of output variables $\mathbf{Y}_h$ which are the resulting variables calculated by the HYSYS simulation.

The input variables of $\mathbf{I}_h$ consist of the manipulated variables and the inlet stream variables. The manipulated variables are the variables that can affect the operation conditions of the process, and the inlet stream variables are the molar flows, temperatures and pressures. In the S-R section as shown in the Figure

B.1, the manipulated variables are the mass flow of cooling water in the coolers CW-2, CW-3 and CW-4, denoted as $m_{BFW,2}$, $m_{BFW,3}$ and $m_{BFW,4}$ respectively, the split fraction of lower substream in split Sp-1, denoted as r_{Sp}, the revolution speed of the air cooler AC-2, denoted as RPM_{AC-2}, the temperature difference between the outlet and inlet flows of the heat exchanger HEx-1, denoted as ΔT_{HEx-1}, and the revolution speed of the compressors C2-1, C2-2 and C2-3, denoted as RPM_{C2}. Since the compressors C2-1, C2-2 and C2-3 have the same draft, the revolution speed of them are hence the same and count for only one input variable. The nominal conditions and corresponding variation ranges of these manipulated variables are shown in Table A.1.

The S-R section has two inlet streams denoted as N_{RH} and N_{MS} as shown in Figure B.1. The stream N_{RH} is from the reaction section and the stream N_{MS} is from the makeup section. Both of the inlet streams contain the input variables as the molar flows of different compositions, the temperatures and the pressures in the streams, which are denoted as $N_{i,j}$, T_j and p_j respectively, where i represents the compositions with $i \in \{NH_3, CH_4, H_2, H_2O, He, Ar, N_2\}$ and j represents the streams with $j \in \{RH, MS\}$. The nominal conditions and variation ranges of the variables in the two inlet steams are identified in Table A.2.

By importing the input variables to the HYSYS model, the resulting variables can be obtained which are hence the output variables defined as the set $\mathbf{Y}_h$. The output variables are essentially the variables in the outlet streams of the S-R section. This section has six outlet streams in total as N_{HM}, N_{SM}, N_S, N_{R1}, N_{R2} and N_{R3} shown in the Figure B.1, which contains the output variables as the molar flows of different compositions $N_{i,k}$, the temperatures T_k and the pressures p_k, where $i \in \{NH_3, CH_4, H_2, H_2O, He, Ar, N_2\}$ and $k \in \{HM, SM, S, R1, R2, R3, R4\}$. The resulting output variables exported from HYSYS with the input variables at nominal conditions are listed in Table A.3.

The input and output variables defined above can identify the HYSYS model of the S-R section. Based on the set definition and the corresponding variables classification, the sets of input variables $\mathbf{I}_h$ and output variables $\mathbf{Y}_h$ of the S-R section are summarized in Table C.1.

Variable Reduction

The variables that are needed to define the HYSYS model has been classified as the set of input variables $\mathbf{I}_h$ and the set of output variables $\mathbf{Y}_h$. In essence, the surrogate model is constructed to describe the mapping relationships from $\mathbf{I}_h$ to $\mathbf{Y}_h$. Hence the computation expense mainly depends how many variables are needed. In order to decrease the computational expense of both sampling and surrogate model generation, the variables of $\mathbf{I}_h$ and $\mathbf{Y}_h$ are aimed to be reduced as many as possible without affecting the accuracy requirement of the resultant surrogate model. This section concerns the possible ways to reduce the variables.

Neglecting Trivial Variables

One straightforward way to reduce variables is neglecting the trivial variables. The trivial variables do not have significant influence on the process so that they can be neglected. In this process, some molar flows can be neglected due to their negligible contributions compared with other molar flows. For example, as shown in the Table A.2, the molar flow of H_2O in the inlet stream $N_{H_2O,RH}$ is about 1000 times smaller compared to $N_{H_2O,MS}$. Since they are both input variables of the process concerning the same composition, $N_{H_2O,55}$ does not have significant contribution to the H_2O amount imported into the process. Hence it can be neglected and reduced form the input variables. The same situations can also be observed for the molar flows of $N_{Ar,MS}$, $N_{CH_4,MS}$ and $N_{He,MS}$ since they are all about 1000 times smaller than the molar flows with the same

compositions. Therefore, $N_{Ar,MS}$, $N_{CH_4,MS}$ and $N_{He,MS}$ can be neglected and reduced from the input variables.

Some of the output variables are also trivial variables and can be reduced. From Table A.3, it can be found out that $N_{H_2,R1}$, $N_{He,R1}$, $N_{Ar,R1}$ and $N_{N_2,R1}$ are too small and hence can be reduced from the output variables. To be noticeable, the composition of H_2O is only present in the inlet stream N_{MS} and the outlet stream N_{R1}, indicating all the H_2O feed to the process will eventually leaves the process through the outlet stream N_{R1}. Hence the molar flow of H_2O is essentially not a design variable for surrogate model generation and the relevant variables can be reduced from both the input and output variables.

Mass Conservation

As mentioned in the Section 3.4.1, mass conservation is a crucial property to be utilized to reduce the variables such as molar flows. In order to illustrate the mass conservation throughout the S-R section, the process is divided into several parts as shown in the Figure B.2 and it is also simplified as a block diagram with only mass flows present as shown in the Figure 4.1.

In the Figure B.2, the S-R section is divided into two sections as separation section and the refrigeration section which are represented by grey shadow blocks. Furthermore, the separation section is divided into two parts as the HEx part, which mainly contains the heat exchanger network plus one separator, and the separator part, which contains only one isolated separator. The HEx part and separator part are represented by dark blue shadow blocks. In the simplified diagram in Figure 4.1 The connecting flows, inlet flows and outlet flows of different sections and parts are clearly visualized. From Figure 4.1, following the flow moving pattern, the mass conservation relationships of the molar flows can be illustrated as Equations (4.1).

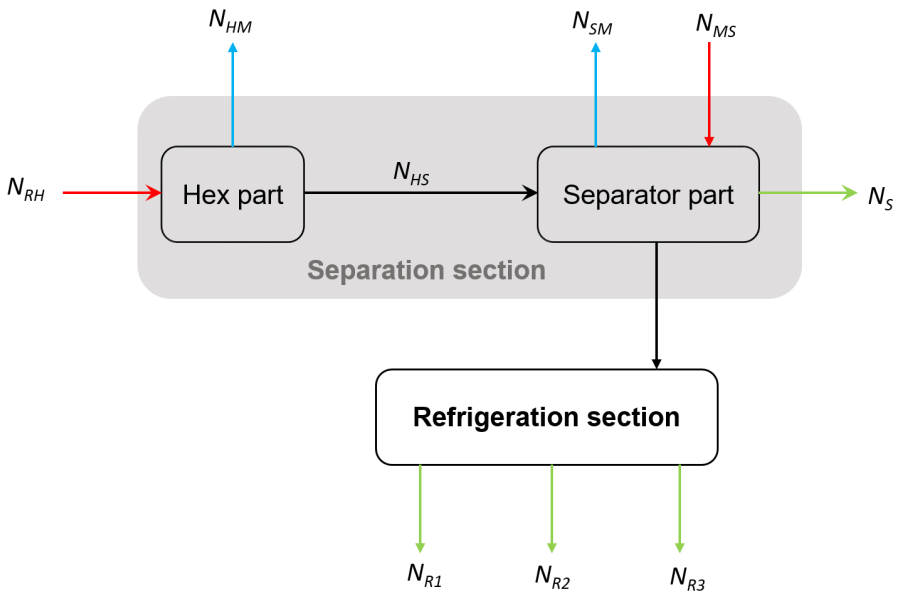


Figure 4.1: The simplified block diagram of the S-R section

$$N_{RH} = N_{HM} + N_{HS} \tag{4.1a}$$

$$N_{HS} + N_{MS} = N_{SM} + N_S + N_{R1} + N_{R2} + N_{R3} \tag{4.1b}$$

From Equation (4.1), it can be seen that some output molar flows are not independent variables and can be calculated by the input molar flows and the other output molar flows. Hence these dependent output molar flows can be reduced based on the Equations (4.1). However, since some trivial flows have been neglected, the mass balance equations (4.1) can not be fulfilled exactly for every composition. In order to address this issue and reduce some of the output molar flows, we introduce the fractional factors γ and β as the ratios between independent outlet molar flows and inlet molar flows, which are defined

in Equations 4.2.

$$\gamma = \frac{N_{HS}}{N_{RH}} \tag{4.2a}$$

$$\beta_1 = \frac{N_{SM}}{N_{HS} + N_{MS}} \tag{4.2b}$$

$$\beta_2 = \frac{N_S}{N_{HS} + N_{MS}} \tag{4.2c}$$

$$\beta_3 = \frac{N_{R1}}{N_{HS} + N_{MS}} \tag{4.2d}$$

$$\beta_4 = \frac{N_{R2}}{N_{HS} + N_{MS}} \tag{4.2e}$$

$$\tag{4.2f}$$

Hence we can use the input variables and the corresponding fractional factors to calculate both the independent and dependent output molar flows. Here we consider molar flows in N_{R3} and N_{HM} as the dependent variables and they can be calculated by input molar flows with fractional factors by Equations 4.3. Therefore, molar flows in N_{R3} and N_{HM} can be reduced from the output variables and fractional factors are introduced to substitute the dependent output molar flows.

$$N_{HM} = (1 - \gamma) \cdot N_{RH} \tag{4.3a}$$

$$N_{R3} = (1 - \beta_1 - \beta_2 - \beta_3 - \beta_4) \cdot (\gamma N_{RH} + N_{MS}) \tag{4.3b}$$

The input variables of $\mathbf{I}_h$ and the output variables of $\mathbf{Y}_h$ have been reduced to a minimal number by variable reduction. Hence at this point the remaining variables are essentially the variables of the set $\mathbf{U}_s$ and the set $\mathbf{Y}_s$ to generate the surrogate models. Since here we use the absolute outlet variables to define $\mathbf{Y}_s$, it is denoted as $\mathbf{Y}_{s,abs}$. The resultant variables of $\mathbf{U}_s$ and $\mathbf{Y}_{s,abs}$ are listed in Table

C.2.

Comparing the Table C.4 with the Table C.2, it can be observed that the number of input variables are reduced from 25 to 21, and the number of output variables are reduced from 54 to 42. The variables have been reduced effectively by the dependency relationships and neglecting trivial variables. The variable reduction saves the computational expense of both sampling and model generation. Besides, it also makes the resulting surrogate model easier to handle for the further optimization.

Variable Identification

The variables to generate surrogate models are determined after variable reduction as shown in Table C.2. But some output variables can be identified in another way while still contain the same variable information as the previous. The output variables such as the pressures and temperatures can be substituted by the pressure differences and temperatures differences instead of the absolute values. The variables differences are defined in Equations 4.4. With the variable differences, the variables can be identified as shown in the Table C.3, where the output variables are denoted as $\mathbf{Y}_{s,diff}$ now.

The different identifications of the input and output variables essentially contain the same data information, but it may have significant influence on the surrogate model generation even using the same modeling techniques. This will be discussed in Chapter 5.

$$
\begin{aligned}
\Delta p_{HM} &= p_{HM} - p_{RH} & \Delta T_{HM} &= T_{HM} - T_{RH} \\
\Delta p_{SM} &= p_{SM} - p_{RH} & \Delta T_{SM} &= T_{SM} - T_{RH} \\
\Delta p_{S} &= p_{S} - p_{RH} & \Delta T_{S} &= T_{S} - T_{RH} \\
\Delta p_{R1} &= p_{R1} - p_{RH} & \Delta T_{R1} &= T_{R1} - T_{RH} \\
\Delta p_{R2} &= p_{R2} - p_{RH} & \Delta T_{R2} &= T_{R2} - T_{RH} \\
\Delta p_{R3} &= p_{R3} - p_{RH} & \Delta T_{R3} &= T_{R3} - T_{RH}
\end{aligned} \tag{4.4}
$$

4.1.2 Sampling of the Design Space

The design space of the input variables in $\mathbf{U}_s$ is determined after variable identifications. The input design space is aimed to be sampled to obtain the input sample space $\mathbf{X}$ which will be imported in the HYSYS model, and hence the corresponding sample space of the output variables $\mathbf{Y}$ can be obtained.

We utilize adaptive sampling mentioned in Section 3.4.2 to sample the input design space. Since there are 21 input variables in this case, the initial sample points must be more than 21 points. We start from $n_0 = 30$ points and then increase the number of the sample points by an interval of $t = 10$ at each sampling step. We set the thresholds $\sigma = 1\%$ and $\theta = 0.5$ for the adaptive sampling.

Unfortunately, the sampling did not work out since the convergence issues arose in the heat exchangers of the HYSYS model, and we cannot get the output sample data. This issue is mainly caused by the recycle loops involving the heat exchangers. Hence, we decided to decompose the S-R section furthermore into HEx part, separator part and refrigeration section as illustrated in Figure B.2. Then surrogate models are constructed for individual parts using the same approach.

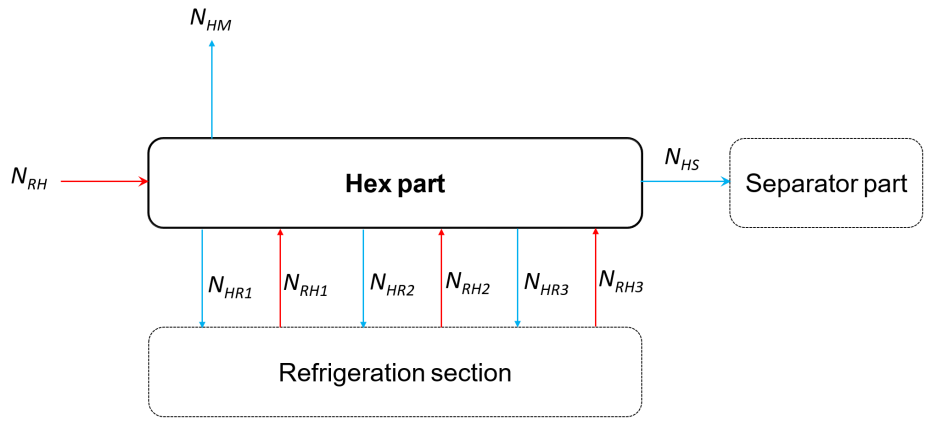


Figure 4.2: The simplified diagram of the HEx part

4.2 Surrogate Model Construction for the HEx Part

Since it was not feasible to model the whole S-R section, we divided it into three parts as shown in Figure B.2 and aim to construct surrogate models for each part. In this section, the surrogate model is constructed for the HEx part. The model construction follows the approach shown in the diagram of Figure 3.2. The following sections describe the detailed procedures of the construction.

4.2.1 Variable Analysis

Similar to the S-R section, $\mathbf{I}_h$ and $\mathbf{Y}_h$ should be defined firstly for HEx part. Since the HEx part is separated with the refrigeration section by breaking the three recycle loops as shown in Figure B.2, three new inlet streams and three new outlet stream are present additionally as N_{RH1}, N_{RH2}, N_{RH3}, N_{HR1}, N_{HR2} and N_{HR3} as shown in the PFD of Figure B.2. The flowsheet can then be simplified as a block diagram with noting flows as Figure 4.2.

However, the molar flows of all the compositions in the six recycle streams are not varying much and can be regarded as constants. Hence these molar

flows are not considered as variables. But the variations of the pressures and temperatures in these streams can not be neglected.

In the HEx part, the manipulated variables are the mass flow of cooling water in the cooler CW-2, denoted as $m_{BFW,2}$ and the split ratio of the split Sp-1, denoted as r_{SP-1}. The HEx part process has four inlet streams, the feed to the process N_{RH} and three inlet streams to the heat exchangers as N_{RH1}, N_{RH2} and N_{RH3}. The input variables with their nominal conditions and variation ranges are listed in Table A.4.

The output variables as the resulting variables yielded from HYSYS simulations are the variables of outlet streams, including three outlet streams leaving the heat exchangers as N_{HR1}, N_{HR2} and N_{HR3}, one stream leaving the HEx part to the makeup section as N_{HM}, and one stream passing to the downstream separator part as N_{HS}. The stream configuration can be visualized in Figure 4.2. The output variables at nominal conditions are listed in Table A.5.

The input and output variables defined above can identify the HYSYS model of the HEx part. Based on the set definition and the corresponding variables classification, the sets of input variables $\mathbf{I}_h$ and output variables $\mathbf{Y}_h$ of the HEx part process in the HYSYS model are identified in Table C.4.

Variable Reduction

The variables defining the HEx part has been classified as the set of the input variables $\mathbf{I}_h$ and the set of the output variables $\mathbf{Y}_h$. The input variables are used to close the degree of freedom of the HYSYS model but they are not exactly independent with each other. For instance, temperatures could be dependent on pressures in a stream with specific compositions. The output variables also have dependency relationships within themselves or with the input variables due to mass conservation throughout the process. The dependency between pressures

and temperatures as well as the mass conservation throughout the process are considered in the following sections to reduce the input and output variables.

Mass conservation

As mentioned in Section 3.4.1, mass conservation is a crucial property to be utilized to reduce the variables such as molar flows. Based on the simplified flow diagram in Figure 4.2, the mass balance through out the HEx part can be expressed as

$$N_{RH} = N_{HM} + N_{HS} \tag{4.5}$$

From the nominal conditions of input variables in Table A.4 and output variables in Table A.5, H_2O is only present in the inlet stream N_{RH} and outlet stream N_{HS}. Hence $N_{H_20,HR}$ is always equal to $N_{H_2O,RS}$, indicating they are not required for surrogate model generation. Therefore, N_{RH} and N_{HS} can be reduced from the input and output variables.

The Equation (4.5) defines the dependency of the outlet molar flows on the inlet molar flows. Similar to the Section 4.1.1, here we define the fractional factor α as Equation 4.6, and consider the stream N_{HM} as dependent variables. Therefore variables in the stream N_{HM} can be calculated by the input variables and the fractional factor α.

$$\alpha = \frac{N_{HS}}{N_{RH}} \tag{4.6}$$

Hence the variables in the stream N_{HM} can be calculated by the input variables N_{RH} and the fractional factor α as

$$N_{HM} = (1-\alpha)N_{RH} \tag{4.7}$$

Antoine equation

From Figure B.2, it can be seen that the inlet streams to the heat exchangers

Composition	N_{RH1} (HEx5)	N_{RH2} (HEx6)	N_{RH3} (HEx7)
NH_3	0.9999	1.0000	0.9995
CH_4	0.0001	0.0000	-
H_2O	-	-	0.0005

Table 4.1: The compositions in the three recycle loops

HEx-5, HEx-6 and HEx-7 come from the corresponding upstream separators S-7, S-6 and S-5 in the refrigeration section. The gas and liquid phase inside the separators are in thermodynamic equilibrium, and hence the the outlet streams of the separators contain the liquid under the saturated vapor pressure, indicating there is dependent relationship between the temperature and the pressure in the outlet streams out of the separators. The composition of the three streams are mainly ammonia with little impurities, and the composition fractions of the streams are shown in the Table 4.1. The three compositions are almost pure ammonia so that the behavior of the temperature and the pressure may be approximately described using the Antoine equation [56] or other modified forms of the Antoine equation. The dependency relationship between the temperature and the pressure is aimed to be described by these equations so that one of these two variables can be reduced form the input variables.

The Antoine equation is a class of semi-empirical correlations describing the relation between vapor pressure and temperature for pure components. A typical Antoine equation is given by

$$\ln p = A - \frac{B}{C+T} \tag{4.8}$$

where P [bar] represents the vapor pressure and T [K] represents the temperature. The A, B and C are the parameters of the specific conditions and vary in different temperature range. The temperature range for considering all the 3

Temperature [K]	A	B	C
239.6 - 371.5	4.86886	1113.928	-10.409

Table 4.2: Parameters of the Antoine equation of ammonia [5]

Temperature [K]	A	B	C	D	E	F
239.8 - 405.55	59.655	-4261.5	0	-6.9048	$1.0017 \cdot 10^{-5}$	2

Table 4.3: Parameters of the modified Antoine equation of ammonia from the HYSYS' library

streams in this case is between 253 K and 288 K. The corresponding parameters in this temperature range are shown in the Table 4.2.

Instead of the original Antoine equation, the HYSYS' library utilize various modified forms of the Antoine equation, of which one can be expressed as Equation (4.9).

$$\ln p = A + \frac{B}{C+T} + D \ln T + E T^F \tag{4.9}$$

where p is the vapor pressure with the unit of kPa and T is the temperature with the unit of K. The A, B, C, D, E and F are the parameters, which are listed in the Table 4.3 for the corresponding temperature range.

The Antoine equation and the modified Antoine equation are implemented by using the pressures p as the independent variable and then the equations calculate the temperatures T as the resulting variables. The idea is to use the equations and the pressures p to get the temperatures T, so that the temperatures T are dependent variables that can be reduced from the input variables. They results from the equations are validated by comparing them with the simulation data extracted from the HYSYS model calculated by the same pressures. It was found out that there exist the deviations between the equations and the HYSYS model. The reason could be that the composition is essentially not pure

ammonia and hence the equations with the proposed parameters are not sufficient to describe the behaviors between the pressures and the temperatures. However, this means the approach of using Antoine equation or modified Antoine equation to describe the dependency of temperatures on pressures did not work out and can not be used to reduce the variable of the temperatures T.

Another way to describe the dependency of T on p is to fit a simple surrogate model for them. Since the typical curve pattern of the Antoine equation is like a quadratic function form, it is reasonable to use the quadratic polynomials to model the relationships between the pressures p and the temperatures T for all the three streams, and hence the corresponding model equations can be obtained. The input variables of this equation are pressures p and the output variables are the corresponding temperatures T. Using this function the temperatures can be reduced from the input variables of the HEx part since the temperatures can be calculated by the pressures. The polynominal function is fitted for each stream using MATLAB. The resulting functions of polynomials are validated by using new data other than the data used to fit the polynomials. And it was shown that the polynomials can describe the dependency relationship quite good without significant deviations. The curves of the Antoine equation, modified Antoine equation and three quadratic polynominal models are visualized in Figure 4.3. The patterns of the curves are almost the same whereas small deviations can be observed among different curves.

After variable reduction, the input variables of $\mathbf{I}_h$ and the corresponding output variables of $\mathbf{Y}_h$ are reduced to $\mathbf{U}_s$ and $\mathbf{Y}_s$ respectively, which contain the variables for surrogate model generation. To note the outlet molar flows are defined in absolute values, the output variables are denoted as $\mathbf{Y}_{s,abs}$. The variables of $\mathbf{U}_s$ and $\mathbf{Y}_{s,abs}$ defined for the HEx part are listed in the Table C.5.

Comparing the Table C.4 with the Table C.5, it can be known that the num-

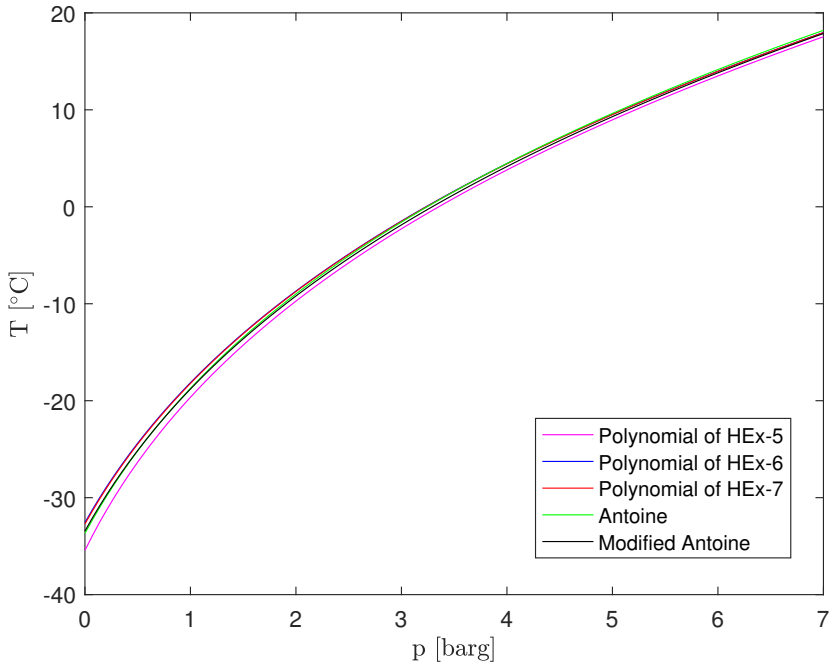


Figure 4.3: The p-T curves of different model

ber of input variables are reduced from 17 to 13 and the number of output variables are reduced from 24 to 16. The variables are reduced effectively using the dependency relationships and hence it could help with saving the computation expense of both sampling and model generation.

Variable Identification

Similar to the Section 4.1.1, the output variables in the HEx part such as the pressures and temperatures can be defined in another way by using the pressure and temperature differences instead of the outlet pressures and temperatures. The variable differences for corresponding outlet pressures and temperatures are defined in Equations (4.10). By substituting the absolute pressures and tem-

pratures with variable differences, the input and output variables denoted as $\mathbf{U}_s$ and $\mathbf{Y}_{s,diff}$ can be identified in Table C.6.

$$\begin{aligned}
\Delta p_{HR1} &= p_{HR1} - p_{RH1} \\
\Delta p_{HR2} &= p_{HR2} - p_{RH2} \\
\Delta p_{HR3} &= p_{HR3} - p_{RH3} \\
\Delta p_{HM} &= p_{HM} - p_{RH} \\
\Delta p_{HS} &= p_{HS} - p_{RH}
\end{aligned}
\qquad
\begin{aligned}
\Delta T_{HM} &= T_{HM} - T_{RH} \\
\Delta T_{HS} &= T_{HS} - T_{RH}
\end{aligned}
\tag{4.10}$$

The different identifications of the input and output variables essentially contain the same data information, but it can have significant influence on the surrogate model generation even though using the same modeling techniques. The detailed discussions are presented in Chapter 5.

4.2.2 Sampling of the Design Space

The design space of the input variables of $\mathbf{U}_s$ is determined after variable identifications. Then we use the adaptive sampling mentioned in Section 4.2.3 to obtain the input sample space $\mathbf{X}$. And the corresponding output sample space $\mathbf{Y}$. Since there are 13 input variables in $\mathbf{U}_s$, the initial sample points must be more than 13 points. We start from $n_0 = 20$ points and then increase the number of the sample points by an interval of $t = 10$ at each sampling step. We set the thresholds $\sigma = 1\%$ and $\theta = 0.5$ for this adaptive sampling.

Since the output variables can be identified in two ways as $\mathbf{Y}_{s,diff}$ in Table C.6 and $\mathbf{Y}_{s,abs}$ Table C.5. The adaptive sampling is implemented for both cases to obtain the sample spaces $\mathbf{X}$ and $\mathbf{Y}$.

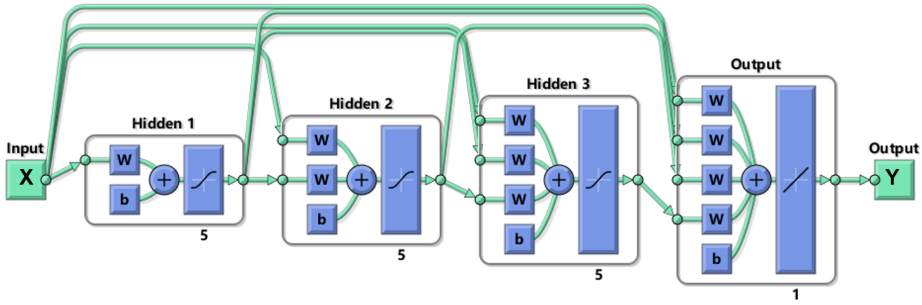


Figure 4.4: The cascade-forward networks representation

4.2.3 Surrogate Model Generation

Using adaptive sampling, the input sample space **X** and the output sample space **Y** are obtained. Then we use ANN to build the surrogate model based on **X** and **Y**. There are many different types of networks in ANN. Here, we use the so called cascade-forward networks [57] to train the networks which determine the mapping from the input **X** to the output **Y**. The cascade-forward networks consist of a series of layers of which the first layer has a connection from the network input **X**. Each subsequent layer has connections from the input and every previous layer and the final layer produces the network's output. The cascade-forward networks has a main advantage of flexibility for any kind of input to output mapping. Hence it can be suitably implemented for a black-box model like HYSYS in this case. In addition, the cascade-forward networks does not include the feedback loops in the networks so that it can save unnecessary computational expense. Therefore it is suitable to be used for surrogate model generation of this HYSYS model process. In this case, we use 3 hidden layers and each hidden layer contains 5 neurons. The configuration of the used cascade-forward networks is shown in the Figure 4.4.

We used MATLAB to build the networks which utilize the Levenberg-Marquardt

algorithm [58] to train the cascade-forward networks. The resulting network functions are essentially the surrogate models we desired. The surrogate models are generated based on $\mathbf{Y}_{s,abs}$ and $\mathbf{Y}_{s,diff}$ respectively, the corresponding resultant functions are saved as net_{abs} and net_{diff}. Hence the output data $\mathbf{Y}$ can be calculated by the functions with the input data $\mathbf{X}$ as

$$\mathbf{Y} = net_{diff}(\mathbf{X}) \tag{4.11}$$

The function net_{abs} and net_{diff} are then validated in Chapter 5.

Chapter 5

Model Validation

Two surrogate models are generated based on different identifications of variables. The model net_{diff} is generated with the variables identified in Table C.6 and the model net_{abs} is generated based on the variables in Table C.5. With the same input validation space, the models are validated by comparing their output data with the output data extracted from the HYSYS model. Validation figures are drawn to show the overall performance of the surrogate models. Furthermore, the mean, median and max relative deviations between the the output data of surrogate models and the HYSYS model are also calculated for validation. At the end, we compared the computation expenses for constructing the two surrogate models.

5.1 Model Validation for net_{diff}

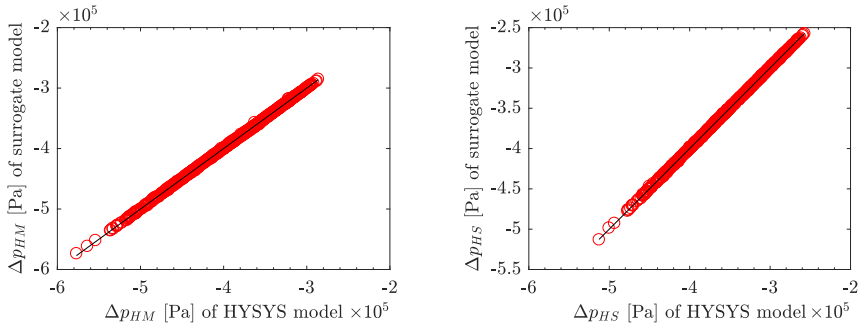
The resultant surrogate model net_{diff} of the HEx part needs to be validated to check if it has the ability to substitute the original HYSYS model with required accuracy. Hence a new sample space of input variables with 1000 points is firstly designed using the LHS sampling method, denoted as $\mathbf{X}_v$. As mentioned in the

Section 4.2.1, there are 13 input variables and 16 output variables in the HEx part, hence $\mathbf{X}_v$ is a 1000-by-13 matrix with each row corresponding to one sample point data and columns corresponding to the input variables. Using $\mathbf{X}_v$ as the input sample space, the output sample space of surrogate model and HYSYS model can be obtained, which are denoted as $\mathbf{Y}_{v,s}$ and $\mathbf{Y}_{v,H}$ respectively. Both $\mathbf{Y}_{v,s}$ and $\mathbf{Y}_{v,H}$ are 1000-by-16 matrices with each row corresponding to one sample point data and columns corresponding to output variables.

We validate the surrogate model by plotting the results of output variables of both the surrogate model and the HYSYS model as shown in Figure 5.1, 5.2 and 5.3, where the red circles represent the resultant points of surrogate model and the black line represents the HYSYS model results. Therefore, if it the red circles fits the black line exactly, the surrogate model is considered has the ability to replace the HYSYS model.

As shown in the Figure 5.1, the surrogate model has fairly good performance for the variables Δp_{HM}, ΔT_{HM}, Δp_{HS} and ΔT_{HS}. As for the fractional factor α, the validation plots are shown in the Figure 5.2. The surrogate model also has a good performance for α. The validation plots for the pressures and temperatures in the three recycle loop streams N_{HR1}, N_{HR2} and N_{HR3} are shown in the Figure 5.3.

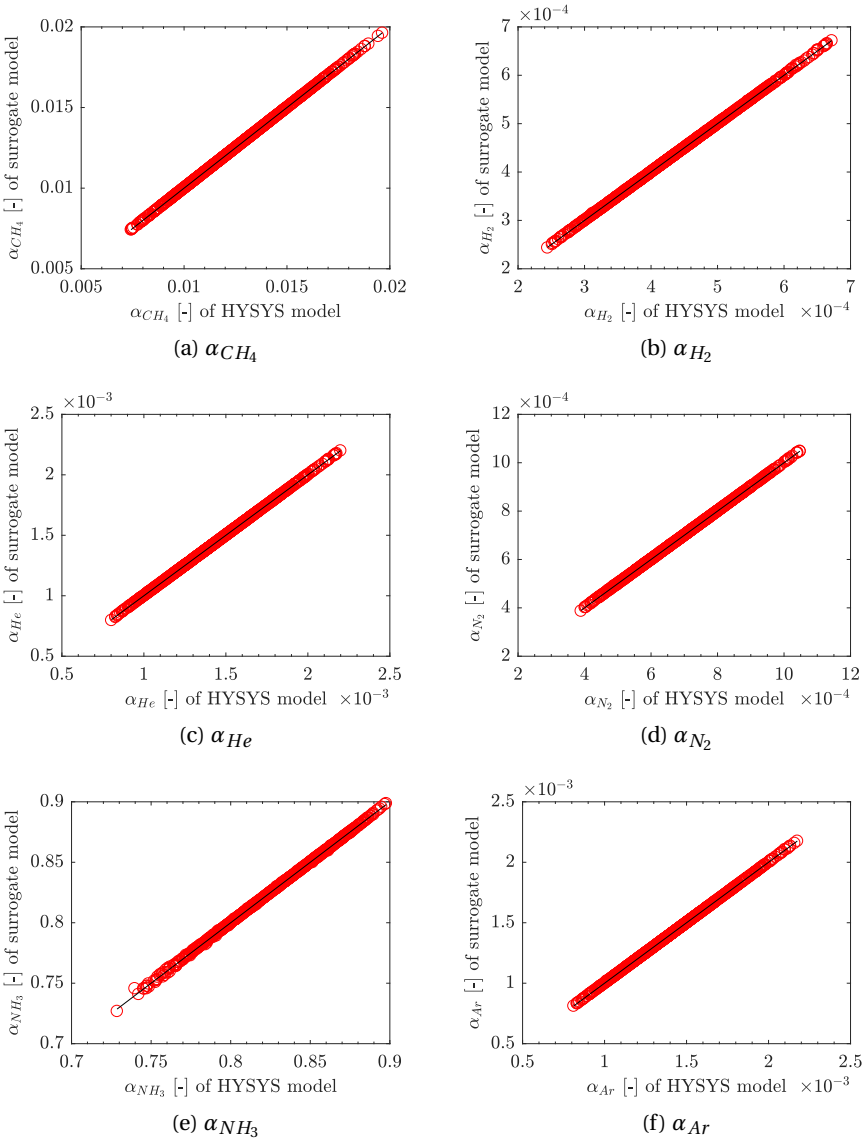
As shown in the Figure 5.3, the temperatures T_{HR1}, T_{HR2} and T_{HR3} can be predicted using the surrogate model with satisfied accuracy. However, the pressure differences Δp_{HR1}, Δp_{HR2} and Δp_{HR3} can not be predicted with satisfied accuracy using the surrogate model. The reasons could be that the pressure differences have fluctuating variations. Hence unstable fitting is present, indicating the surrogate model can not predict Δp_{HR1}, Δp_{HR2} and Δp_{HR3} stably. As mentioned in the Section 3.4.1, the identification could affect the resulting surrogate model. In Section 5.2, the surrogate models can have a good perfor-



(a) Δp_{HM} (b) ΔT_{HM}

(c) Δp_{HS} (d) ΔT_{HS}

Figure 5.1: Validation of Δp_{HM}, ΔT_{HM}, Δp_{HS} and ΔT_{HS} for net_{diff}

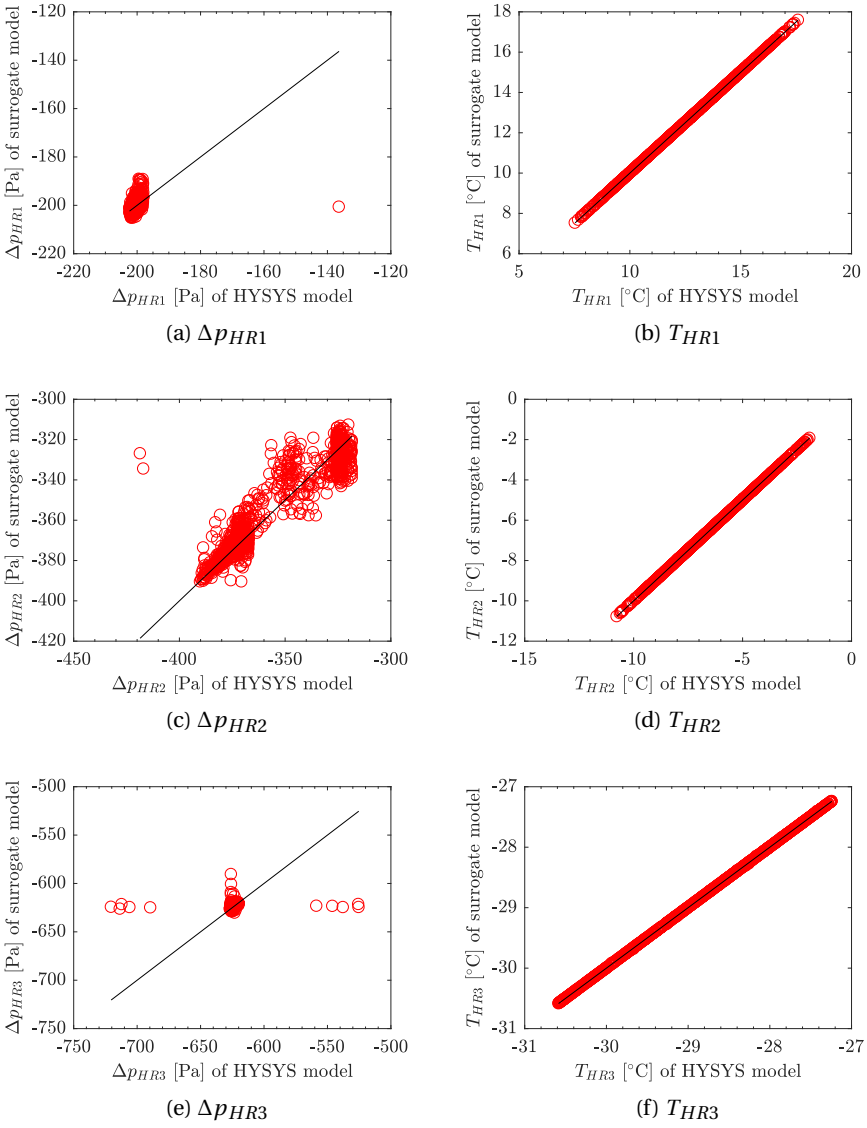


(a) α_{CH_4} (b) α_{H_2}

(c) α_{He} (d) α_{N_2}

(e) α_{NH_3} (f) α_{Ar}

Figure 5.2: Validation of the fractional factors α for net_{diff}



(a) Δp_{HR1} (b) T_{HR1}

(c) Δp_{HR2} (d) T_{HR2}

(e) Δp_{HR3} (f) T_{HR3}

Figure 5.3: Validation of Δp_{HR1}, T_{HR1}, Δp_{HR2}, T_{HR2}, Δp_{HR3} and T_{HR3} for net_{diff}

mance with the absolute pressures of p_{HR1}, p_{HR2} and p_{HR3} as output variables instead of the pressure differences.

To show the deviations more clearly, for each output variable m, we calculated relative deviations of the $n = 1000$ sample points and obtain the mean values by $\varepsilon_{mean,m}$ and the maximum values denoted by $\varepsilon_{max,m}$ as expressed in the Equations (5.1). The median value of the relative deviations are also found out which is denoted as $\varepsilon_{med,m}$. The resulting relative deviations are listed in the Table 5.1. From the table it can been observed that $\varepsilon_{max,m}$ of most variables are fairly small. But the $\varepsilon_{max,m}$ of Δp_{HR1}, Δp_{HR2} and Δp_{HR3} are 47.179%, 21.904% and 18.886% respectively, which are certainly not acceptable. Hence the surrogate model net_{diff} can not be used to replace the original HYSYS model.

$$\varepsilon_{mean,m} = \frac{\sum_{n=1}^{1000} |Y_{v,s}(n,m) - Y_{v,H}(n,m)|}{1000} \tag{5.1a}$$

$$\varepsilon_{max,m} = \max_{n} |Y_{v,s}(n,m) - Y_{v,H}(n,m)| \tag{5.1b}$$

5.2 Model Validation for net_{abs}

In this section, the resultant surrogate model net_{abs} is generated with the absolute output variables defined in $\mathbf{Y}_{s,abs}$. And it is validated using the same approach as mentioned in Section 5.1. The validation plots are shown in Figure 5.4, 5.5 and 5.6. And the calculated deviations are listed in Table 5.2.

Comparing Figure 5.6 with Figure 5.3. It can be seen that the prediction of p_{HR1}, p_{HR2} and p_{HR3} using net_{abs} have much better performance than prediction of Δp_{HR1}, Δp_{HR2} and Δp_{HR3} using net_{diff}. The improvements can also be observed by comparing the Table 5.1 and Table 5.2, the $\varepsilon_{max,m}$ for the

Output variable $\mathbf{Y}_{s,diff}$	$\varepsilon_{max,m}$ [%]	$\varepsilon_{mean,m}$ [%]	$\varepsilon_{med,m}$ [%]
Δp_{HM}	1.6818	0.14376	0.10983
ΔT_{HM}	0.94493	0.063332	0.038143
Δp_{HS}	2.4543	0.11496	0.087914
ΔT_{HS}	0.81318	0.043036	0.022058
α_{H_2}	0.86128	0.060409	0.040706
α_{NH_3}	0.80756	0.02981	0.018972
α_{He}	0.56454	0.050002	0.034496
α_{Ar}	0.60886	0.03899	0.026946
α_{N_2}	0.65954	0.060827	0.039767
α_{CH_4}	0.35584	0.029217	0.019267
T_{RH1}	0.12508	0.0043488	0.0032052
Δp_{RH1}	47.179	0.75797	0.54603
T_{RH2}	0.071066	0.0075646	0.0061685
Δp_{RH2}	21.904	1.613	1.0349
T_{RH3}	0.058634	0.001476	0.00070335
Δp_{RH3}	18.886	0.26541	0.079716

Table 5.1: The relative deviations between the surrogate model net_{diff} and the HYSYS model for output variables

three pressures are reduced from 47.179%, 21.904% and 18.886% to 0.0084084%, 0.031511% and 0.058145%. Hence net_{abs} has a good performance to predict p_{HR1}, p_{HR2} and p_{HR3}. The predictions of other variables also have satisfied performance. The different performance for net_{diff} and net_{abs} indicates that the variable identification may have significant influence on surrogate model generation.

5.3 Comparison of net_{diff} and net_{abs}

As mentioned in the previous sections, a lot of approaches are utilized to reduce the variables and the adaptive sampling is used to minimize the sample points required to generate the surrogate models. These strategies are implemented

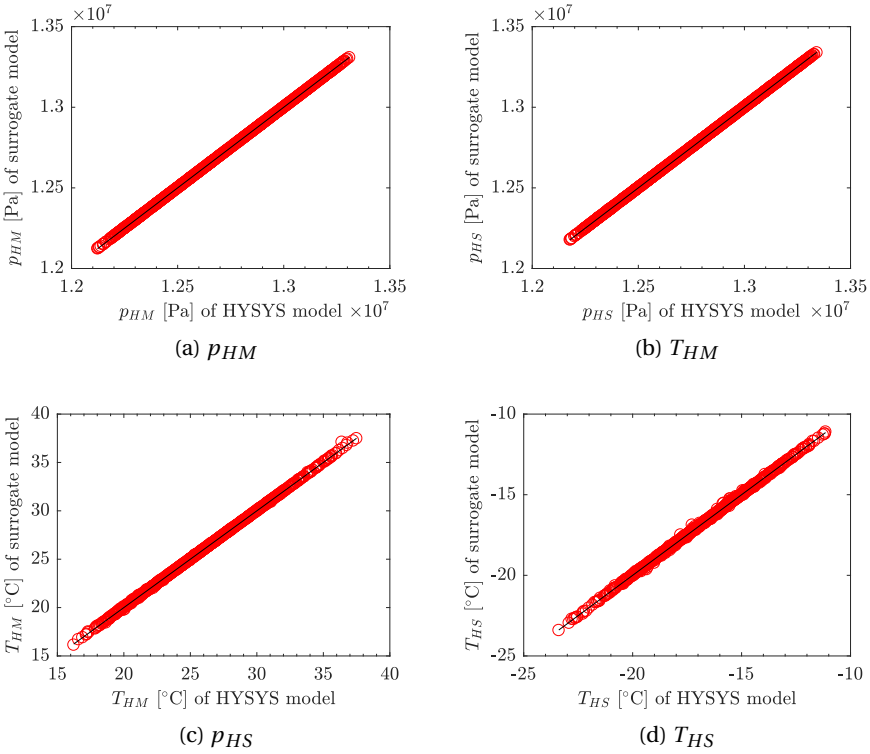
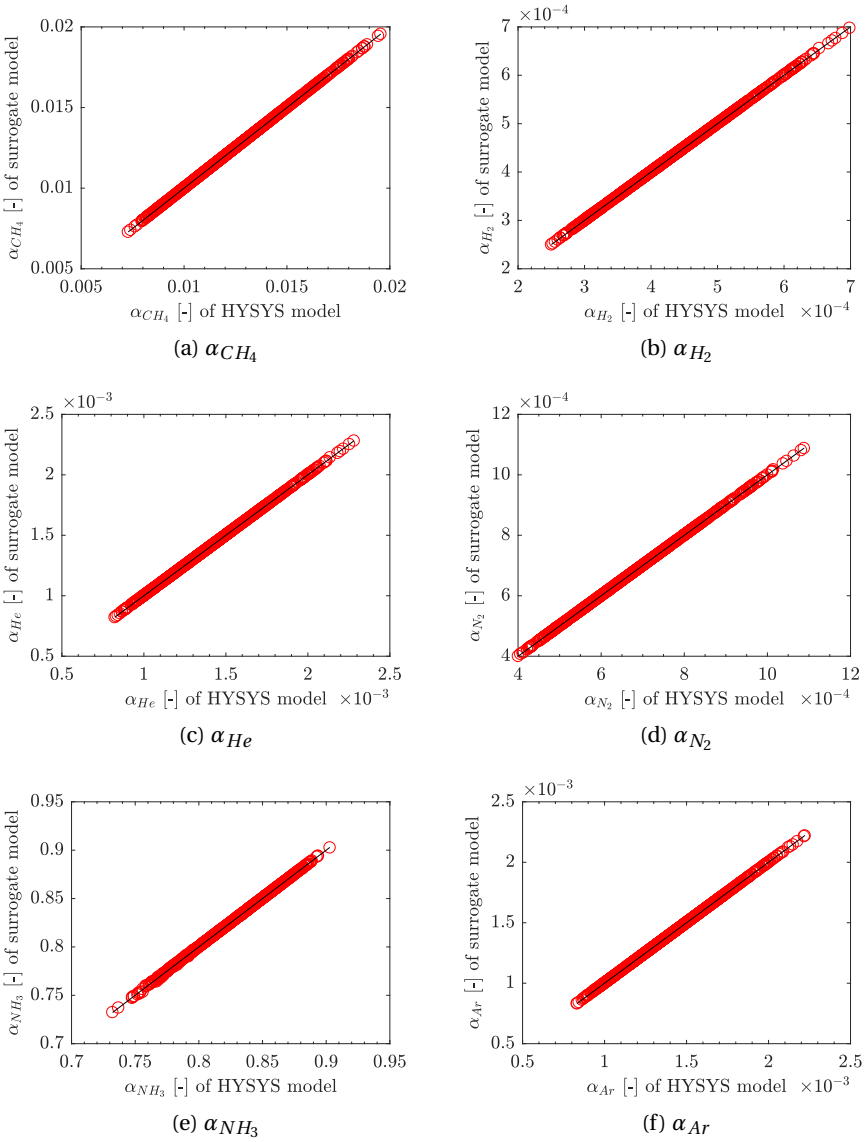


Figure 5.4: Validation of p_{HM}, T_{HM}, p_{HS} and T_{HS} for net_{abs}

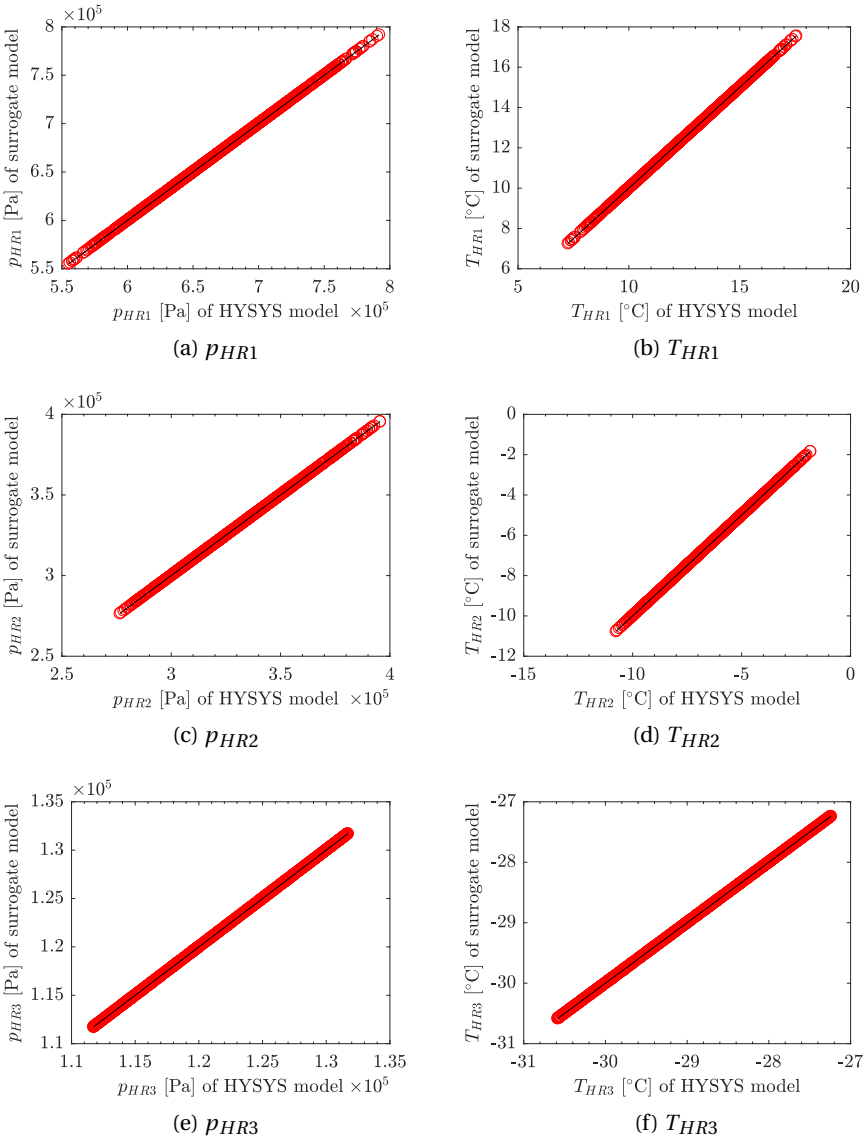


(a) α_{CH_4} (b) α_{H_2}

(c) α_{He} (d) α_{N_2}

(e) α_{NH_3} (f) α_{Ar}

Figure 5.5: Validation of the fractional factors α for net_{abs}



(a) p_{HR1} (b) T_{HR1}

(c) p_{HR2} (d) T_{HR2}

(e) p_{HR3} (f) T_{HR3}

Figure 5.6: Validation of p_{HR1}, T_{HR1}, p_{HR2}, T_{HR2}, p_{HR3} and T_{HR3} for net_{abs}

Output variable $\mathbf{Y}_{s,abs}$	$\varepsilon_{max,m}$ [%]	$\varepsilon_{mean,m}$ [%]	$\varepsilon_{med,m}$ [%]
p_{HM}	0.014595	0.0022376	0.0015808
T_{HM}	1.9882	0.084529	0.049647
p_{HS}	0.038062	0.0032719	0.0024392
T_{HS}	2.3274	0.1613	0.096058
α_{H_2}	0.4368	0.040509	0.028473
α_{NH_3}	0.34788	0.02375	0.015088
α_{He}	0.41258	0.042658	0.030454
α_{Ar}	0.30183	0.038395	0.02696
α_{N_2}	0.34068	0.037824	0.026545
α_{CH_4}	0.40724	0.014016	0.0086241
T_{RH1}	0.031611	0.0034433	0.0023714
p_{RH1}	0.0084084	$1.4397 \cdot 10^5$	$3.7482 \cdot 10^6$
T_{RH2}	0.16463	0.01224	0.0087845
p_{RH2}	0.031511	0.0022238	0.0015407
T_{RH3}	0.03914	0.00009259	0.00049691
p_{RH3}	0.058145	0.00059877	0.00030716

Table 5.2: The relative deviations between the surrogate model net_{abs} and the HYSYS model for output variables

to decrease the computational expense as much as possible. We evaluate the computational expense mainly by the total time used to construct the surrogate models, including sampling time and model generation time. The number of required sample points determined by adaptive sampling and the construction time for both surrogate models net_{diff} and net_{abs} are compared in Table 5.3. It can be observed that the construction of net_{diff} requires more sample points and hence takes more construction time than net_{abs}. The reason could be that net_{diff} is constructed based on the set $\mathbf{Y}_{s,diff}$ with the variable differences. The variable differences actually contain both input and output data information, indicating they can provide the information of the process more efficiently than absolute output variables. Hence, to generate the surrogate model for the same process, less variable differences are required than absolute vari-

Surrogate model	net_{diff}	net_{abs}
Required sample points	520	790
Construction time [s]	$1.57 \cdot 10^3$	$2.09 \cdot 10^3$

Table 5.3: The required sample points and construction time of two surrogate models

ables. However, the drawbacks of using variable differences have been shown in Section 5.1. The resulting surrogate model could be not stable for some variable differences which vary not significantly. Therefore, variable identification is crucial for construction of surrogate models and there is trade-off between ensuring accuracy and saving computational expense when constructing surrogate models.

As for the accuracy of the two model, net_{abs} has a better performance to predict the pressures p_{HR1}, p_{HR2} and p_{HR3} as discussed in Section 5.1. For net_{diff}, the $\varepsilon_{max,m}$ of the variables are all below 0.5% except Δp_{HM} and Δp_{HS}, of which $\varepsilon_{max,m}$ are 1.6818% and 2.4543% respectively. Meanwhile, the $\varepsilon_{max,m}$ of p_{HM} and p_{HS} are fairly small, being 0.014595% and 0.038062% respectively. Hence if we replace Δp_{HM} and Δp_{HS} by p_{HM} and p_{HS} for surrogate model generation, the resultant model function might have good performance to predict all the variables. This could be done in the future work.

Chapter 6

Summary and Recommendations for Further Work

In this chapter, we summarize the work in this thesis and give recommendations for future work.

6.1 Summary and Conclusions

In this thesis, the existing ammonia synthesis process modeled in HYSYS is aimed to be optimized. The process consists of four interconnected sections, the reaction section, the makeup section, the separation section and the refrigeration section. The whole process is described in Chapter 2.

Since the HYSYS simulator is a black-box model, indicating the derivative information are not available, it is difficult to utilize some optimization solvers to optimize the process in the HYSYS model. Hence, the surrogate model is introduced to address this issue. The relevant theories are mentioned in Chapter 3.

The surrogate models are construed in Chapter 4. First, the variables defin-

ing the process are classified by input variables and output variables. Then the variables are reduced as many as possible using the dependency relationships. After the input and output variables are determined by variable identification, adaptive sampling is implemented to sample the input variables to obtain the input sample space with minimized number of sample points. The resultant input sample data is imported to HYSYS to obtain the corresponding output sample data. However, the HYSYS model of S-R section was not able to calculate the corresponding output samples due to convergence issues. In order to address this issue, we divided the S-R section furthermore into the HEx part, the separator part and the refrigeration section. Then we use the same approach to construct a surrogate model for the HEx part. We used two different variable identifications to define the output variables, of which one used the absolute output variables and the other one used variables differences. Sample spaces of both cases are obtained and surrogate models are generated using artificial neural network. The resulting surrogate models are validated by comparing their output prediction with the HYSYS' output results and the deviations are calculated in Chapter 5.

It can be concluded that the surrogate models can be efficiently constructed based on the approach used in this thesis. Using the adaptive sampling, the number of sample points required to generate surrogate models can be successfully minimized, which simplified the model generation effectively. In this approach, the variable identification is extremely crucial in surrogate model construction. It can affect both the computational expense of surrogate model generation and accuracy of resultant surrogate models. Generally, the variable differences can save the computational expense but results in less accurate surrogate models than absolute variables. The approach used in this thesis can also be implemented to construct the surrogate models for other simulators.

6.2 Recommendations for Further Work

In the short term, the surrogate models for other parts of the S-R section need to be generated based on the same approach. After model generation, the individual models will be combined to form the complete separation section. However, since the decomposition of the S-R is due to convergence issues in HYSYS, it is possible that the similar convergence issues arise when combining the individual models. If the convergence issues occur, a possible solution is to decompose the S-R section in another way and construct surrogate models for new decomposed parts. If the model of the S-R section is obtained successfully by combining individual models, then surrogate models of the reaction section and makeup section need to be constructed. By combining the models of all the sections, the complete model for the whole process can be obtained so that the process can be optimized using some optimization solvers.

Appendix A

Nominal Conditions and Variation Ranges of Variables

This appendix includes the nominal conditions and variation ranges of corresponding variables in the S-R section and the HEx part.

Manipulated variables	Nominal conditions	Upper bounds	Lower bounds
$m_{BFW,2}$ [tonne/h]	545.2	907.4	544.5
$m_{BFW,3}$ [tonne/h]	399.2	500	300
$m_{BFW,4}$ [tonne/h]	1500	1875	1125
r_{Sp-1} [-]	0.535	0.54	0.53
RPM_{AC-2} [RPM]	173	300	100
RPM_{C2} [RPM]	7739	10180	7000
ΔT_{HEx-1} [°C]	0.068	0.090	0.068

Table A.1: The variation ranges of manipulated variables in the S-R section

Inlet variables	Nominal conditions	Upper bounds	Lower bounds
$N_{NH_3,RH}$ [kmole/h]	3713.2681	+20%	-20%
$N_{CH_4,RH}$ [kmole/h]	1274.7321	+20%	-20%
$N_{H_2,RH}$ [kmole/h]	15350.1440	+20%	-20%
$N_{H_2O,RH}$ [kmole/h]	0.0002	+20%	-20%
$N_{He,RH}$ [kmole/h]	345.1839	+20%	-20%
$N_{Ar,RH}$ [kmole/h]	1658.7192	+20%	-20%
$N_{N_2,RH}$ [kmole/h]	4800.2236	+20%	-20%
T_{RH} [°C]	53.73	63.73	43.73
p_{RH} [bar]	131.2	136.2	126.2
$N_{NH_3,MS}$ [kmole/h]	27.2754	+20%	-20%
$N_{CH_4,MS}$ [kmole/h]	0.003526	+20%	-20%
$N_{H_2,MS}$ [kmole/h]	0.01384	+20%	-20%
$N_{H_2O,MS}$ [kmole/h]	1.4474	+20%	-20%
$N_{He,MS}$ [kmole/h]	0.0001	+20%	-20%
$N_{Ar,MS}$ [kmole/h]	0.0003	+20%	-20%
$N_{N_2,MS}$ [kmole/h]	0.009486	+20%	-20%
T_{MS} [°C]	-34.38	-24.38	-44.38
p_{MS} [bar]	15.12	20.12	10.12

Table A.2: The variation ranges of inlet variables in the S-R section

Output variables	Nominal conditions	Output variables	Nominal conditions
$N_{NH_3,HM}$ [kmole/h]	629.1933	$N_{NH_3,SM}$ [kmole/h]	131.0039
$N_{CH_4,HM}$ [kmole/h]	1259.1391	$N_{CH_4,SM}$ [kmole/h]	0.6625
$N_{H_2,HM}$ [kmole/h]	15343.6344	$N_{H_2,SM}$ [kmole/h]	0.2768
$N_{H_2O,HM}$ [kmole/h]	0.0000	$N_{H_2O,SM}$ [kmole/h]	0.0000
$N_{He,HM}$ [kmole/h]	344.7035	$N_{He,SM}$ [kmole/h]	0.02043
$N_{Ar,HM}$ [kmole/h]	1656.42654	$N_{Ar,SM}$ [kmole/h]	0.09749
$N_{N_2,HM}$ [kmole/h]	4797.0018	$N_{N_2,SM}$ [kmole/h]	0.1370
T_{HM} [°C]	26.28	T_{SM} [°C]	-17.03
p_{HM} [bar]	127.3	p_{SM} [bar]	127.7
$N_{NH_3,S}$ [kmole/h]	3.0845	$N_{NH_3,R1}$ [kmole/h]	2838.1790
$N_{CH_4,S}$ [kmole/h]	4.3073	$N_{CH_4,R1}$ [kmole/h]	0.02283
$N_{H_2,S}$ [kmole/h]	5.8286	$N_{H_2,R1}$ [kmole/h]	0.000003622
$N_{H_2O,S}$ [kmole/h]	0.0000	$N_{H_2O,R1}$ [kmole/h]	1.4476
$N_{He,S}$ [kmole/h]	0.3718	$N_{He,R1}$ [kmole/h]	0.000009673
$N_{Ar,S}$ [kmole/h]	1.7305	$N_{Ar,R1}$ [kmole/h]	0.00006962
$N_{N_2,S}$ [kmole/h]	2.7377	$N_{N_2,R1}$ [kmole/h]	0.00001695
T_S [°C]	-14.53	T_{R1} [°C]	-29.56
p_S [bar]	15.11	p_{R1} [bar]	1.1749
$N_{NH_3,R2}$ [kmole/h]	152.7128	$N_{NH_3,R3}$ [kmole/h]	5.9285
$N_{CH_4,R2}$ [kmole/h]	0.7142	$N_{CH_4,R3}$ [kmole/h]	9.8879
$N_{H_2,R2}$ [kmole/h]	0.001711	$N_{H_2,R3}$ [kmole/h]	0.4216
$N_{H_2O,R2}$ [kmole/h]	0.0000	$N_{H_2O,R3}$ [kmole/h]	0.0000
$N_{He,R2}$ [kmole/h]	0.001125	$N_{He,R3}$ [kmole/h]	0.08748
$N_{Ar,R2}$ [kmole/h]	0.005849	$N_{Ar,R3}$ [kmole/h]	0.4607
$N_{N_2,R2}$ [kmole/h]	0.002584	$N_{N_2,R3}$ [kmole/h]	0.3568
T_{R2} [°C]	-33.03	T_{R3} [°C]	0.4084
p_{R2} [bar]	1.013	p_{R3} [bar]	13.59

Table A.3: The nominal conditions of output variables in the S-R section

Input variables	Nominal conditions	Upper bounds	Lower bounds
$N_{NH_3,RH}$ [kmole/h]	3700.9387	+20%	-20%
$N_{CH_4,RH}$ [kmole/h]	1270.4051	+20%	-20%
$N_{H_2,RH}$ [kmole/h]	15300.8893	+20%	-20%
$N_{H_2O,RH}$ [kmole/h]	0.00024	+20%	-20%
$N_{He,RH}$ [kmole/h]	344.0365	+20%	-20%
$N_{Ar,RH}$ [kmole/h]	1653.2608	+20%	-20%
$N_{N_2,RH}$ [kmole/h]	4784.5849	+20%	-20%
T_{RH} [°C]	53.73	500	300
p_{RH} [bar]	131.2	1875	1125
T_{RH1} [°C]	12.14	-	-
p_{RH1} [bar]	6.698	7.698	5.698
T_{RH2} [°C]	-6.147	-	-
p_{RH2} [bar]	3.337	3.837	2.737
T_{RH3} [°C]	-28.92	-	-
p_{RH3} [bar]	1.2202	1.3202	1.1202
$m_{BFW,2}$ [°C]	725.9	907.4	544.5
r_{SP-1} [-]	0.535	0.54	0.53

Table A.4: The variation ranges of input variables in the HEx part

Output variables	Nominal conditions	Output variables	Nominal conditions
$N_{NH_3,HM}$ [kmole/h]	625.8417	$N_{NH_3,HS}$ [kmole/h]	625.8417
$N_{CH_4,HM}$ [kmole/h]	1254.8587	$N_{CH_4,HS}$ [kmole/h]	1254.8587
$N_{H_2,HM}$ [kmole/h]	15294.4030	$N_{H_2,HS}$ [kmole/h]	15294.4030
$N_{H_2O,HM}$ [kmole/h]	0.0000	$N_{H_2O,HS}$ [kmole/h]	0.0002
$N_{He,HM}$ [kmole/h]	343.5579	$N_{He,HS}$ [kmole/h]	343.5579
$N_{Ar,HM}$ [kmole/h]	1650.9762	$N_{Ar,HS}$ [kmole/h]	1650.9762
$N_{N_2,HM}$ [kmole/h]	4781.3751	$N_{N_2,HS}$ [kmole/h]	4781.3751
T_{HM} [°C]	12.61	T_{HS} [°C]	12.61
p_{HM} [bar]	12.61	p_{HS} [bar]	12.61
T_{HR1} [°C]	12.61	p_{HR1} [bar]	6.696
T_{HR2} [°C]	-6.168	p_{HR2} [bar]	3.333
T_{HR3} [°C]	-28.85	p_{HR3} [bar]	1.2140

Table A.5: The nominal conditions of output variables in the HEx part

Appendix B

Process Flow Diagrams for Separation-Refrigeration Section

This appendix includes the PFDs for the S-R section.

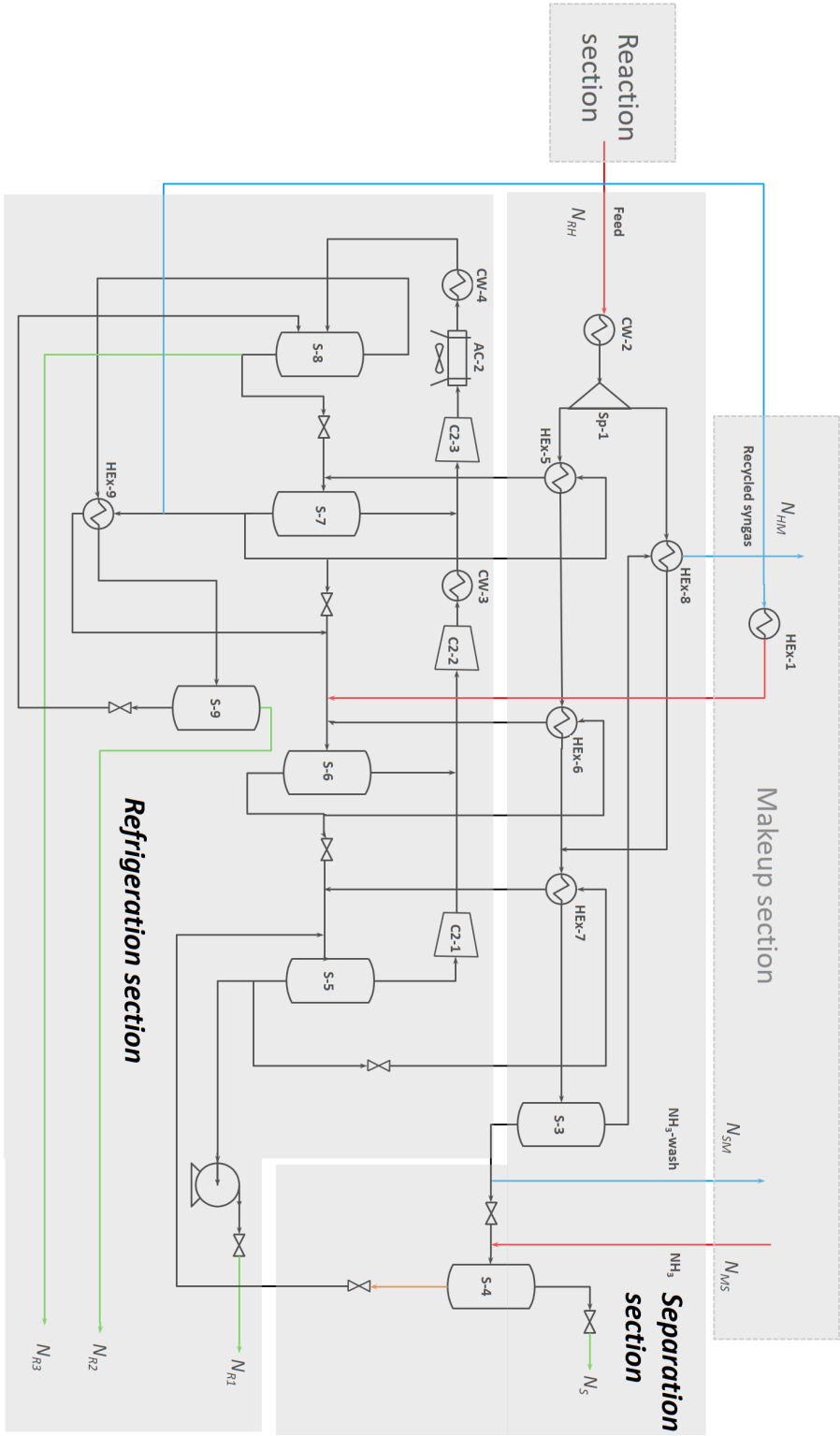


Figure B.1: The PFD of the S-R section

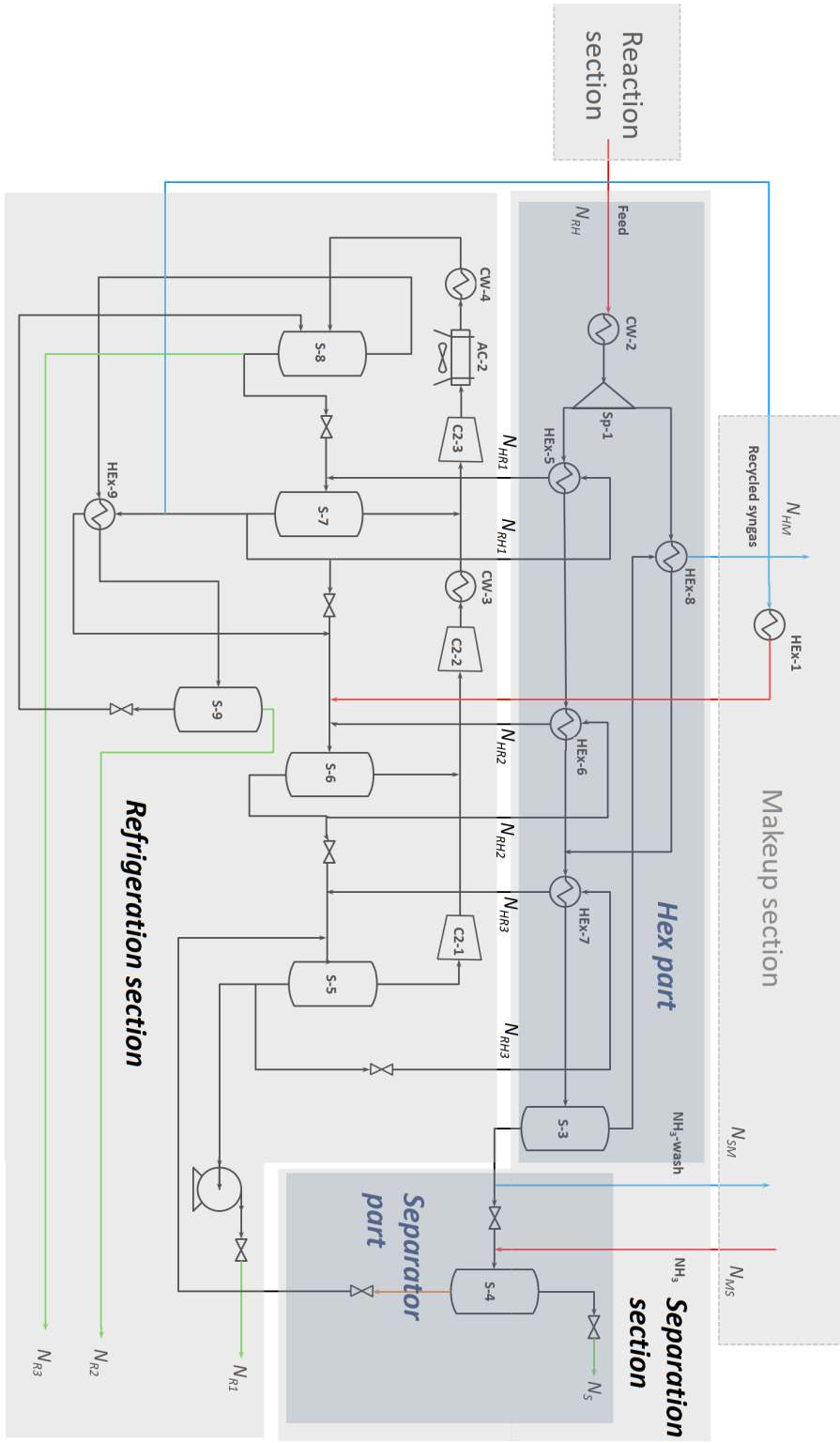


Figure B.2: The PFD of the decomposed S-R section

Appendix C

Variable Definitions

This appendix includes the variables defined in the S-R section and HEx part.

$\mathbf{I}_h$	Inlet stream variables: Manipulated variables:	$N_{i,j}$, T_j, p_j $m_{BFW,2}$, $m_{BFW,3}$, $m_{BFW,4}$, r_{Sp-1}, RPM_{AC-2}, RPM_{C2}, ΔT_{HEx-1}
$\mathbf{Y}_h$	Outlet stream variables:	$N_{i,k}$, T_k, p_k
	Number of variables: Total: Input: Output:	 79 25 54
		$i \in \{NH_3, CH_4, H_2, H_2O, He, Ar, N_2\}$ $j \in \{RH, MS\}$ $k \in \{HM, SM, S, R1, R2, R3\}$

Table C.1: The sets of $\mathbf{I}_h$ and $\mathbf{Y}_h$ defining the S-R section

$\mathbf{U}_s$	Inlet stream variables:	N_{RH}:	$N_{NH_3,RH}$, $N_{CH_4,RH}$, $N_{H_2,RH}$, $N_{He,RH}$, $N_{Ar,RH}$, $N_{N_2,RH}$, T_{RH}, p_{RH},
		N_{MS}:	$N_{NH_3,MS}$, $N_{H_2,MS}$, $N_{H_2O,MS}$, $N_{N_2,MS}$, T_{MS}, p_{MS}
	Manipulated variables:		$m_{BFW,2}$, $m_{BFW,3}$, $m_{BFW,4}$, r_{Sp-1}, RPM_{AC-2}, RPM_{C2}, ΔT_{HEx-1}
$\mathbf{Y}_{s,abs}$	Outlet stream variables:	N_{HM}:	T_{HM}, p_{HM},
		N_{SM}:	T_{SM}, p_{SM}
		N_S:	T_S, p_S
		N_{R1}:	T_{R1}, p_{R1}
		N_{R2}:	T_{R2}, p_{R2}
		N_{R3}:	T_{R3}, p_{R3}
		α:	α_{NH_3}, α_{CH_4}, α_{H_2}, α_{He},α_{Ar}, α_{N_2},
		$\beta 1$:	β_{1,NH_3}, β_{1,CH_4}, β_{1,H_2}, $\beta_{1,He}$,$\beta_{1,Ar}$, β_{1,N_2},
		$\beta 2$:	β_{2,NH_3}, β_{2,CH_4}, β_{2,H_2}, $\beta_{2,He}$,$\beta_{2,Ar}$, β_{2,N_2},
		$\beta 3$:	β_{3,NH_3}, β_{3,CH_4}, β_{3,H_2}, $\beta_{3,He}$,$\beta_{3,Ar}$, β_{3,N_2},
		$\beta 4$:	β_{4,NH_3}, β_{4,CH_4}, β_{4,H_2}, $\beta_{4,He}$,$\beta_{4,Ar}$, β_{4,N_2},
	Number of variables:		
	Total:	63	
	Input:	21	
	Output:	42	

Table C.2: The sets of $\mathbf{U}_s$ and $\mathbf{Y}_{s,abs}$ identified in the S-R section

$\mathbf{U}_s$	Inlet stream variables:	N_{RH}:	$N_{NH_3,RH}$, $N_{CH_4,RH}$, $N_{H_2,RH}$, $N_{He,RH}$, $N_{Ar,RH}$, $N_{N_2,RH}$, T_{RH}, p_{RH},
		N_{MS}:	$N_{NH_3,MS}$, $N_{H_2,MS}$, $N_{H_2O,MS}$, $N_{N_2,MS}$, T_{MS}, p_{MS}
	Manipulated variables:		$m_{BFW,2}$, $m_{BFW,3}$, $m_{BFW,4}$, r_{Sp-1}, RPM_{AC-2}, RPM_{C2}, ΔT_{HEx-1}
$\mathbf{Y}_{s,diff}$	Outlet stream variables:	N_{HM}:	ΔT_{HM}, Δp_{HM},
		N_{SM}:	ΔT_{SM}, Δp_{SM}
		N_S:	ΔT_S, Δp_S
		N_{R1}:	ΔT_{R1}, Δp_{R1}
		N_{R2}:	ΔT_{R2}, Δp_{R2}
		N_{R3}:	ΔT_{R3}, Δp_{R3}
		α:	α_{NH_3}, α_{CH_4}, α_{H_2}, α_{He},α_{Ar}, α_{N_2},
		$\beta 1$:	β_{1,NH_3}, β_{1,CH_4}, β_{1,H_2}, $\beta_{1,He}$,$\beta_{1,Ar}$, β_{1,N_2},
		$\beta 2$:	β_{2,NH_3}, β_{2,CH_4}, β_{2,H_2}, $\beta_{2,He}$,$\beta_{2,Ar}$, β_{2,N_2},
		$\beta 3$:	β_{3,NH_3}, β_{3,CH_4}, β_{3,H_2}, $\beta_{3,He}$,$\beta_{3,Ar}$, β_{3,N_2},
		$\beta 4$:	β_{4,NH_3}, β_{4,CH_4}, β_{4,H_2}, $\beta_{4,He}$,$\beta_{4,Ar}$, β_{4,N_2},
	Number of variables:		
	Total:	63	
	Input:	21	
	Output:	42	

Table C.3: The sets of $\mathbf{U}_s$ and $\mathbf{Y}_{s,diff}$ identified in the S-R section

$\mathbf{I}_h$	Inlet stream variables: T_j, p_j	$N_{i,RH}$, T_{RH}, p_{RH},
	Manipulated variables:	$m_{BFW,2}$, r_{Sp-1}
$\mathbf{Y}_h$	Outlet stream variables:	$N_{i,k}$, T_k, p_k T_l, p_l
	Number of variables:	
	Total:	41
	Input:	17
	Output:	24
		$i \in \{NH_3, CH_4, H_2, H_2O, He, Ar, N_2\}$ $j \in \{RH1, RH2, RH3\}$ $k \in \{HM, HS\}$ $l \in \{HR1, HR2, HR3\}$

Table C.4: The sets of $\mathbf{I}_h$ and $\mathbf{Y}_h$ defining the HEx part

$\mathbf{U}_s$	Inlet stream variables: Manipulated variables:	$N_{i,RH}$, T_{RH}, p_{RH}, p_j $m_{BFW,2}$, r_{Sp-1}
$\mathbf{Y}_{s,abs}$	Outlet stream variables: α:	T_k, p_k T_l, p_l α_i
	Number of variables:	
	Total:	29
	Input:	13
	Output:	16
		$i \in \{NH_3, CH_4, H_2, He, Ar, N_2\}$ $j \in \{RH1, RH2, RH3\}$ $k \in \{HM, HS\}$ $l \in \{HR1, HR2, HR3\}$

Table C.5: The sets of $\mathbf{U}_s$ and $\mathbf{Y}_{s,abs}$ identified in the HEx part

$\mathbf{U}_s$	Inlet stream variables:	$N_{i,RH}$, T_{RH}, p_{RH}, p_j
	Manipulated variables:	$m_{BFW,2}$, r_{Sp-1}
$\mathbf{Y}_{s,diff}$	Outlet stream variables:	ΔT_k, Δp_k T_l, Δp_l α_i
	Number of variables:	
	Total:	29
	Input:	13
	Output:	16
		$i \in \{NH_3, CH_4, H_2, He, Ar, N_2\}$ $j \in \{RH1, RH2, RH3\}$ $k \in \{HM, HS\}$ $l \in \{HR1, HR2, HR3\}$

Table C.6: The sets of $\mathbf{U}_s$ and $\mathbf{Y}_{s,diff}$ in the HEx part

Bibliography

[1] Timothy W Simpson, JD Poplinski, Patrick N Koch, and Janet K Allen. Metamodels for computer-based engineering design: survey and recommendations. *Engineering with computers*, 17(2):129–150, 2001.

[2] Carlos A Henao and Christos T Maravelias. Surrogate-based superstructure optimization framework. *AIChE Journal*, 57(5):1216–1232, 2011.

[3] Ruichen Jin, Wei Chen, and Timothy W Simpson. Comparative studies of metamodelling techniques under multiple modelling criteria. *Structural and multidisciplinary optimization*, 23(1):1–13, 2001.

[4] Gérard Bloch and Thierry Denoeux. Neural networks for process control and optimization: Two industrial applications. *ISA transactions*, 42(1):39–51, 2003.

[5] Daniel R Stull. Vapor pressure of pure substances. organic and inorganic compounds. *Industrial & Engineering Chemistry*, 39(4):517–540, 1947.

[6] Alison Cozad, Nikolaos V Sahinidis, and David C Miller. Learning surrogate models for simulation-based optimization. *AIChE Journal*, 60(6):2211–2227, 2014.

[7] Lorenz T Biegler, Ignacio E Grossmann, and Arthur W Westerberg. Systematic methods for chemical process design. 1997.

[8] Minimizing the complexity of surrogate models for optimization. In Zdravko Kravanja and Miloš Bogataj, editors, *26th European Symposium on Computer Aided Process Engineering*, volume 38 of *Computer Aided Chemical Engineering*, pages 289 – 294. Elsevier, 2016.

[9] Donald R Jones. A taxonomy of global optimization methods based on response surfaces. *Journal of global optimization*, 21(4):345–383, 2001.

[10] G Gary Wang and Songqing Shan. Review of metamodeling techniques in support of engineering design optimization. *Journal of Mechanical design*, 129(4):370–380, 2007.

[11] Arne Stolbjerg Drud. Conopt—a large-scale grg code. *ORSA Journal on computing*, 6(2):207–216, 1994.

[12] Andreas Wächter and Lorenz T Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical programming*, 106(1):25–57, 2006.

[13] Philip E Gill, Walter Murray, and Michael A Saunders. Snopt: An sqp algorithm for large-scale constrained optimization. *SIAM review*, 47(1):99–131, 2005.

[14] Luis Miguel Rios and Nikolaos V Sahinidis. Derivative-free optimization: a review of algorithms and comparison of software implementations. *Journal of Global Optimization*, 56(3):1247–1293, 2013.

[15] Donald R Jones, Matthias Schonlau, and William J Welch. Efficient global

optimization of expensive black-box functions. *Journal of Global optimization*, 13(4):455–492, 1998.

[16] José A Caballero and Ignacio E Grossmann. Rigorous flowsheet optimization using process simulators and surrogate models. *Computer Aided Chemical Engineering*, 25:551–556, 2008.

[17] Nestor V Queipo, Raphael T Haftka, Wei Shyy, Tushar Goel, Rajkumar Vaidyanathan, and P Kevin Tucker. Surrogate-based analysis and optimization. *Progress in aerospace sciences*, 41(1):1–28, 2005.

[18] Jack PC Kleijnen. *Statistical tools for simulation practitioners*. Marcel Dekker, Inc., 1986.

[19] Alexander Forrester, Andras Sobester, and Andy Keane. *Engineering design via surrogate modelling: a practical guide*. John Wiley & Sons, 2008.

[20] John Winsor Pratt, Howard Raiffa, and Robert Schlaifer. *Introduction to statistical decision theory*. MIT press, 1995.

[21] Ilya M Sobol. *A primer for the Monte Carlo method*. CRC press, 1994.

[22] Michael D McKay, Richard J Beckman, and William J Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 42(1):55–61, 2000.

[23] TJ Santner, BJ Williams, and WI Notz. The design and analysis of computer experiments springer-verlag. *New York. 283pp*, 2003.

[24] Boxin Tang. Orthogonal array-based latin hypercubes. *Journal of the American statistical association*, 88(424):1392–1397, 1993.

[25] Kenny Q Ye. Orthogonal column latin hypercubes and their application in computer experiments. *Journal of the American Statistical Association*, 93(444):1430–1439, 1998.

[26] Thomas M Cioppa and Thomas W Lucas. Efficient nearly orthogonal and space-filling latin hypercubes. *Technometrics*, 49(1):45–55, 2007.

[27] Fasheng Sun, Min-Qian Liu, and Dennis KJ Lin. Construction of orthogonal latin hypercube designs with flexible run sizes. *Journal of Statistical Planning and Inference*, 140(11):3236–3242, 2010.

[28] Richard F Gunst. Response surface methodology: process and product optimization using designed experiments, 1996.

[29] Norman Richard Draper, Harry Smith, and Elizabeth Pownell. *Applied regression analysis*, volume 3. Wiley New York, 1966.

[30] Timothy W Simpson, Timothy M Mauery, John J Korte, and Farrokh Mistree. Kriging models for global approximation in simulation-based multidisciplinary design optimization. *AIAA journal*, 39(12):2233–2241, 2001.

[31] SJ Yakowitz and F Szidarovszky. A comparison of kriging with nonparametric regression methods. *Journal of Multivariate Analysis*, 16(1):21–53, 1985.

[32] Carlos A Henao and Christos T Maravelias. Surrogate-based process synthesis. *Computer Aided Chemical Engineering*, 28:1129–1134, 2010.

[33] P Hajela and L Berke. Neural networks in structural analysis and design: an overview. *Computing Systems in Engineering*, 3(1-4):525–538, 1992.

[34] Nira Dyn, David Levin, and Samuel Rippa. Numerical procedures for surface fitting of scattered data by radial functions. *SIAM Journal on Scientific and Statistical Computing*, 7(2):639–659, 1986.

[35] Jerome H Friedman. Multivariate adaptive regression splines. *The annals of statistics*, pages 1–67, 1991.

[36] Timothy W Simpson, Jesse Peplinski, Patrick N Koch, and Janet K Allen. On the use of statistics in design and the implications for deterministic computer experiments. *Design Theory and Methodology-DTM'97*, pages 14–17, 1997.

[37] Wei Chen, Janet K Allen, Dimitri N Mavris, and Farrokh Mistree. A concept exploration method for determining robust top-level specifications. *Engineering Optimization+ A35*, 26(2):137–158, 1996.

[38] Anthony Giunta and Layne Watson. A comparison of approximation modeling techniques-polynomial versus interpolating models. In *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, page 4758, 1998.

[39] Russell R Barton. Metamodels for simulation input-output relations. In *Proceedings of the 24th conference on Winter simulation*, pages 289–299. ACM, 1992.

[40] William J Welch, Robert J Buck, Jerome Sacks, Henry P Wynn, Toby J Mitchell, and Max D Morris. Screening, predicting, and computer experiments. *Technometrics*, 34(1):15–25, 1992.

[41] Carlos Gershenson. Artificial neural networks for beginners. *arXiv preprint cs/0308031*, 2003.

[42] Richard Lippmann. An introduction to computing with neural nets. *IEEE Assp magazine*, 4(2):4–22, 1987.

[43] Simon Haykin and Neural Network. A comprehensive foundation. *Neural Networks*, 2(2004):41, 2004.

[44] Kürt Meert and Marcel Rijckaert. Intelligent modelling in the chemical process industry with neural networks: a case study. *Computers & chemical engineering*, 22:S587–S593, 1998.

[45] Fabiano AN Fernandes. Optimization of fischer-tropsch synthesis using neural networks. *Chemical engineering & technology*, 29(4):449–453, 2006.

[46] Iqbal M Mujtaba, Norashid Aziz, and MA Hussain. Neural network based modelling and control in batch reactor. *Chemical Engineering Research and Design*, 84(8):635–644, 2006.

[47] J Straus and S Skogestad. Variable reduction for surrogate modelling.

[48] Clement John Adkins. *Equilibrium thermodynamics*. Cambridge University Press, 1983.

[49] Joseph Mauk Smith. *Introduction to chemical engineering thermodynamics*. PhD thesis, Rensselaer Polytechnic Institute, 1975.

[50] Nicholas W Tschoegl. *Fundamentals of equilibrium and steady-state thermodynamics*. Elsevier, 2000.

[51] Timothy W Simpson. *A concept exploration method for product family design*. PhD thesis, Georgia Institute of Technology, 1998.

[52] S.K. Thompson. *Sampling*. CourseSmart Series. Wiley, 2012.

[53] Stephen Leary, Atul Bhaskar, and Andy Keane. Optimal orthogonal-array-based latin hypercubes. *Journal of Applied Statistics*, 30(5):585–598, 2003.

[54] Julian Straus and Sigurd Skogestad. Adptive sampling, in prep.

[55] Sijmen De Jong. Simpls: an alternative approach to partial least squares regression. *Chemometrics and intelligent laboratory systems*, 18(3):251–263, 1993.

[56] George Wm Thomson. The antoine equation for vapor-pressure data. *Chemical reviews*, 38(1):1–39, 1946.

[57] Howard Demuth and Mark Beale. Matlab neural network toolbox user's guide version 6. the mathworks inc. 2009.

[58] Jorge J Moré. The levenberg-marquardt algorithm: implementation and theory. In *Numerical analysis*, pages 105–116. Springer, 1978.