

Bi-functional Optimization of a Fixed-bed Reactor using Surrogate Models

Shaghayegh Bashiri, Elham Yasari*

Department of Chemical Engineering, Faculty of Engineering, Ferdowsi University of
Mashhad, Mashhad, Iran
elhamyasari@ferdowsi.um.ac.ir

Abstract

In this paper, a bi-functional optimization problem is applied in the two most common surrogate models, Multilayer Perceptron and Radial Basis Function, both of which are based on LHS sampling. Furthermore, they are compared in terms of mean square error rate to determine the more accurate one. To examine the two types of alternative models, the initial multivariate database is required which has been produced by the modeling of catalyst fixed-bed reactor. The mathematical model has been done using MATLAB 2017b. The results indicate that Multilayer Perceptron is much more efficient than the other one. Non dominated sorting genetic algorithm is used as an optimization algorithm to maximize and minimize the two components of reactor products: dimethyl ether and carbon dioxide. Since the objective functions are in contrast to each other, model optimization has various sets of answers instead of a single solution. The best set of optimization solution is Pareto front. The result of optimization demonstrates that the exact and the best surrogate model do in fact match.

Key words: Multi-objective optimization, Modeling, Latin Hypercube Sampling, Dimethyl ether, Surrogate models

1. Introduction

To recognize the minimum and maximum of the objective function, an optimization algorithm is employed. The optimization problems include three elements that are necessary to determine; Firstly, there is the objective function that is minimized or maximized in the problem. Secondly, variables, and Thirdly, there is the condition of being constrained or

unconstrained. When there is no equal or unequal constrained in the optimization problem, an unconstrained problem is, thus, left to be dealt with. When there is more than one objective function, a multi-objective function is used to find the minimum or maximum. The aim of solving multi-objective function is to receive a set of insuperable answers which is called the Pareto front [1].

Surrogate models are used as function approximation to estimate the unknown process, in case there is a lack of data or receiving experimental data is not available. Since accessing to commercial simulator code is costly and nearly impossible, and more importantly only input and output are considered in optimization methods, black-box method can be a good alternative for simulator code[2].

The set of data sampling plays a crucial role in the accuracy of the research results, therefore, the strategies of data collection and usage is important. To select data from a population of unknown parameters, sampling is used due to the fact of its being time-consuming. Low cost and fast data collection are the major advantages of sampling methods[3]. Latin hypercube sampling is one of the best technique to select data when there is a lack of experimental data. The LHS involves the design space division in 'n'(number of sampling) equal partitions and selecting a random data point in each partition[2]. Predicting the behavior of reactors has been the debate of engineering sciences throughout the last decade. Artificial intelligence extracts pattern, knowledge and algorithms based on numerical data[4]. Artificial Neural Network is one of the subcategories of AI which is used to predict, model, optimize and classify by using any number of sampling data[5].

Although there are various networks such as ADALINE, ANFIS, MLP and RBF which have been previously used, MLP and RBF are the most popular and appropriate models to represent inputs and outputs widely. Feed forward network concludes MLP and RBF. ANN is able to learn the relationship between phenomenon equations. A Neural Network has neurons, each containing an input, output and one or more hidden layers. Input neurons receive and process input signal, then transfer the output to the other neurons. The major feature of this ANN is the capability to learn and generalize[6]. Using pairs of data set, called training set, the network parameters are adjusted in order to achieve an acceptable model. Tuning parameters of MLP are the number of hidden layer and neurons[7]. The value of spread and number of neurons in the network need to be determined before training with RBF[8]. The common activation function for MLP and RBF are sigmoid and Gaussian, respectively.

2.Experimental

There are two ways to acquire the data: Experimental data and data extracted from modeling or simulations. Requiring an input-output data set is crucial for creating alternative models. In this case, due to a lack of experimental data, the fixed-bed reactor of dimethyl ether produced from syngas is modeled in MATLAB software. Later, among four sampling methods, LHS is selected to generate input data by 200 sample size. By this set of data, two types of proxy models are created: RBF and MLP. One of the applications of surrogate models is in the optimization problem. In this paper, multi-objective optimization is used to maximize and minimize the dimethyl ether and carbon dioxide, respectively. To evaluate the performance of

the best alternative model in the optimization case, the optimization results of both proxy models and mathematical models are compared.

2.1 LHS

Generating a multi-variable sample by the LHS method is relatively simple. In this technique, all samples related to each variable must be stratified optimally. This technique divides the boundary of every input variable into equal distribution. A sample data is placed in every row and column preventing clusters to formulate. In this method, the range of change of each variable is covered completely. The necessity of this method in natural resources and related topics is to consider the limit values in the maps or input variables that are not considered in the conventional sampling of these values[9].

2.2 Neural network

The Neural network is the combination of neurons which are allocated to different layers. The input layer is the first layer that receives information from the surrounding. This layer depends on independent variables. Hidden layers must convert inputs to outputs through various processes. The last layer is related to the response of the dependent variable to the stimulation of independent variables in the neural network. A feature of neurons is that each and every one works independently. They also need local information, meaning that all information for the output of neurons are defined in neurons themselves and their inputs.

2.3 Radial basis function

Radial basis function is a tool for interpolation data that is expanded. RBFs are the most commonly used radial functions. Radial structure and the ability to learn are the features of this type of function. Among all activation functions which are used in the output layer such as Multi-Quadric Function and Thin Plate Spline Function, Gaussian is the most common (σ is spread) and most applicable, because when r is willing to be infinite, the function will become zero[10].

$$\varphi_r = \exp\left(-\frac{r^2}{2\sigma^2}\right)$$
$$\text{if } r \rightarrow \infty \text{ then } \varphi_r \rightarrow 0$$

There are two major differences between MLP and RBF:

- 1) In RBF, rather than internal multiplication between input vector and weights, the distances between them are calculated.
- 2) In RBF, the bias multiplies the distances between input vector and weights, but in MLP, bias is added to the multiplication of the distances between the input vector and the weights.

2.4 REAL MODEL

To produce the output data and objectively evaluate the performance of the different types of proxy models, a mathematical model is required. The reactions that occur in a pseudo-homogeneous fixed-bed reactor of direct synthesis of dimethyl ether ($\text{CH}_3\text{-O-CH}_3$) from syngas

(CO + H₂) are to be considered as an exact model. All the reactions are exothermic, as mentioned in Table 1.

Table 1: Reactions of dimethyl synthesis production from syngas

Reactions	Type of Reaction	Heat of reactions
$CO + 2H_2 \leftrightarrow CH_3OH$	Methanol Synthesis	-90.85
$CO + H_2O \leftrightarrow CO_2 + H_2$	Water Gas Shift	-41.1
$CO_2 + 3H_2 \leftrightarrow CH_3OH + H_2O$	Methanol synthesis	-50.1
$2CH_3OH \leftrightarrow CH_3OCH_3 + H_2O$	Methanol Dehydration	-23.4

The model is evaluated according to Hu et al. (2008) work in the same condition to establish a model for sampling.

The mass conservation equations were derived from the following assumptions: the ratio of length to diameter is large [11] so radial dispersion can be neglected. The internal and external diffusion are omitted and the gas was assumed as an ideal gas. By writing mass and energy balances, the following equations are utilized.

Mass equations:

$$\frac{\varepsilon C_{u_s}}{L_r} \frac{\partial x_i}{\partial z} = \rho_b \left(\sum_{j=1}^{n_r} v_{ij} r_j - x_i \sum_{i=1}^{n_c} \sum_{j=1}^{n_r} v_{ij} r_j \right)$$

Where ε is bed porosity, U_s is gas velocity in the reactor (m/s), L_r is reactor length (m), x_i is mole fraction of i component, ρ_b is bed density (kg/m³), r_j is reaction rate (mol/skg).

Thermal equations:

$$\frac{\rho_g C_{p_g} u_s}{L_r} \frac{\partial T}{\partial z} = \left(\rho_b \left(\sum_{i=1}^{n_r} (-\Delta H)_i r_i \right) + \frac{4U_{eff}}{d_{out}} (T - T_c) * \text{config}_{coeff} \right)$$

C_{p_g} , ρ_g , U_{eff} , d_{out} , T_c , u_g are specific heat capacity of gas (J/ kgK), gas density (kg/m³), overall heat transfer coefficient (W/m² K), external diameter of tube (mm), shell temperature (K) and gas velocity in the tube (m/s), respectively.

Ergun equation is used to calculate pressure drop:

$$\frac{dP}{dz} = -10^{-5} L_r \left(1.75 + 150 \left(\frac{1 - \varepsilon}{d_p \rho_g u_s / \mu_g} \right) \right) \frac{u_g^2 \rho_g}{d_p} \frac{1 - \varepsilon}{\varepsilon^3}$$

Where μ_g is gas phase viscosity (kg/ms) and d_p is equivalent diameter of catalyst particle (mm).

The present mathematical model should respect the following conditions:

$$Z=0 \quad C_i = C_0 \quad T_i = T_0 \quad P_i = P_0$$

$$Z=L \quad C_i = C_0 \quad T_i = T_0 \quad P_i = P_0$$

The purpose of evaluation is to judge the efficiency and accuracy of the techniques. There are different variations of validation equations: the two most common, MSE and MAE, are introduced below. Since both error values have the same trend, MSE is selected as the base.

MSE

The general error between prediction and real output is calculated by mean square error.

$$MSE = \frac{1}{N} \sum (d_i - y_i)^2$$

MAE

The ideal value of mean absolute error is zero. Maximum and Minimum illustrate overestimate and underestimate the real value.

$$MAE = \frac{1}{N} \sum (d_i - y_i)$$

3. optimization

One of the most popular, evolutionary algorithms is the non-dominated sorting genetic algorithm. A Non-dominated sorting genetic algorithm is a problem-solving method that uses genetic science to solve problem models. This algorithm is a search method to find approximate solutions to optimization problems. The non-dominated sorting genetic algorithm deals with a population of solutions. The optimal solution is the one that maximizes the value of the fitness function, but if the current issue is to minimize a cost function, it's easy to adapt. The cost function simply must be converted to a fit function. Almost all conventional optimization methods search for a single point, while NSGA is always applied to a population of points. One of the challenges in multi-objective optimization compared to the single-objective optimization, which recognizes only one optimal point, is to use the model that has high accuracy in all points. Furthermore, the aim is to find a model that has an ability to generalize well. Optimization has three parts that need to be recognized before it being performed; objective functions, variables and constraints.

Objective functions

Each model has at least one objective function which consists of the minimum and the maximum; This function ranks the outcome.

Variables

The layout of system is expressed by a set of parameters, some of which are fixed while others are variable. There are two types of variables, continuous and discontinuous. Designing variables cannot be selected optionally, but they must satisfy the requirements of a set of finite and functional norms.

Constraints

When the function boundaries are limited, the optimization problem is said to have constraints. Therefore, the optimum values can occur just at the boundaries of constraints or between them. As in this paper, maximization of carbon dioxide and minimization of dimethyl ether are objective functions. The input temperature, input pressure, and shell temperature are the variables. The maximum temperature in the reactor necessary to prevent temperature runaway is known as the constraint.

Pareto Front

The set of points in the target space that do not overpower each other (the set of non-dominant points) is called the set of Pareto front. In multi-objective optimization problems, unlike single-

objective problems where there is only one optimal solution, all points of the Pareto front are possible answers for the optimization problem. None of the points on Pareto front is superior to the other, and depending on the choice of the user, one or more of the points can be selected as the final answer. The concept of Pareto front is that by maximizing dimethyl ether mole fraction, carbon dioxide will be increased too; Likewise, by minimizing the amount of carbon dioxide, the amount of dimethyl ether is reduced.

Non-dominated sorting genetic algorithm is one of the most widely used and powerful algorithms for solving multi-objective optimization problems. Its efficiency in solving various problems has been frequently proved. In multi-objective problems, there exists more than a single optimal solution, instead there are sets of solutions called effective non dominated solutions. Among these sets of finite solutions, there will be appropriate solutions that have acceptable performance for all purposes. Two operators have been added to the genetic algorithm to rank the answer: Non-dominated Ranking and Crowded Distance [12].

Non-dominated Solution:

Members of the population are organized according to the concept of non-dominance. In the non-dominated sorting method, the members of the population that are not superior to others are placed in the first front and excluded from the population. The remaining members, then, are reselected, placed in the second front, and temporarily excluded from the crowd. This process will continue until all members of the population are placed in different fronts.

Crowded Distance:

After allocating the points, another criterion is used to evaluate points on a front, and that is the crowded distance. Crowded distance is calculated separately for each member of the population. In order to calculate the crowded distance of a particular answer at a boundary, two things are taken into consideration: the largest rectangle in which that answer is placed, and its two neighbors which are placed on its both sides; Later, the sum of the rectangle's length and width is measured as the crowded distance for that particular answer.

4. Results and discussion

The set of differential equations of dimethyl ether production in the pseudo-homogeneous reactor were numerically solved with Runge-Kutta first and fifth-order methods in MATLAB. To evaluate the accuracy of the mathematical model, the CO₂, DME mole-fractions and temperatures at the end of the reactor are compared at the same condition (Table 2) between our modeling and analysis by Hu et al[13]. The model is reliable since the error is less than 5%.

Table 2: Comparison between Hu and our model

Mole fraction	Hu et al.	Model	Error %
CO ₂	0.066033	0.0637	3.53
DME	0.048857	0.0492	0.702
Temperature	245.928	246.036	0.043

To produce a set of data in MATLAB, input-output parameters and their ranges are determined in Table 3.

Table 3: Range of input-output variables

inputs	Min	Max	outputs
Inlet temperature(K)	468	538	Mole fraction
Shell temperature(K)	468	538	Max temperature(K)
Pressure(Kpa)	3000	8000	Pressure(Kpa)
Catalyst ratio	0.1	0.9	Mole-fraction of components

LHS Sampling

This method is one of the random classification methods. The LHS will sample random inputs and their permutations through a very controlled selection. There is only one sample in each column and row of space design. The plotting of three variables of temperature, shell temperature, and pressure in three dimensions based on LHS method is shown in Figure 1.

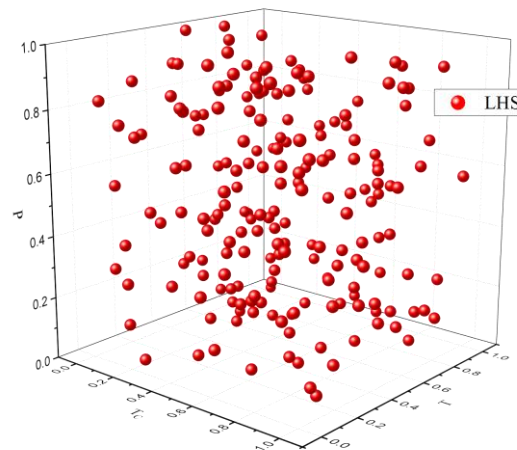


Figure 1: Three parameters of 50 samples are plotted together using the Latin cube sampling technique

Neural network fitting tool box in MATLAB is used for fitting data set. Inlet temperature, pressure, shell temperature and catalyst ratio are input parameters. Mole fraction of output components, pressure and max temperature are the output for networks. Three hidden layers with [20, 15, 10] neurons are calculated as the best state by trial and error. RBF is coded in MATLAB, too. The number of neurons in hidden layer and spread is tuned on as 20, 8, respectively. As mentioned before, both types of surrogate models are modeled by LHS sampling method with 200 sample size. As in Figure 2, MLP offers the lowest dispersion and does not present such extreme values as the RBF model does.

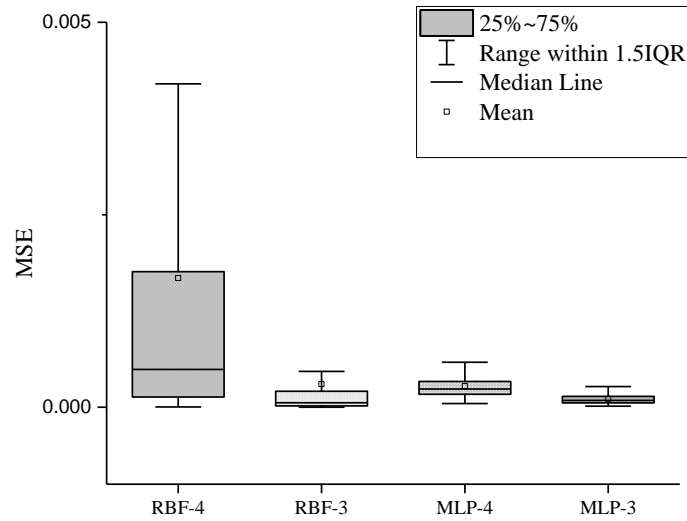


Figure 2: Comparison between RBF and MLP based on MSE in 200 sample size

In this section, the Pareto front is plotted by optimizing the exact model and placing the optimal values of the variables from the exact model into the alternative model. A comparative diagram of mole-fraction of the two components, carbon dioxide, and dimethyl ether, is drawn between the multilayer neural network model and the exact model in **Error! Reference source not found.** The objective functions and design constraints are expressed below. It should be noted that the most important constraint is to avoid temperature runaway.

$$\left\{ \begin{array}{l} \text{Max} : X_{DME} \\ \text{Min} : X_{CO2} \\ \text{subject to:} \\ 30\text{bar} < P < 80\text{bar} \\ 450\text{K} < T_{in}, T_{shell} < 550\text{K} \\ T_{max} < 548 \end{array} \right.$$

The Non Dominated Sorting Non-dominated sorting genetic algorithm is used to compare the optimization of the exact model structure with the alternative model. According to the Figure 3 both models have the same trend, and the relative error rate is very low.

There is a slight difference between the results of optimization based on mole fraction of components. This means that the alternative model can predict the values of the optimal variables correctly. Therefore, it can also be concluded that the alternative model has high accuracy (the error is less than 3%) in an optimization process.

In other words, checking out the proxy model is considered in multi-objective optimization. to illustrates the reliability of the model. The significance of Figure 3 is that the results of the objective functions of both models do match, when optimizing the exact model and proxy model.

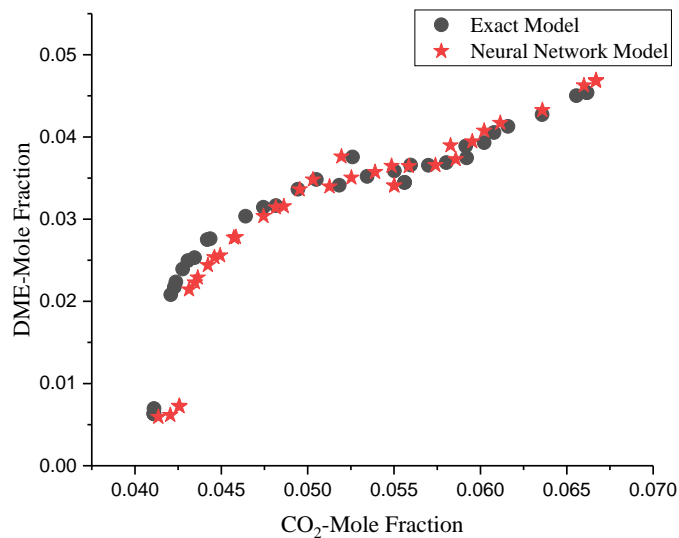


Figure 3: Comparing between MLP and Exact model objective functions at the same condition

One of the most important factors that have led researchers to optimize the use of the alternative model discussed before, is the time required for implementation. Although the time spent on network optimization and exact model in this study are short (less than an hour), Table 4 indicates a shorter time required for MLP optimization compared to the exact model. The exact model optimization is approximately 93 times slower than the selected alternative model, and this is one of the advantages of the alternative model.

Table 4: Runtime of optimization Model

Model	Time(second)
Exact Model	177.819878
Neural Network	1.8976

5. Conclusions

The set of differential equations of dimethyl ether production in the pseudo-homogeneous reactor were numerically solved with Runge-Kutta first and fifth-order methods in MATLAB. The study has provided results on the performance of two types of proxy models which are MLP and RBF. These models are formed by the 200 sets of data extracted from the best sampling method, LHS. These comparisons reveal that for any problem a particular alternative model is applicable, while a single model is not to be used for all problems.

Between both MLP and RBF models that have been tested, the MLP appears to give the best results in terms of high accuracy to the fixed-bed reactor of dimethyl ether production. Nevertheless, RBF is not fit for accurate estimation. The best application of these models is an optimization problem. With respect to efficiency, it is obvious that MLP is the winner. Thus, it can be concluded that an efficient model is a tool for optimizing this type of fixed-bed reactors. Comparing between these two types of meta models, the evaluation is based on MSE has been

the major topics of this paper which have been thoroughly explored. Also, in terms of optimization runtime, MLP is faster than the exact model.

References

- [1] Deb, Kalyanmoy. Multi-objective optimization using evolutionary algorithms. Vol. 16. John Wiley & Sons, 2001.
- [2] F. Jurecka, 'Robust design optimization based on metamodeling techniques', Technische Universität München, 2007.
- [3] H. Taherdoost, 'Sampling Methods in Research Methodology; How to Choose a Sampling Technique for Research', SSRN Journal, 2016.
- [4] Flasiński, Mariusz. Introduction to artificial intelligence. Springer, 2016.
- [5] H. Kukreja, 'AN INTRODUCTION TO ARTIFICIAL NEURAL NETWORK', vol. 1, no. 5, p. 5, 2016.
- [6] M. Nwadiugwu, 'NEURAL NETWORK, ARTIFICIAL INTELLIGENCE AND THE COMPUTATIONAL BRAIN', 2015.
- [7] Section of Agricultural Statistics, Department of Farm Engineering, Institute of Agricultural Sciences, Banaras Hindu University, Varanasi-221 005, India and H. G. R., 'Multilayer Perceptron Method of Artificial Neural Network for Classification of Farmers Based on Adoption of Drought Coping Mechanisms', Int. J. Pure App. Biosci., vol. 6, no. 2, pp. 1408–1414, May 2018.
- [8] Ch. S. K. Dash, A. K. Behera, S. Dehuri, and S.-B. Cho, 'Radial basis function neural networks: a topical state-of-the-art survey', Open Computer Science, vol. 6, no. 1, Jan. 2016.
- [9] S. Koziel, D. E. Ciaurri, and L. Leifsson, 'Surrogate-based methods', in Computational optimization, methods and algorithms, Springer, 2011, pp. 33–59.
- [10] A. G. Bors, 'Introduction of the Radial Basis Function (RBF) Networks', p. 8.
- [11] E. Yasari, 'A green industrial scale di-methyl ether reactor with aiming to CO₂ reduction: staging and multi-objective optimization approach', Journal of the Taiwan Institute of Chemical Engineers, vol. 81, pp. 110–118, Dec. 2017.
- [12] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan, 'A Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-objective Optimization: NSGA-II', in Parallel Problem Solving from Nature PPSN VI, vol. 1917, M. Schoenauer, K. Deb, G. Rudolph, X. Yao, E. Lutton, J. J. Merelo, and H.-P. Schwefel, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, pp. 849–858.
- [13] Y. Hu, Z. Nie, and D. Fang, 'Simulation and model design of pipe-shell reactor for the direct synthesis of dimethyl ether from syngas', Journal of Natural Gas Chemistry, vol. 17, no. 2, pp. 195–200, Jun. 2008.