Surrogate modeling applications in chemical and biomedical processes

Kazemzadeh Farizhandi, Amir Abbas

2017

Kazemzadeh Farizhandi, A. A. (2017). Surrogate modeling applications in chemical and biomedical processes. Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/72705>

<https://doi.org/10.32657/10356/72705>

NANYANG
TECHNOLOGICAL
UNIVERSITY

Surrogate modeling applications in chemical and biomedical processes

Amir Abbas Kazemzadeh Farizhandi

SCHOOL OF CHEMICAL AND BIOMEDICAL ENGINEERING

2017

Surrogate modeling applications in chemical and biomedical processes

Amir Abbas Kazemzadeh Farizhandi

School of Chemical and Biomedical Engineering

A thesis submitted to the Nanyang Technological University

In partial fulfillment of the requirement for the degree of

Doctor of Philosophy

2017

Abstract

Surrogate modeling is an efficient alternative for computation-intensive process simulations in engineering problems. Surrogate model is developed using experimental or computer data, which are collected from experiments or simulation runs. The use of surrogate model allows efficient and cost-effective computation for different applications. With this purpose, two systems: 1) particle size distribution (PSD) in gas-solid fluidized beds and 2) carrier-based dry powder inhalation (DPI) efficiency have been considered as case studies. In this study, artificial neural network (ANN) coupled with genetic algorithm (GA) was employed as a surrogate modeling tool.

PSD plays a crucial role in performance and operation of the fluidized bed. Since monitoring of the change in PSD in computational fluid dynamic (CFD) simulation is computationally expensive, PSD usually considers being constant during fluidization in CFD simulation. Therefore, surrogate modeling has been proposed as a fast and cheap computation method to estimate PSD during fluidization. Planetary ball milling is employed to derive descriptive parameters to account for the effect of material properties in the particle attrition process. Gas-solid fluidized bed experiments have been conducted to provide required data for surrogate model construction. The results show that the Rosin-Rammler (RR) distribution is able to describe the PSD reasonably well ($R\text{-square} > 0.97$) for fluidization and ball milling processes. Two ANN-GA models were developed based on the RR parameters (d and n) obtained from least-square fitting of experimental PSD results. $R\text{-square}$ values of leave-one-out cross-validation for the developed ANN-GA models were more than 0.9589 which

shows that the surrogate model can estimate PSD during fluidization reasonably well. With adding the developed surrogate model to CFD simulation, more accurate and reliable results can be provided in the simulation of gas-solid fluidized beds.

On the other hand, finding the effect of variables interaction on the efficiency of DPI by experiments is not possible because usually changing a variable will change the other variables inevitably. Therefore, ANN-GA approach as a surrogate model has been employed to evaluate the effect of different variables on DPI efficiency. In vitro aerosolization performance and drug delivery efficiency of a DPI are generally represented by emitted dose (ED) and fine particle fraction (FPF). Image analysis is employed to obtain various descriptive parameters for surface morphologies of carriers based on scanning electron microscopy (SEM) images. Variable selection is used to reduce the number of input variables needed for surrogate model development. R-square values of leave-one-out cross-validation for the developed surrogate models were more than 0.7546 in prediction of ED and FPF. Sensitivity analysis was also performed to determine the key variables affecting ED and FPF. With this developed model, one variable can be isolated and its effect on DPI efficiency can be evaluated. In fact, it provides a tool for better understanding of DPI formulation and it can be used for the design and optimization of DPI.

Acknowledgement

I would like to express my sincere thanks and appreciation to my supervisor, Dr. Lau Wai Man, Raymond for his invaluable guidance, support and suggestions. His knowledge, suggestions, and discussions help me to become a capable researcher. His encouragement also helps me to overcome the difficulties encountered in my research. I also want to thank my colleagues in the lab, for their generous help. I want to thank Dr. Wang ke for her explanation of the surrogate modeling, which saved me a lot of time, and Zhao, for his generous help in my experiments in fluidized bed. I am very grateful to my lovely wife who always supports me. Last but not least, I want to thank my parents in Iran, for their constant love and encouragement.

Table of content

Abstract	i
Acknowledgement	iii
Table of content	iv
List of figures	vii
List of tables	x
Chapter 1 Introduction	1
1.1 Background	1
1.2 Motivation of this research.....	4
1.3 Objectives and scope.....	8
1.4 Organization of the thesis.....	10
Chapter 2 Literature Survey	12
2.1 Review of surrogate modeling	12
2.2 Data distribution methods	17
2.3 Surrogate modeling techniques	20
2.4 Surrogate model fitting methods	26
2.5 Surrogate model validation and accuracy	27
2.6 Review of surrogate modeling applications in chemical engineering.....	29
Chapter 3 Modeling Techniques	33
3.1 Preface.....	33
3.2 Artificial neural network (ANN) as a surrogate modeling technique	33
3.3 Variables selection	35
3.4 Sensitivity analysis (SA)	37
3.5 Symbolic regression (SR)	41
3.6 Genetic algorithms (GA).....	43
3.7 Accuracy and validation of surrogate model.....	45
3.8 ANN-GA as an integrated approach for process modeling.....	46
3.9 Particle size distribution (PSD).....	52
Chapter 4 Modeling the change in particle size distribution in a gas-solid fluidized bed due to particle attrition using a hybrid artificial neural network- genetic algorithm approach	55
4.1 Preface.....	55
4.2 Experimental setup.....	58
4.3 Data collection	60
4.4 Design of the ANN model for prediction of PSD	61

4.5 Results and Discussion.....	62
4.5.1 Application of the Rosin–Rammmler model to the IBA particle size distribution analysis in fluidization.....	63
4.5.2 Accuracy and validation of surrogate model.....	65
4.5.3 Effect of glass beads on particle attrition	69
4.6 Summary	73
Chapter 5 Modeling of particle size distribution in a gas-solid fluidized bed by planetary ball milling results using a hybrid artificial neural network-genetic algorithm approach	75
5.1 Preface.....	75
5.2 Experimental setup.....	76
5.3 Data collection	77
5.4 Accuracy and validation of surrogate model.....	78
5.5 Genetic algorithms (GA) design for different purposes	78
5.6 Results and Discussion.....	78
5.6.1 Application of the Rosin–Rammmler model to the particle size distribution analysis in ball milling and fluidization	79
5.6.2 Determination of attrition related material properties by ball milling	81
5.6.3 Accuracy and Validation of ANN models	82
5.6.4 Symbolic regression (SR) of d and n	87
5.7 Summary	88
Chapter 6 Evaluation of hydroxyapatite size and morphology in dry powder inhalation for carrier-based pulmonary delivery formulations by response surface methodology	90
6.1 Preface.....	90
6.2 Dataset.....	92
6.3 Surface and shape analysis.....	94
6.4 Design of ANN	95
6.4 Genetic algorithms (GA) parameters	96
6.5 Results and discussion	96
6.5.1 Analysis of surface roughness.....	96
6.5.2 Selection of important variables.....	98
6.5.3 Design of the ANN model for prediction of FPF and ED.....	100
6.5.4 The sensitivity analysis of input variables on ED and FPF.....	101
6.5.5 Effects and interactions of various factors on ED and FPF	103
6.5.5.1 Effect of particle average size and size distribution on ED and FPF ..	103
6.5.5.2 Effect of flow rate and carrier-to-drug ratio on ED and FPF	105

6.5.5.3 Effect of surface morphology on ED and FPF.....	106
6.6 Summary	111
Chapter 7 Modeling of emitted dose and fine particle fraction in dry powder inhalation for carrier-based pulmonary delivery formulations by using neural networks and genetic algorithms.....	113
7.1 Preface.....	113
7.1 Dataset.....	113
7.2 Surface and shape analysis.....	115
7.3 Design of ANN	115
7.4 Genetic algorithms parameters.....	116
7.5 Results and discussion	117
7.5.1 Analysis of surface roughness.....	117
7.5.2 Selection of important variables.....	119
7.5.3 Design of the ANN model for prediction of FPF and ED.....	121
7.5.4 Sensitivity analysis of input variables.....	122
7.5.5 Effects of carrier materials and interactions of various factors on ED and FPF	123
7.5.5.1 Carrier materials.....	123
7.5.5.2 Effect of carrier particle average size and size distribution	124
7.5.5.2 Effect of carrier-to-drug ratio and drug particles average size.....	126
7.5.5.3 Effect of flow rate and carrier tap density.....	128
7.5.5.4 Effect of carrier surfaces morphology.....	130
7.5.6 Symbolic regression (SR)	134
7.6 Summary	136
Chapter 8 Conclusion and outlook.....	138
8.1 Conclusions.....	138
8.2 Outlooks	140
References.....	145
Appendix.....	188

List of figures

Figure 1. 1. Data to knowledge process by surrogate modeling.	3
Figure 2. 2. A typical structure for construction of a surrogate model.	13
Figure 3. 1. Artificial neural network structure.....	34
Figure 3. 2. Evolution flow of genetic algorithm.....	45
Figure 3. 3. The structure ANN-GA as a hybrid intelligent system model for the process modeling.....	49
Figure 3. 4. Effect of d and n on RR distribution.....	54
Figure 4. 1. Fluidized bed experimental setup.	60
Figure 4. 2. Artificial neural network structure for prediction of d (y1) and n (y2) as RR distribution parameters..	62
Figure 4. 3. Fitting of PSD using RR distribution function: a) Original IBA PSD with, b) Pure IBA PSD at time = 30 min, c) IBA PSD at time = 300 min with using 50% small glass beads	65
Figure 4. 4. Parity plots of experimental and predicted RR parameters values calculated by the models for training data	67
Figure 4. 5. Experimental data of IBA particle size with fitted and predicted RR distribution from ANN models	69
Figure 4. 6. Three-dimensional surfaces of ANN models for d as a function of the time and glass beads percentage at $d_0 = 0.9$ and $n_0 = 1.6$ for a) small glass beads; b) large glass beads	70
Figure 4. 7. Three-dimensional surfaces of ANN models for n as a function of the time and glass beads percentage at $d_0 = 0.9$ and $n_0 = 1.6$ a) small glass beads; b) large glass beads	71
Figure 5. 1. Fitting of PSD in ball milling using RR distribution function: a) Silica PSD at time = 108 min, b) Gypsum PSD at time = 120 min	80
Figure 5. 2. Fitting of PSD in fluidization using RR distribution function: a) Silica PSD at time = 240 min and $U_g = 1.3$ m/s, b) Activated carbon PSD at time = 300 min and $U_g = 0.73$ m/s	80
Figure 5. 3. a) Change of d and n in ball milling for gypsum, b) determination of B_d and B_n for gypsum.....	81
Figure 5. 4. Comparison of materials hardness with B_d and B_n	82
Figure 5. 5. Parity plots of experimental and predicted RR parameters values calculated by the models for testing data	83
Figure 5. 6. Experimental data of particle size with fitted and predicted RR distribution from ANN models for testing points: a) Silica PSD at time = 180 min and $U_g = 1.3$ m/s as the best prediction, b) Gypsum PSD at time = 1200 min and $U_g = 0.84$ m/s as the medium prediction, c) Activated carbon PSD at time = 60 min and $U_g = 1.23$ m/s as the worst prediction	85

Figure 5. 7. Experimental data of IBA particle size with fitted and predicted RR distribution from ANN models for IBA testing points: a) IBA PSD at time = 60 min and $U_g = 0.78$ m/s as the best prediction, b) IBA PSD at time = 180 min and $U_g = 0.78$ m/s as the medium prediction c) IBA PSD at time = 1200 min and $U_g = 0.78$ m/s as the worst prediction.....	86
Figure 5. 8. Parity plots of real and calculated d and n by the developed equations by GA SR.....	88
Figure 6. 1. A sample workflow of surface roughness analysis: SEM image, 2D, 3D surface plots, and surface properties	97
Figure 6. 2. Frequency of variable usage in models; a) ED, b) FPF	100
Figure 6. 3. Parity plots of experimental and predicted values based on training data a) ED; b) FPF	101
Figure 6. 4. Sensitivity analysis of input variables in prediction of (a) ED and (b) FPF	103
Figure 6. 5. Three-dimensional surfaces of ANN models for ED and FPF as a function of the HA particle average size and size standard deviation; a) ED; b) FPF	105
Figure 6. 6. Three-dimensional surfaces of ANN models for ED and FPF as a function of the flow rate and carrier-to-drug ratio; a) ED; b) FPF.....	106
Figure 6. 7. Three-dimensional surfaces of ANN models for ED and FPF as a function of the HA surface roughness variables; a) ED as a function of R_a and R_q ; b) ED as a function of SA and $MFOV$; c) FPF as a function of R_a and R_q ; c) FPF as a function of SA and FPO	107
Figure 6. 8. Relationship between FPO and peak distance.	111
Figure 7. 1. A sample workflow of surface property analysis.	118
Figure 7. 2. Frequency of variable usage in models; a) ED, b) FPF.	120
Figure 7. 3. Parity plots of experimental and predicted ED and FPF values calculated by the models for training data.	122
Figure 7. 4. Sensitivity analysis of input variables in prediction of (a) ED and (b) FPF.	123
Figure 7. 5. Three-dimensional surfaces of ANN models for ED and FPF as a function of the carrier particle average size and size standard deviation; a) ED; b) FPF.....	125
Figure 7. 6. Three-dimensional surfaces of ANN models for ED and FPF as a function of the carrier-to-drug ratio and drug particles average size; a) ED; b) FPF.	127
Figure 7. 7. Three-dimensional surfaces of ANN models for ED and FPF as a function of the flow rate and tap density; a) ED; b) FPF.	129
Figure 7. 8. Three-dimensional surfaces of ANN models for ED and FPF as a function of the carrier surface roughness variables; a) ED as a function of R_a and SA ; b) FPF as a function of R_q and R_a ; c) FPF as a function of FPO and SA	134

Figure 7. 9. Parity plots of real and calculated ED and FPF by the developed equations by GA SR, a) calculated ED versus experimental ED b) calculated FPF versus experimental FPF	136
Figure A1: Generation versus fitness value for ANN-GA in IBA fluidization (chapter 4)	188
Figure A2: Generation versus fitness value for ANN-GA in all materials fluidization (chapter 5).....	188
Figure A3: Generation versus fitness value for ANN-GA in HA carrier DPI (chapter 6)	189
Figure A4: Generation versus fitness value for ANN-GA in all carriers (chapter 7)	189

List of tables

Table 1.1. Different methods of design of experiment (data distribution), surrogate models, and model fitting.....	4
Table 4.1. Input and output variables their range of values	60
Table 4.2. GA parameters for ANN optimization.....	62
Table 4.3. Validation results of surrogate models.....	67
Table 5.1. Input and output variables and their range of values	77
Table 5.2. GA parameters for variable selection and ANN optimization	78
Table 5.3. Calculated materials attrition properties by ball milling.....	82
Table 5.4. Accuracy and validation results of ANN models.....	84
Table 6.1. Complete list of input and output variables considered in the study	93
Table 6.2. GA parameters for variable selection and ANN optimization	96
Table 6.3. Validation results of surrogate models.....	101
Table 6.4. Analysis of roughness parameters based on cropped carrier surface images	110
Table 7.1. General description of created database.....	114
Table 7.2. GA parameters for variable selection and ANN optimization	117
Table 7.3. Validation results of surrogate models.....	121
Table 7.4. Chemical structure and properties of carrier materials	124
Table A1: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of d	190
Table A2: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of n	191
Table A3: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of d	193
Table A4: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of n	194
Table A5: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of ED for HA carrier	196
Table A6: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of FPF for HA carrier.....	199
Table A7: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of ED for all carriers	202
Table A8: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of FPF for all carriers	205

Nomenclature

ACOSSO	Adaptive Component Selection Shrinkage Operator
ANN	Artificial neural network
BP	Back propagation
CCD	Central composite design
CFD	Computational fluid dynamics
CSTR	Continuous stirred-tank reactor
DoE	Design of experiments
DPI	Dry powder inhalation
EA	Evolutionary algorithms
ED	Emitted dose
FAD	Direction of azimuthal facets
FEA	Finite element analysis
FPF	Fine particle fraction
FPO	Mean polar facet orientation
GA	Genetic algorithms
GPR	Gaussian process regression
HA	Hydroxyapatite
HMs	Heavy metals
I/O	Input-Output
IBA	Incineration bottom ash
LA	Lactose
LHS	Latin hypercube sampling
MARS	Multivariate Adaptive Regression Splines
MAX	Maximum absolute error
MBLHS	Minimum bias Latin hypercube sampling

MFOV	Variation of the polar facet orientation
MN	Mannitol
MRV	Mean resultant vector
NNs	Neural networks
PS	Pattern search
PSD	Particle size distribution
PSO	Particle swarm optimization
Ra	Arithmetical mean deviation
RBF	Radial Basis Functions
Rc	Average height of an unleveled surface
RF	Random forests
Rku	Kurtosis of the assessed profile
RMSE	Root-mean-square-error
RMSECV	Root mean square error cross validation
Rp	Highest peak
RPM	Revolutions Per Minute
Rq	Root mean square deviation
RR	Rosin-Rammler
Rsk	Skewness of the assessed profile
RSM	Response surface methodology
Rt	Total height of the profile
Rv	Lowest valley
SA	Surface area
SDR	State dependent parameter regression
SEED	Sequential exploratory experiment design
SEM	Scanning electron microscopy

SR	Symbolic regression
SVM	Support vector machine
U_g	Gas velocity

Chapter 1 Introduction

1.1 Background

The design and optimization of many industrial processes involve the use of computer simulation models. However, certain systems are complex and thus making the simulation computationally expensive and time-consuming. Despite continual advances in computer speed and capacity, some simulations such as computational fluid dynamics (CFD) simulation can still be difficult or even impossible for some processes [1]. In the recent years, approximation methods, such as surrogate modeling, have attracted intense attention due to ease of use [2]. These methods approximate the complicated physical models with simple analytical models [3]. These simple models are called surrogate models, meta-models, response surface models, emulator models, or auxiliary models, etc. A surrogate model is constructed based on input-output (I/O) data, and these data are provided by experimental or simulated models, so the developed surrogate model is a simplified model of an actual model. Hence, the surrogate model is called a model of a model. The process of developing a surrogate model is called surrogate modeling [4].

The developed surrogate model can be used in the explanation of system behavior, optimization, sensitivity analysis (SA), etc. [5]. As it was mentioned, the surrogate models are built by real-world (experimental data) or simulation model input/output (I/O) data. Surrogate models try to find the general trend of the scattered data. Therefore, the accuracy of developed surrogate models in the prediction of system behavior depends on scattered data and the precision of the surrogate modeling process [2].

Figure 1.1 shows the process of data to knowledge conversion. If the appropriate conditions are provided, surrogate modeling usually starts with the design of experiments (DoE) [6, 7]. Different methods for DoE or sample design space have been listed in Table 1.1. The classical methods usually arrange the samples on boundaries of sample space and only a few data are put in the center of the design space [8]. Despite classical methods, space filling methods try to spread samples in all the design space [9]. It should be noted that, due to system complexity in many engineering disciplines and experiment costs, DoE is omitted in the surrogate modeling process and real data is usually used in surrogate modeling [10]. Therefore, as presented in Figure 1.1, data distribution methods are carried out after data collection [11].

These methods try to arrange gathered real data in the design space. Some of the data are used for surrogate modeling construction, which are called training data, and some of the data are kept for accuracy and validation evaluation as testing data for the developed surrogate model [12].

Figure 1. 1. Data to knowledge process by surrogate modeling.

The next step is surrogate modeling construction by training data. There are many different surrogate modeling methods; most usable surrogate models have been mentioned in Table 1.1. Each surrogate model has some parameters called hyper-parameters that should be determined by model fitting methods and training data [13]. There is no single surrogate modeling and model fitting method that is the best for all problems [14]. Of course, some sophisticated techniques, such as artificial neural network (ANN) and Gaussian process regression (GPR), can provide surrogate models with high accuracy [2]. Among of the model fitting methods, due to advances in modern computers, evolutionary algorithms (EA), such as genetic algorithms (GA), have been proven to be useful for finding the global optimum in the surrogate modeling process [15].

Table 1.1. Different methods of design of experiment (data distribution), surrogate models, and model fitting.

Design of Experiments or Data Distribution Method	Surrogate model Methods	Model Fitting Techniques
Classic methods (fractional (factorial), central composite, box-Behnken, alphabetical optimal, Plackett-Burman)	Artificial neural networks (ANN)	Weighted least squares regression
	Splines (linear, cubic, Non-Uniform Rational B-splines (NURBS))	Weighted squares regression
		Back propagation
Space filling methods (simple grids, Latin hypercube, orthogonal arrays, Hammersley sequence, uniform designs, minimax and maximin)	Multivariate adaptive regression splines (MARS)	Best linear unbiased predictor (BLUP)
	Gaussian process regression (GPR)	Particle swarm optimization (PSO)
	Interpolation model	Simulated annealing
Hybrid methods	Support vector machine (SVM)	Evolutionary algorithms (EAs)
Random or human selection	Ensemble and heterogenetic models	
Importance sampling		
Directional simulation		
Discriminative sampling		
Sequential or adaptive methods		

1.2 Motivation of this research

The reductions of energy consumption and raw material are always an important issue in chemical, petrochemical, and biomedical processes. Modeling of the system can provide a practical tool in the process of decision making. Due to the growing cost of energy, the lack of raw materials, and intense competition, among other reasons, the principal objective of industries is to improve the efficiency of existing processes.

In the field of systems approach to process engineering, the development of mathematical models plays a paramount role to achieve various goals ranging

from process understanding, offline optimal design, on-line real-time optimization to process control. A notable trend in process systems engineering is the ever-increasing model complexity, which may be defined as the amount of computation required to solve the model. In chemical engineering, complex models mainly originate from the physical scales being considered. For example, a complex plant-wide model (i.e. flowsheet simulation) is typically implemented by combining the models for individual processing units. Another example is that a simple reactor model based on ordinary differential equations becomes more complex if the spatial variation within the reactor is not negligible, and thus partial differential equations have to be applied. Process models are even more demanding in terms of computation if meso- and micro-scale phenomena are considered, such as computational fluid dynamic (CFD) models and molecular simulations. In general, complex models are capable of representing the underlying process more realistically and accurately. Therefore, the computational cost is among the major obstacles for the wide acceptance of complex models in practice.

To address the computational challenge, several techniques have been proposed in the literature. The method of “model reduction” is primarily designed to reduce the number of ordinary differential equations, which are typically the result of discretizing partial differential equations, using principal component analysis (PCA) [16, 17] and approximate inertial manifolds [18]. As indicated by Romijn et al. [19], purely reducing the number of equations does not automatically reduce computation, since the complexity in evaluating the nonlinear equations is intact. Following this argument, Romijn et al. [19] combined PCA with a grey-box approach, whereby the nonlinear part of the ordinary differential equations is

approximated by an empirical neural network (NN) model. The resulting reduced model runs sufficiently fast for real-time applications, such as model-based predictive and optimizing control.

As opposed to on-line applications, an alternative category of techniques is originally targeted at off-line process understanding and design. Early work in this category was presented in the community of applied statistics [20, 21]. The basic concept is to gather data from computer simulation or physical/chemical experiments, and then apply the surrogate modeling to study the impact of process inputs (e.g. operating conditions) on outputs (e.g. process yield). The data (input–output pairs) are used to develop a surrogate model, which can be used in place of the original complex model for process analysis and design. Compared with the grey-box model reduction technique, surrogate modeling is a black-box approach and is especially suitable to be used with third-party simulation tools, such as commercial flowsheet software and CFD tools. Recently, surrogate modeling has been introduced into process systems engineering for the optimization of radiant-convective drying [22], flowsheet simulations [23-25], multivariate spectroscopic calibration [26], and development of high-performance catalysts for CO oxidation [27]. Gomes et al. [28] also demonstrated the extension of surrogate modeling for real-time optimization.

In this regard, surrogate modeling is a powerful tool that can be used in modeling chemical and biomedical processes. Using surrogate modeling in modeling different processes would be helpful in developing this method in chemical and biomedical processes. In this study, we show applications of surrogate modeling in two different processes: 1- prediction of particle size distribution in gas-solid

fluidized bed and 2- evaluation of different factors in carrier-based dry powder inhalation.

Solid particle size plays a crucial role in performance and operation of gas-solid fluidized bed, for example in catalytic fluidized bed. However, it is not enough to consider only the average sizes of the particles, since also particle size distribution (PSD) plays a vital role in the performance and operation of fluidized beds. For example, in circulating fluidized beds, it is typical that the largest particles tend to remain near the bottom of the bed in dense suspension while the smaller particles flow more freely in the upper region. If one performs a simulation using only the average diameter for the whole bed, it can be hard to predict the proper solid distribution in the vertical direction. Since taking PSD into account in computational fluid dynamics (CFD) simulation needs to consider all particle-particle interactions that imposes a large number of equations to simulation procedure, CFD simulation will be a computationally expensive process. Hence, many studies usually consider constant PSD during fluidization in CFD simulation constant throughout the entire process. Surrogate modeling is introduced as a fast and cheap-to-compute alternative for computation-intensive problems such as CFD simulation. Therefore, the objective of gas-solid fluidized bed study is developing a surrogate model to estimate PSD during fluidization. Finally, with adding the developed surrogate model to CFD simulation, more accurate and reliable results can be provided. In addition, the time-behavior of PSD change under various process conditions such as different gas velocities or different glass beads size as foreign particles can be tested by developed surrogate model.

Similar to gas-solid fluidized bed, dry powder inhalation (DPI) is a process between gas and solid phases (carrier with drug is solid phase in DPI). In simulation of DPI by physical equations, there is a similar issue with simulation of PSD in fluidized beds. Due to a large number of particle-particle interactions (there is an equation for each interaction), study the fluidization process of a powder bed is computationally expensive. On the other hand, finding the effect of variables interaction on the efficiency of DPI by experiments is not possible because usually change in one variable, will change other variables inevitably. For example, change in carrier particle size will change the carrier surface roughness. So, same as gas-solid fluidized bed study, ANN-GA approach as a surrogate model was developed to evaluate the effect of different variables on DPI efficiency. With this developed model, one variable can be isolated and its effect on DPI efficiency can be evaluated. In fact, it provides a tool for better understanding of DPI formulation and it can be used for the design and optimization of DPI.

Therefore, the major contribution of this study is to apply surrogate modeling as a cheap, fast, and accurate method, to modeling of these two complex processes. These developed models can provide a powerful tool for design, optimization and sensitivity analysis of processes.

1.3 Objectives and scope

The overall objective of this study is to apply surrogate modeling for prediction of particle size distribution (PSD) in a gas-solid fluidized bed as a chemical process and evaluation of different factors in carrier-based dry powder inhalation (DPI) system as a biomedical process.

Hence, an ANN with a GA as a surrogate modeling tool is employed to model the change in PSD during fluidization. The fluidization study is divided into two parts. In the first work, experiments are conducted using incineration bottom ash (IBA) as the fluidizing particles, and different mass percentages of large and small glass beads are used as the grinding medium. The Rosin–Rammmler (RR) distribution is used to describe the IBA PSD. The developed ANN-GA models are subsequently used to study the effect of fluidization time, the mass percentage of glass beads, and the size of glass beads used on the IBA particle attrition during fluidization. For the second study in fluidization, to generalize the developed model, the attrition property of material is introduced by the planetary ball milling process. Then, time, gas velocity, initial particle size parameters, and the attrition property are used in modeling using ANN-GA. Data for three different materials, including activated carbon (graphite), gypsum, and silica, are used as training data.

For DPI system modeling in this study, after variable selection, ANN with GA as a surrogate modeling tool was employed to model ED and FPF. Similar to fluidization study, evaluation of DPI system efficiency is divided into two parts. In the first part, hydroxyapatite (HA) is used as a carrier, while, in the second part, HA in addition to lactose (LA) and mannitol (MN) are utilized to provide a surrogate model for DPI. The developed ANN-GA models are subsequently used to investigate sensitivity analysis to determine the most important variables in DPI formulation. Then, the effect of carrier properties, flow rate, and carrier-to-drug ratio on ED and FPF are studied.

Particularly, the main innovation and benefits of this study can be summarized as follows. (I) Surrogate modeling was adopted as a simple and effective method

to model gas-solid fluidized bed and DPI processes. (II) The effects of the initial PSD, time, and foreign particles on IBA particle attrition are studied. The results can be used to maximize recovery of heavy metals from IBA as a power plant waste. (III) A comprehensive model is introduced that can provide PSD information during fluidization for each material. This model requires only ball milling results of materials, gas velocity, and time to determine PSD in fluidization. (IV) Modeling of DPI by surrogate modeling can help to improve the design of drug formulation. (V) Based on sensitivity analysis results of input variables, the most effective variables on DPI efficiency can be determined. (VI) Then, a general formula is presented to use for rough computation of ED and FPF.

Finally, these case studies show that surrogate modeling as a simple and powerful tool can employ for different chemical and biomedical processes. Then, the developed surrogate model can apply for optimization, sensitivity analysis, and prototyping of process.

1.4 Organization of the thesis

This thesis comprises nine chapters. **Chapter 1** is the introduction, which gives a brief background of the research, the objective and significance of the work, and the organization of the thesis.

Chapter 2 covers the literature survey about surrogate modeling methods and techniques. Application of surrogate modeling in different disciplines and comparison of different techniques is also included in this chapter.

Chapter 3 reviews the modeling techniques that are used in the next chapters.

Chapter 4 reports the results of modeling the change in IBA PSD in a gas-solid fluidized bed due to particle attrition using a hybrid ANN-GA approach.

Chapter 5 presents the modeling of PSD in a gas-solid fluidized bed using planetary ball milling results using a hybrid ANN-GA approach.

The effect of HA size and morphology in DPI for carrier-based pulmonary delivery formulations is evaluated in **Chapter 6** through surrogate modeling.

Based on the study reported in Chapter 6, **Chapter 7** is focused on further carriers to find a comprehensive model for prediction of DPI efficiency.

Chapter 8 covers the conclusions and recommendations.

Chapter 9 covers all references.

Chapter 2 Literature Survey

2.1 Review of surrogate modeling

Computation-intensive design problems are becoming increasingly common in manufacturing industries. The computation burden is often caused by expensive analysis and simulation processes to reach a comparable level of accuracy as physical testing data. To address such a challenge, surrogate modeling techniques are often used. Surrogate modeling techniques have been developed from many different disciplines including statistics, mathematics, computer science, and various engineering disciplines [29-31]. Figure 2.1 illustrates a typical structure for construction of a surrogate model. Surrogate modeling involves (a) choosing an experimental design for generating data, (b) choosing a surrogate model to represent the data, (c) fitting the surrogate model to the observed data, and then (d) accuracy evaluation [30]. Many studies have been done on data distribution methods, surrogate modeling techniques, model fitting techniques, surrogate model accuracy, and validation, and surrogate models applications such as optimization, sensitivity analysis, prototyping, and prediction [29, 32, 33].

Figure 2. 2. A typical structure for construction of a surrogate model.

Today, surrogate modeling is known as a powerful tool in decision-making for design engineers [34, 35]. There are comprehensive reviews of surrogate modeling applications in mechanical and aerospace systems [36], structural optimization [37], and multidisciplinary design optimization [38]. According to the literature [39], some of the areas in which surrogate modeling can play a role in engineering sciences are:

- Model prediction or approximation: Surrogate modeling can provide an approximate model to use for system behavior prediction with low computation costs. For example, surrogate modeling has been used to predict clock tree synthesis as a key aspect of on-chip interconnect [40], friction factor of alluvial channel [41], and aircraft noise [42].
- Design space exploration: Surrogate modeling can help engineers in the understanding of the design problem by working on a cheap-to-run surrogate model. For instance, in the face of the actual demand for sustainable design, the use of simulation has attained high relevance in determining the energy performance of building designs. Simulation is required for examining the dynamic thermal effects of energy efficiency. However, a major problem of applying dynamic building simulation in the design process is the long computation time and the resulting delayed response. Due to surrogate modeling ability to provide quick responses compared to other methods, it is proposed for design space exploration [43]. Another example is processor architecture design space exploration by surrogate modeling [44]. Most of today's design tools such as computer aided design aim at improving the productivity of a design engineer. The relationship between design variables and product performance is usually embedded in complex equations or models in finite element or CFD codes. Engineers, by experience, often only have a vague idea about such relationship. The metamodeling approach can assist the engineer to gain insight to the design problem, currently, through two channels. The first is through the surrogate model itself. Given the surrogate model, one can analyze the properties of the surrogate

model to gain a better understanding of the problem. A good example is the quadratic polynomial surrogate model, if all the design variables are normalized to $[-1, 1]$, then the magnitude of the coefficients in the surrogate model indicates the sensitivity or importance of the corresponding term [45]. This is in fact used for screening of design variables. The second way of enhancing the understanding is through visualization. Visualization of multi-dimensional data alone has been an interesting topic, and many methods have been developed over the years [46, 47]. Winer and Bloebaum developed a visual design steering method based on the concept of Graph Morphing [48, 49]. Eddy and Kemper proposed cloud visualization for the same purpose [50]. Also, Ford integrated parallel computation and surrogate modeling for rapid visualization of design alternatives [51].

- Problem formulation: Surrogate model with associated sensitivity analysis can contribute to reducing the number of variables, variables size range, removing unnecessary constraints. On the other hand, the optimization problem becomes easier with a new problem formulation. Building a design optimization model is the first and yet critical step for design optimization. The quality of the optimization model directly affects the feasibility, cost, and effectiveness of optimization. The optimization problem, however, is usually formulated only from experience in making following decisions: 1) the objective function and, in certain cases, goals, 2) the constraint functions and limits, 3) the design variables, and 4) the search range of each design variable. Surrogate modeling and design space exploration can help the engineer to decide on

a reasonable goal for objectives and limits on constraints. Some of the objectives or constraints can be eliminated, combined, or modified. More importantly, surrogate modeling helps significantly in reducing the number of design variables and their range of search. In design engineering optimization, engineers tend to give very conservative lower and upper bounds for design variables at the stage of problem formulation. This is often due to the lack of sufficient knowledge of function behavior and interactions between objective and constraint functions at this early stage which this issue can be solved by surrogate modeling. Multivariate spectroscopic calibration [26], development of high-performance catalysts for CO oxidation [27], and carrier-based drug delivery formulations [52] are three examples of surrogate modeling application in the problem formulation.

- Optimization application: There are many optimization problems, such as global, multiobjective, multidisciplinary, and probabilistic optimization in engineering disciplines. Surrogate modeling can solve various kinds of optimization problems according to their challenges and constraints. In general, classical gradient-based optimization methods have several limitations that hinder the direct application of these methods in modern design. First, gradient-based optimization methods require explicitly formulated and cheap-to-compute models, while engineering design involves implicit and computation-intensive models such as finite elements, CFD, and other simulation models with unreliable and expensive gradient information. Second, gradient-based methods often output a single optimal solution, while engineers prefer multiple design

alternatives. Third, the gradient-based optimization process is sequential, non-transparent, and provides nearly no insight to engineers. Lastly, to apply the optimization methods, high-level expertise on optimization is also required for engineers. The advantages of applying surrogate modeling in optimization are manifold: 1) the efficiency of optimization is greatly improved with surrogate models; 2) because the approximation is based on sample points, which could be obtained independently, parallel computation is supported (assuming an optimization requires 50 expensive function evaluations and each takes 2 hours, these 50 evaluations can be computed in parallel and thus the total amount of time is 2 hours as compared to 100 hours.); 3) the approximation process can help study the sensitivity of design variables, and thus give engineers insights to the problem; and 4) this method can handle both continuous and discrete variables. Multi-objective optimization of an industrial crude distillation unit [53], optimization of crude oil distillation unit for optimal crude oil blending and operating conditions [54, 55], and optimization of steady state flowsheet simulations [56] are the tangible example of surrogate modeling application in the chemical processes.

Because of surrogate modeling application, advances in surrogate modeling have been achieved in four primary fields which make the surrogate modeling structure, namely, data distribution methods, surrogate modeling techniques, surrogate model fitting methods, and surrogate modeling accuracy and validation [30].

2.2 Data distribution methods

There are two different methods for data gathering. Data are provided by experiments or computer simulations of the process by commercial software or mathematical correlations. In general, a sample size increase can improve the surrogate model accuracy, but it imposes extra costs. There is an appropriate sample size, which should be determined based on the number of involved variables and surrogate model complexity. With four or more variables, understanding the impact of altering variable values becomes complex, particularly if the effects of interactions between variables are considered. Interaction effects are often more significant to output characteristics than single variable effects. Design of Experiments (DoE) enables this complex situation to be understood, thus gaining an in-depth knowledge of the process. This in turn can direct the engineering team to select the right control variables and allowable ranges for the setting and adjustment of those variables.

DoE deals with identifying variable model input parameters and setting the parameter values at which an experiment or simulation model is run. The set of experiment or simulation runs specified by the DoE will be used to fit a surrogate model. Numerically, the result is an experiment run matrix X , with k columns (one for each variable model parameter), and n rows (each specifying the parameter settings for an experimental run).

$$X = \begin{matrix} & \begin{matrix} \text{var } I & \cdots & \text{var } j & \cdots & \text{var } k \end{matrix} \\ \begin{pmatrix} x_{11} & \cdots & x_{1j} & \cdots & x_{1k} \\ \vdots & \ddots & \vdots & & \vdots \\ x_{i1} & \cdots & x_{ij} & \cdots & x_{ik} \\ \vdots & & \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nj} & \cdots & x_{nk} \end{pmatrix} \end{matrix} \quad (2.1)$$

There exist many types of experimental designs, which are used under different circumstances. They can be classified into two main groups: classical designs and designs for computer experiments.

Classical data distribution methods such as factorial or fractional factorial [57], central composite design (CCD) [57, 58], Box–Behnken [57], alphabetical optimal [59, 60], and Plackett–Burman designs [57] usually focus on the planning of physical experiments, so random error in physical experiments has minimum influence on the model accuracy [61].

These methods tend to spread the data points around boundaries of the design space and just put a few points at the centre of design space. In contrast, in computer simulations, systematic error is more than the random error [4, 61, 62]. In the presence of a systematic error, a space filling method such as maximum entropy design [63], mean-squared-error designs, minimax and maximin designs [64], Latin hypercube designs [65-69], orthogonal arrays [70-72], and Hammersley sequences can provide more accurate results than classical methods [73-76]. Simpson et al. [77] confirmed that space filling methods distribute data points in a reliable manner. Orthogonal arrays, various Latin hypercube designs, Hammersley sequences, and uniform designs are the most widely used methods in space filling [78-86]. Hammersley sampling is found to provide the best uniformity of data points space filling [87, 88]. Sequential and adaptive sampling methods such as sequential exploratory experiment design (SEED) [89, 90], Bayesian method [91], and inheritable Latin hypercube design [92] have also gained popularity in recent years.

Since the total number of design points in the DoE is given by the product of the number of levels of each factor, the DoE is expensive for a large number of design variables. In an experiment, each level is a particular setting of a variable input parameter or factor. For instance, the temperature factor in a chemical process might have levels of 20°C, 25°C, 30°C, 35°C, 40°C, 45°C, and 50°C, varying between experimental runs. Therefore, for example, they require q^k runs for full factorial design, where q is the number of levels and k the number of factors. As it can be understood, the number of points required becomes prohibitively large as the number of design variables increases. The number of runs can be specified in some DoE methods, but decreasing of experimental runs for a process with a large number of design variables will reduce the DoE accuracy. Hence, implementation of DoE for such experimental study is not practical and DoE usually uses when data for surrogate modeling are provided by computer simulation. Therefore, DoE can be ignored in surrogate modeling processes for experimental study with a large number of design variables [93, 94].

2.3 Surrogate modeling techniques

As stated earlier, a surrogate model is a general purpose mathematical approximation to input-output functions. Let X be a matrix of n experiment runs, with each row vector $\vec{x}_i, i = 1 \dots n$, specifying a design location based on k input variables. Further, let Y be a matrix of output responses, with each row vector $\vec{y}_i, i = 1 \dots n$, containing the performance measures of p output responses. Different types of surrogate models can be used as surrogates for the complex systems.

There are several surrogate modeling techniques that have been summarized in Table 1.1. A review about different types of surrogate modeling technique was

provided by Kajero et al. [95] in 2016. The simplest technique of surrogate modeling involves rational and polynomial functions, which are widely used in different engineering problems. Besides these functions, a stochastic model called Kriging was proposed to find the most accurate model based on random functions [96, 97]. In addition, neural networks (NNs) have been applied in surrogate modeling in various engineering problems for system approximation [98]. Other surrogate modeling techniques include Radial Basis Functions (RBF) [99, 100], Multivariate Adaptive Regression Splines (MARS) [101], interpolation model [102], and inductive learning [103]. A combination of these models is also used in some studies [104]. Other techniques are an extension or combination of the mentioned techniques. Mullur and Messac [105] introduced a new RBF model by adding a new term to the regular RBF. Turner and Crawford [106] developed a new spline surrogate model by adding a new parameter that can be used for low-dimensional problems.

There are no general comments about which surrogate model technique is superior to the others. The choice of surrogate model type and its functional form is not a simple one, and there are many criteria that need to be considered [4, 107]. Some criteria for choosing the type of surrogate model are listed below:

1. The ability to gain insight from the form of the surrogate model. Can the surrogate model be used to determine which variables are important in the model? For instance, the coefficients of a regression model provide information about the variables in the model; on the other hand, the coefficients of a radial basis function or kriging surrogate model are not interpretable.

2. The ability to capture the shape of arbitrary smooth functions based on observed values, which may be perturbed by stochastic components with general distribution. How well does the surrogate model capture the shape of the true (unknown) response? An approximation based on a low degree polynomial model will not be able to capture the shape of a highly non-linear response as well as a nonparametric model.
3. The ability to characterize the accuracy of fit through confidence intervals. How certain are we that the surrogate model predictions are correct?
4. The robustness of the prediction away from observed $(X; Y)$ pairs. Is the surrogate model sensitive to the points sampled with the experimental design?
5. The ease of computation of the approximating function. As an example, consider fitting a second-order polynomial surrogate model with least squares versus fitting a kriging surrogate model, which requires solving an optimization problem for estimating the model parameters.
6. The numerical stability of the computations, and consequent robustness of predictions to small changes in the parameters defining the approximating function. For instance, it has been pointed out that the condition number deteriorates with increasing problem dimension as well as increasing number of data values to be fit when solving the linear system for computing the coefficients for the radial basis function surrogate model [99]. The conditioning problem has also been observed with kriging surrogate models [108].
7. Does software exist for computing the surrogate model, characterizing its fit, and using it for prediction?

8. For a given a problem setting, are there empirical studies that advocate the use of one particular strategy over another?

9. How well does the surrogate model perform when it is used for optimization? For example, are the convergence properties of the surrogate model the same as for the disciplinary model?

10. The range of application scenarios. That is, can a particular surrogate model type be used for different problems varying in type, size, etc.?

In the recent works, there has been an interest in comparing different techniques on the same problem [109-116]. Some of them introduced Kriging as a successful technique in engineering problems. There are many modeling codes for Kriging in MATLAB, which are downloadable from open sources [117]. According to studies results, the Kriging models are more accurate for nonlinear problems and it is a flexible method for problems with noisy data. However, the finding of optimum hyper-parameters from likelihood estimators becomes involved with nonlinearity increase and finding the optimum one will be difficult. In contrast, polynomial techniques are simple, easy and cheap to use and clear on variable sensitivity, but their accuracy is less than that of the Kriging technique [4]. On the other hand, the polynomial model cannot be used to interpolate the sample points and this ability is limited by the chosen function type. For example, Palmer and Realff [56] have tested two case studies, a continuous stirred-tank reactor (CSTR) and an ammonia synthesis plant in which both problems included seven input variables. Minimum bias Latin hypercube sampling (MBLHS) with Kriging and polynomial have been used to build surrogate models. The developed model for CSTR by Kriging was the most accurate model. Fourteen

different engineering problems with different degrees of nonlinearity, different dimensions and noisy/smooth behaviors have been applied to test the Polynomial, Kriging, MARS, and RBF models. The number of inputs was between 2 and 16, which were organized with LHS. In general, RBF was the best surrogate model for a low order of nonlinearity while Kriging results were more accurate in the large-scale problem [4]. Another technique that is widely employed in surrogate modeling is the support vector machine (SVM) [118]. A study shows that the SVM model provides a higher accuracy than other models, including Kriging, polynomial, MARS, and RBF for a test problem. The reason for SVM's better performance over other models is not clear [102, 119]. In addition, artificial neural network (ANN), due to its precise performance, has been used in different engineering problems [120-132]. Li et al. [2] used 16 stochastic simulation problems with 2 to 8 inputs that were designed by Latin hypercube sampling (LHS). Five different surrogate models (ANN, RBF, SVM, Kriging, and MARS) were compared together, and their results show that ANN achieves the best accuracy and robustness. Villa-Vialaneix et al. [133] utilized a set of 19000 data (80% as training data and 20% for testing data). Two parametric linear techniques and six nonparametric approaches (Adaptive Component Selection Shrinkage Operator (ACOSSO), state dependent parameter regression (SDR), Kriging, ANN, SVM, and random forests (RF)) were compared together. The ANN showed the most accurate performance in this large-scale problem. Jin et al. and Zhao and Xue have also performed other comparative studies for different surrogate modeling techniques and confirmed ANN accurate performance [3, 131].

ANNs are mathematical models that attempt to imitate the behavior of biological brains. They have universal function approximation characteristics and also the ability to adapt to changes through training. Instead of using a pre-selected functional form, ANNs are parametric models that are able to learn underlying relationships between inputs and outputs from a collection of training examples. ANNs have very good generalization capability when processing unseen data and are robust to noise and missing data. Moreover, ANN can theoretically approximate any function to any level of accuracy, which is very interesting when the governing physical mechanisms are highly non-linear [134]. Several other advantages of using ANN for surrogate modeling when compared to classical regression-based techniques have been reported [135, 136]. All these advantages make ANNs very suitable to be used as the surrogates for computationally expensive simulation models. The ANN training process is in principle an optimization problem by itself because the goal is to find the optimal topology and parameters (e.g., weights and bias) to minimize the mean squared error (MSE), which is common to many ANN training algorithms. In summary, the advantages of ANNs are listed in below.

- A neural network can perform tasks that a linear program cannot.
- When an element of the neural network fails, it can continue without any problem by their parallel nature.
- A neural network learns and does not need to be reprogrammed.
- It can be implemented in any application.
- It can be implemented without any problem.

According to the literature results, ANN is adopted in this study.

2.4 Surrogate model fitting methods

Each model type has a set of parameters that control the complexity of the model. For example a polynomial model has a degree parameter, an SVM has a kernel function, Kriging has theta parameters, etc. We refer to these parameters as hyper-parameters or model parameters. To generate a good model you need to search for a good set of model parameters. In essence this is an optimization problem in model parameter space or hyper-parameter space [137, 138]. The fitting model methods are optimization methods that try to minimize the defined error for the system [139]. The error is usually determined based on differences between real data (the experimental or simulated data) and predicted data (the surrogate model responses). Different optimization methods, such as genetic algorithms (GA), pattern search (PS), particle swarm optimization (PSO), and simulated annealing have frequently been utilized in the optimization of hyper-parameters.

About ANN, sufficient volume of input/output data is required to train the neural network. The procedure to find the set of weights which minimize the errors between the predicted and the target outputs of the network is called the training of the network. Training a neural network is an iterative process. Back propagation (BP) algorithm is one of the most effective methods of ANN training. Any continuous function in a closed interval can be approximated by using a BP ANN with one hidden layer. For any complicated system, if its samples of input and output are enough, a BP ANN model that reflects the relationships between the input and output variants can be constructed after repeated learning and training. However, previous studies have shown that BP may not be an ideal option for training ANNs [140-142]. Since the initial

interconnecting weights of BP ANN are often stochastically given, the learning times and final interconnecting weights of the network are therefore changed for different times of training. That is to say, the trained network is not unique and sometimes the network possibly plunges into local optima. Gupta and Sexton [141] found that BP tends to converge to local optima. In addition, the blindness of the determination of initial interconnecting weights always results in too many training times and slow convergence [15, 143]. These shortages of BP ANN seriously impact its precision of modeling and effects of application.

Genetic Algorithm (GA) is an iterative algorithm that is parallel and global. According to the theory of GA, the possible solution in the field of problem is considered to be an individual or a chromosome of the colony, and all the individuals are then coded to be symbol strings. By simulating the evolutionary processes of organisms such as natural selection and elimination, the colony is repeatedly selected, intercrossed and mutated. Based on the evolutionary rules of survival of the fittest, and elimination of the unfittest, as well as the adaptive estimation of every individual, better and better colony is gradually evolved. At the same time, the best adaptive individuals in the optimized colony are also searched by global and parallel ways. Because the processed objects of GA are gene individuals that have been coded with parameter strings, GA can directly operate the structures of these objects. Especially, since GA evaluates multi-solutions in the searching space simultaneously, it has very strong ability of global searching and also easy to be parallelized. GA has been proven to be useful for finding the global optimum in NN training [141, 143, 144]. Thus, GA was adopted in this study.

2.5 Surrogate model validation and accuracy

The accuracy and validation of a surrogate model should be examined before being used as a surrogate model [137]. The surrogate model validation process, similar to other computational model validation, is a challenging task [138]. The primary validation method is cross-validation [139]. The training data set, S , consists of N data points $(\mathbf{x}, y)$ where y is the response data and $\mathbf{x}$ is the input data points. In P -fold cross-validation, the training data splits to P subsets and the surrogate model is fit P times omitting one subset each time, then the omitted subset is used for error computation. The P results from the folds can then be averaged to produce a single estimation.

Another validation method is the leave- k -out approach [145]. In this method, all possible $\binom{N}{k}$ subsets of size k are left out, and the surrogate model is fitted to each remaining set. Each time, the error measure of interest is computed at the omitted points. This approach is a computationally more expensive version of P -fold cross-validation.

Previous studies show that only cross-validation is insufficient for surrogate models evaluation; employing additional points as testing points are essential in surrogate model validation [89]. When testing points are used for validation, there are several different error measures for model accuracy measurement. The first two are the root-mean-square error (RMSE) and the maximum absolute error (MAX):

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (2.2)$$

$$MAX = |y_i - \hat{y}_i|, i = 1, 2, \dots, m \quad (2.3)$$

where y_i is the experimental output of test point i , $\hat{y}_i$ is the surrogate model predicted value of the test point i , and m is the number of test points. The lower the value of RMSE and/or MAX, the more accurate the surrogate model. RMSE is used to gauge the overall accuracy of the model, while MAX is used to gauge the local accuracy of the model. An additional measure that is also used is the R^2 (R-square) value:

$$R^2 = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y}_i)^2} \quad (2.4)$$

where $\bar{y}_i$ denotes the mean of experimental outputs of the test points.

2.6 Review of surrogate modeling applications in chemical engineering

As stated in section 2.1, surrogate modeling has different applications in the engineering sciences. This section reviews the applications of surrogate modeling in chemical engineering. Applications of surrogate modeling in chemical engineering include:

- **Process design and optimization:** The most straightforward application of surrogate modeling is process design and optimization. Surrogate model optimization has already been extensively used in design and optimization of many different processes. A wide variety of applications include flow- sheeting [146-150], boiler and combustion processes [151-153], separation processes such as simulated moving bed chromatography [154], pressure swing adsorption [155], heated integrated column [150], divided wall column [156], CO₂ capture process [157], reactor operation such as iron oxide reduction [158], nano

particle synthesis [159], bacteria cultivation [160], polymer processing [161], chemical processes in semiconductor industry [162-164], etc. Some of these works used actual experiments and rest of them utilized simulations to provide required data for surrogate modeling.

- **Process control:** There are numerous studies surrogate models such as ANN [165], RBF [166], SVM [167], GPR [168] can be used to represent nonlinear time series. Such models have been used in soft-sensors development to predictive important quality variables online [169-174]. They can, of course, be used in nonlinear model predictive control (NMPC) [175-178]. Tsen et al. [178] proposed a hybrid approach in which first principle simulation data were trained together with experimental data to obtain an ANN model for use in control. Such hybrid models [179, 180] were developed because of the need of using prior first principle knowledge to avoid unreasonable extrapolations and the necessity to accommodate with experimental information, i.e., migration to a more accurate and realistic model for control purposes.
- **Model calibration:** The surrogate model can also be used to improve predictions of computer simulations. Typically, a simulator requires a set of physically meaningful parameters to make predictions. For example, in CFD simulations of reactors, may consist of transport properties such as viscosity, thermoconductivity, diffusion coefficient, surface tension, thermodynamic properties such as heat capacities, model parameters for vapor-liquid equilibrium calculations, as well as kinetic parameters such as rate constants and activation energies. Theoretically, these parameters can be measured by independent experiments. In practice, they have to

be calibrated by fitting simulation results with experimental data. To do so the simulations have to be carried out at different parameter settings for each of the experiment conditions. This is of course computationally laborious and often impossible when the number of parameters to be determined is large. Alternatively, a surrogate model can be constructed that includes the parameter as input and characteristic experimental observations as output. For example, GPR has been used for multivariate spectroscopic calibration [26].

- Sensitivity analysis: A surrogate model can also help us to evaluate the sensitivity of the response to a certain input. Sensitivity analysis can also be applied uncertain parameters of a model. Sensitivity can be characterized locally by carrying out one-at-time changes to each input and examine the effect on output. Chang et al. provided an example of such approach [181]. The biochemical network was analyzed and simplified. Alternatively global variance based index such as the Sobol indices [182], fast amplitude sensitivity test (FAST) [183, 184], high dimensional model representation (HDMR) [185], polynomial chaos expansion (PCE) [186, 187], etc. can be calculated. Calculation of these global sensitivity indices is of course time- consuming using the computer simulation. However, these indices can be relatively easy using surrogate models [188-191]. Applications of sensitivity analysis to chemical engineering related problems include reaction kinetics [192, 193], biological system modeling [194], process design [195], enhanced oil recovery simulation [196], vapor cloud dispersion [197], etc.

As it was mentioned, the surrogate model can be used for different applications in chemical engineering. There are still many complicated issues in chemical engineering which can be analyzed by surrogate modeling. Hence, in this study, the surrogate modeling has been applied for prediction of PSD during fluidization which can help chemical engineer in gas-solid fluidized bed design, operation and CFD model calibration as well as modeling of drug delivery efficiency in DPI which can be applied for development of new inhalation formulations and new carriers for carrier-based DPI.

Chapter 3 Modeling Techniques

3.1 Preface

The background and literature review related to this chapter is presented in chapter 2. As it was mentioned, ANN was selected among surrogate modeling techniques. Then GA as a powerful optimization tool is applied as model fitting method. In addition to surrogate modeling, different methods such as variable selection, sensitivity analysis, symbolic regression, and particle size distribution are used in various case studies which are introduced in this chapter. Moreover, the proposed combined method for surrogate modeling that is ANN-GA is described briefly.

3.2 Artificial neural network (ANN) as a surrogate modeling technique

Surrogate modeling is an approximation method developed for prediction, calibration, and optimization of the process behavior. Selection of a suitable model usually requires the use of empirical evidence in the data, knowledge of the process and some trial-and-error experimentation. It should be noted that model building is always an iterative process [198].

ANN is an excellent surrogate model for systems that are difficult to express by physical equations. An ANN structure contains interconnected neurons that link the input, output and hidden layers [199]. A typical mathematical form of ANN with three layers and one single neuron output is [2]:

$$\hat{y} = \hat{f}(X) = \sum_{j=1}^J w_j f\left(\sum_{i=1}^I v_{ij} f(x_i) + \alpha_j\right) + \beta + \varepsilon \quad (3.1)$$

where X is a k -dimensional vector with $x_1, x_2, \dots, x_k$ as its elements, f is the user defined transfer function, ε is a random error with a mean of 0, v_{ij} is the weight on the connection between the i th input neuron and the j th hidden neuron, α_j is the bias in the j th hidden neuron, w_j is the weight on connection between the j th hidden neuron and the output neuron, I is the total number of input neurons, J is the total number of hidden neurons, and β is the bias of the output neuron. Figure 3.1 depicts this neural network (three layers and one single neuron output) with working of a single neuron explained separately. The weights and biases (hyper-parameters) can be determined by a training procedure that minimizes the training error [2]. The most important parameters of ANN are the number of hidden layers, the number of hidden neurons in each hidden layers and the transfer function which directly affect the ANN's prediction.

Figure 3. 1. Artificial neural network structure.

As stated, back-propagation (BP) algorithm is known as one of the most classical training algorithms [200]. However, some studies have shown that evolutionary strategies (ESs) can outperform back-propagation techniques in the determination of weights in neural networks [201, 202]. Genetic algorithms (GA) is one of the ESs which has been taken from Charles Darwin's evolution theory [203]. GA is a search technique that is used for optimization. Initially, GA selects random population as solutions to the optimization problem, known as parents. The population is then evolved by means of selection, mutation and crossover operators to form children solutions. Due to these operators, different sets of population can be generated in an iterative process to form generations. Finally, the generation that gives the best solution to the problem is selected. In the determination of weights in ANN, GA is used to find the best weights that minimize the error involved in the network constructed by a special training rule.

3.3 Variables selection

It is a challenge to select the variables needed to build an adequate surrogate model, especially when there is a large number of variables involved and the relative importance of the various input variables on the output variables are unknown. On the other hand, it is often not feasible to utilize all the available variables because it can cause issues such as usage of irrelevant variables, overfitting, increased complexity and difficulty in model convergence.

Variable selection is the process of finding the most relevant variables for a predictive model. These techniques can be used to identify and remove unneeded, irrelevant and redundant features that do not contribute or decrease the accuracy of the predictive model. Variable selection is used to reduce the number of inputs

in the database before the modeling process to simplify the models created, to find the most relevant variables, and to save time and computational resources [204]. In fact, an optimization algorithm is used in variable selection to select the most optimal subset which can best describe the behavior of the model. Here the function to optimize is the generalization performance of the predictive model, represented by the error on a selection data set. The design variables are the inclusion (1) or the exclusion (0) of the variables. An exhaustive selection of features would evaluate lots of different combinations (2^N , where N is the number of variables). This process requires lots of computational work and, if the number of features is big, becomes impracticable. Therefore, we need intelligent methods that allow performing feature selection in practice. Stepwise regression is the simplest approach in variable selection. However, variable interaction is not taken into consideration and the developed model based on stepwise regression is not reliable. GA is a promising alternative that is one of the best simultaneous selection methods [205].

GA selects series of variables, which can describe the correlation between the input and the output variables. There is an X -block of input data which should predict the Y -block as the output variables. So GA can choose a random subset of variables from X and locate the variable subset (or subsets). As the genetic algorithm is a stochastic optimization method, the genes of the individuals are usually initialized at random. Once we have generated and initialized the population, we need to assign the fitness to each individual. The partial least-squares (PLS) Matlab function has been applied as fitting function in this study. The root-mean-square error of cross validation (RMSECV) obtained when using only that subset of variables in a regression model. Then evaluate its selection

error with the selection data. Obviously, a high selection error means a low fitness. Those individuals with greater fitness will have a greater probability of being selected for recombination. After fitness assignment has been performed, the selection operator chooses the individuals that will recombine for the next generation. The individuals most likely to survive are those more fitted to the environment. The whole fitness assignment, selection, recombination and mutation process is repeated until a stopping criterion is satisfied. Finally, after the training and calibration, variables which have a higher frequency of inclusion in the solutions, are more likely to make the solutions with low fitness value. Variable selection by GA was performed by “PLS toolbox” [206] in Matlab R2014a.

3.4 Sensitivity analysis (SA)

A sensitivity analysis is carried out to investigate output variability against changes in input variables. It can provide important information about deterministic variables for output prediction.

Sensitivity analysis (also known as what-if analysis) is used to determine the value of an individual output change by varying an independent input. Typical questions addressed by sensitivity analysis are: What input factors cause the largest variation in the output? Is there any factor whose variability has a negligible effect on the output? Are there interactions that amplify or dampen the variability induced by individual factors? We can distinguish different types of sensitivity analysis depending on how these questions are formulated and addressed [193].

- Local versus global: Local sensitivity analysis considers the output variability against variations of input factors around a specific value, while global sensitivity analysis considers variations within the entire space of variability of the input factors. The application of local sensitivity analysis obviously requires the user to specify a nominal value for the input factors. While global sensitivity analysis overcomes this possible limitation, it still requires specifying the input variability space. When the latter is poorly known, the conclusions drawn from global sensitivity analysis should be taken with care [194, 195].
- Quantitative versus. Qualitative: We use the term quantitative sensitivity analysis to refer to methods where each input factor is associated with a quantitative and reproducible evaluation of its relative influence, normally through a set of sensitivity indices. In a qualitative sensitivity analysis, instead, sensitivity is assessed qualitatively by visual inspection of model predictions or by specific visualization tools like, for instance, tornado plots, scatter (or dotty) plots or representations of the posterior distributions of the input factors [193].
- One-At-a-Time (OAT) versus. All-At-a-Time (AAT): Another distinction often made is between One-[factor]-At-a-Time (OAT) methods and what we propose to call All-[factors]-At-a-Time (AAT) methods. This distinction refers to the sampling strategy used to estimate the sensitivity indices. In fact, in general, sensitivity indices cannot be computed analytically due to the complexity of the input–output relationship and thus they are numerically approximated from a sample of input factors and associated output evaluations. The distinction

between OAT and AAT methods is based on the approach adopted to select input samples. In OAT methods, output variations are induced by varying one input factor at a time, while keeping all others fixed. In AAT methods, output variations are induced by varying all the input factors simultaneously, and therefore the sensitivity to each factor considers the direct influence of that factor as well as the joint influence due to interactions [193].

The main purposes of sensitivity analysis are:

- Ranking aims at generating the ranking of the input factors $x_1, x_2, \dots, x_n$ according to their relative contribution to the output variability.
- Screening aims at identifying the input factors, if any, which have a negligible influence on the output variability.
- Mapping aims at determining the region of the input variability space that produces significant, e.g. extreme, output values.

The purpose of sensitivity analysis (screening, ranking or mapping) defines the ultimate goal of the analysis. It, therefore, guides the choice of the appropriate sensitivity analysis method since different methods are better suited to address different questions.

The main constraint of local One-At-a-Time (OAT) approach is that it cannot detect the presence of interactions between input variables. Therefore, when we use a surrogate model such as ANN which can provide interactions between input variables by its parameters (weights), global sensitivity analysis and model calculation are interlaced in a more complex, often iterative way. Hence, a simple Monte Carlo simulation has been adopted to do SA for screening in this study.

Monte Carlo simulation is a mathematical technique that produces distributions of possible outcome values. One main application is to account for risks in quantitative analysis and decision making. By using probability distributions, variables can have different probabilities of different outcomes occurring. Probability distributions are a much more realistic way of describing uncertainty in variables of a sensitivity analysis. Common probability distributions include normal, lognormal, uniform, other user defined distributions such as triangular and discrete. During a Monte Carlo simulation, values are sampled at random from the input probability distributions. Each set of samples is called an iteration, and the resulting outcome from that sample is recorded. Monte Carlo simulation does this hundreds or thousands of times, and the result is a probability distribution of possible outcomes. In this way, Monte Carlo simulation provides a much more comprehensive view of what may happen. It tells us not only what could happen, but how likely it is to happen.

Monte Carlo simulation has been adopted for sensitivity analysis in this study because it provides a number of advantages over deterministic analysis: 1) results show not only what could happen, but how likely each outcome is, 2) because of the data a Monte Carlo simulation generates, it's easy to create graphs of different outcomes and their chances of occurrence, 3) usually deterministic analysis makes it difficult to see which variables impact the outcome the most. In Monte Carlo simulation, it's easy to see which inputs had the biggest effect on bottom-line results [207]. Monte Carlo simulation for SA has three steps: 1) identification of mathematical modeling, 2) creation of input data, and 3) computing and analyzing outputs. The mathematical model is developed ANN model in this study. For input data sampling, the normal distribution (also known as a Gaussian

or bell curve) is often used to approximate distributions when the true distribution is not known. An important statistical theorem, the central limit theorem, states that in the limit, a sum of independent random variables is approximately normally distributed [208]. Moreover, when the variable under consideration is just one among many variables (more than 10 variables) which affect the overall problem or system, the output is often not very sensitive to the choice of the distribution. This can be likened to the phenomenon of summing a series of random variables, none of which dominates the sum. In this case the total tends to have a normal distribution irrespective of the individual distributions [209]. Large samples give us a great deal of statistical power. In addition, since output data have been shown as normal distribution, large samples presents smooth curve [210]. Hence, in this study, 1,000 Monte Carlo samples were generated as inputs by normal distribution to evaluate the SA of output variables. Each input variable mean and 0.2 times each input variable means (this variance covers equal or more than 95% range of input data) were considered the mean and standard deviation of the normal distribution to produce random input data. Then, outputs were calculated using the surrogate model. The output mean and variance were used to determine the sensitivity of output against input changes.

3.5 Symbolic regression (SR)

Symbolic regression (SR) is an automatic search that can express regression problems with an equation. According to the number of weights and biases, the developed ANN-model mathematical equation is complex and useless; therefore, SR is an appropriate alternative to indicate the outputs based on sensitive variables using a simple equation.

The SR can find an equation for an output based on input variables without having any expert's pre-knowledge or predefined regression structures, such as linear or nonlinear regression [211]. Alternatively, the equation is created by randomly combining mathematical operators, functions, constants, and independent variables. The SR discovers an equation by simultaneous search in the huge space for finding potential solutions [212]. For discovering the potential equations, different operators are used. The operators for SR in this work include add (+) two and three variables, minus (-), multiply ($\times$) two and three variables, divide (/), exponential function (exp), and logarithmic function (log). The developed model using SR is also evaluated by error metering.

The interpretability of symbolic regression models is often hampered by overly complex and large formulas, bloating and introns, and the excessive usage of variables. Furthermore, complex models tend to overfit and memorize the training data, which results in poor prediction performance on unseen data. Hence, simpler models with similar training accuracy are generally preferred to complex ones. Symbolic regression problems are commonly solved by genetic programming, where the formulas are generated during the optimization process and internally represented as symbolic expression trees [213]. Idea how to solve various problems by SR by means of evolutionary algorithms (EAs) come from John Koza who used a genetic algorithm (GA) [214, 215]. A tree is created by genetic algorithm. A generation of such trees is first produced at random:

1. First, randomly pick a top-level operator node (e. g. "+" node). An operator (rather than a constant or variable) must be chosen here so that the tree will have at least one branch.

2. Randomly choose next-lower-level nodes from the set of operators and terminals (constants and variables) and add them as sub-nodes of the top-level node.
3. Continue until there are no more operator nodes requiring sub-nodes, i.e., every branch ends in a terminal node.
4. Build more trees in the same way, until a pre-determined generation size is created.

Next, each tree is evaluated for its fitness. That is, each tree, in turn, is used to predict the output for each input data set. This predicted value for output is compared with the output value of the data point, and the error in prediction is accumulated over all the data rows. Then the best tree is selected. An open-source MATLAB toolbox [216] was employed for SR in this study.

3.6 Genetic algorithms (GA)

The GA is a search technique that is used for optimization. The original ideas of GA, which inspired by biological evolution is efficient domain independent search method. Figure 3.2 depicts evolution flow of genetic algorithm. Initially, the GA randomly generates a population based on sets of solutions, known as the first generation. This population can comprise various sets of solutions from several to thousands. These individual sets of solutions can be possible solutions to the optimization problem. For the case of variable selection, an individual is a subset of input variables. For ANN optimization, an individual is a set of all the weights and biases of an ANN network. Regarding SR, one individual will generate one corresponding equation. Then, the fitness value of each individual is calculated, and the individuals are arranged based on their fitness value. In the

next step, the best individuals are selected as parents of the next generation by selection techniques, such as the stochastic uniform method. The next generation is then created from the selected parents by applying elite, crossover, and mutation operators. Some parents with lower fitness values are chosen as elite and are passed to the next generation. Two parents combine using a method like the scattered method to produce the new individual for the next generation in a crossover operation. A crossover fraction specifies the fraction of the next generation, other than elite individuals, that are produced by crossover. For production of the remaining individuals of the new generation, the GA makes small random changes in the individuals using a function, such as Gaussian, to create the new individuals in the mutation operation. After new generation creation, the selection of best individuals and new generation production processes is repeated until one of the stopping criteria, such as the predefined maximum generation, time limitation, or fitness limitation, is met.

Figure 3. 2. Evolution flow of genetic algorithm.

3.7 Accuracy and validation of surrogate model

Cross-validation is a statistical method of evaluating and comparing learning algorithms by dividing data into two segments: one used to learn or train a model and the other used to validate the model. In typical cross-validation, the training and validation sets must cross over in successive rounds such that each data point has a chance of being validated against. The basic form of cross-validation is P -fold cross-validation. Other forms of cross-validation are special cases of P -fold

cross-validation or involve repeated rounds of P -fold cross-validation. Leave-one-out cross-validation is a special case of cross-validation where the number of folds equals the number of training data. The learning algorithm is applied once for each data, using all other data as a training set and using the selected data as a single item test set [217].

First, the accuracy of surrogate models is checked by leave-one-out cross-validation. Then the accuracy of built surrogate models is tested using the testing points. The testing samples are selected randomly to cover all the range of input variables. For validation test, additional validation points are used. Criteria for a good model for protective purposes has been specified by Alexander et al [218]. According to these criteria, Root-mean-square error (RMSE) and R^2 (R-square) were used for the evaluation of model accuracy and validation:

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (3.2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y}_i)^2} \quad (3.3)$$

where y_i is the experiment output of test point i , $\hat{y}_i$ is the model predicted value of the test point i , m is the number of test points, and $\bar{y}_i$ denotes the mean of experiment outputs of test points.

3.8 ANN-GA as an combined approach for process modeling

The idea of the proposed combined approach was to provide a model to design various manufacturing processes, without assumptions related to the type of

process, interdependences in a process, type and number of process parameters and responses or correlations among them. Various authors employed artificial neural networks (ANNs) to design a process e.g. Hsieh [219]. As stated earlier in section 2.3, ANN is a powerful technique to generate complex process models without referring to a particular mathematical model, proven effective in various applications. On the other hand, genetic algorithms (GAs) are robust adaptive optimization techniques that allow an efficient probabilistic search in a high dimensional space. A GA is proven as a potent multiple directional heuristic search method for optimizing highly nonlinear, non-convex and complex functions, and it is less likely to get trapped at a local optimum than traditional gradient-based search methods.

The recent works have discussed the application of the method of integrating artificial neural networks (ANNs) and genetic algorithms (GAs) to optimize the parameters setting of industrial processes. ANN is used to construct the response function of a process; a GA is then applied to the ANN model for searching the optimal parameters setting that leads to the desired responses. Malakooti and Raman [220] develop an ANN-based method for machine setup problem for a turning operation. Markopoulos et al. [221] used ANN to predict a single response (surface roughness) in electrical discharge machining. Hou et al. [222] presented an integrated system for setting the optimal parameters in chip-package wire bonding processes in semiconductor industry, based on application of ANN and GA. Su and Chiang [223] also used ANN and GA to optimize the wire bonding process. Tong's approach [224], based on case-based reasoning, ANN and GA, is designed to optimize transfer molding of electronic packages. Chang and Chen [225] proposed a general neuro-genetic approach to dynamic parameter

design based on exponential desirability function, similar to the Noorossana's approach [226]. Mok et al. [227] presented an intelligent system based on case-based reasoning, ANN and GA, to optimize injection molding process. Lau et al. [228] also employed GA for optimization of molding operations. Wang et al. [229] performed multi-objective process optimization of the serial-parallel hybrid polishing machine tool using ANN and GA. Huang and Tang [230] utilized ANN and GA for parameter optimization in melt spinning. Sedighi and Afshari [231] used ANN-GA approach to optimize creep feed grinding process. Holden and Serearuno [232] introduced an artificial intelligence method that integrates rule-based knowledge representation and fuzzy logic for precious stone manufacturing improvement.

Sibaliya [233] proposed hybrid intelligent system model that presents a generic and integrative approach to the process design based on combination of ANN and GA as presented in Figure 3.3. In the suggested approach, after collection training data, ANNs were trained to represent the surrogate model of a process that is capable of predicting the corresponding performance measure by giving any parameter (factor) combinations within the feasible solution space. Then, GAs were utilized to obtain the best performance measure and the corresponding factors combination. By employing ANN to model the process behavior, the process is considered as a black-box. Hence, it does not require any prior knowledge about process interrelationships. This feature essentially contributes to the generality of the proposed approach, because the model does not depend on the type of relations between responses and process parameters or their correlations, thus making it applicable to different processes.

Figure 3. 3. The structure ANN-GA as a hybrid intelligent system model for the process modeling.

The general structure of the ANN-based process model used in this approach is shown in Figure 3.1. Neurons in the input layer are process parameters (factors), so the number of input neurons equals the number of parameters; the output layer

has one neuron corresponding to a single synthetic performance measure. Neurons in the hidden layer are computational units that perform non-linear mapping between inputs and outputs. The training and testing data for ANNs were obtained from experimental results. Input data set contains process parameter values for all experimental trials; output set accommodates the synthetic performance measure. The most successful training technique the ANN learning method, improved by a hybrid of the Gauss-Newton nonlinear regression method and gradient steepest descent method (Levenberg-Marquardt algorithm), was adopted in this approach [234].

Selection of a transfer function is very important way of information processing of an artificial neural network. Many studies tried to compare transfer functions for different problems [235-239]. There are several questions in selection of the best transfer function: 1) what type of neural network we are dealing with? 2) What is the learning paradigm? 3) Whether we want to use the same transfer function in the hidden layer and in the output layer? And (the most important question) 4) what is our plan to use ANN? (For function approximation or pattern recognition). Almost all the literature are agreed that using the tangent sigmoid as transfer function of the input and hidden neurons, while a linear transfer function is used in the output neurons is reasonable for function approximation tasks [235].

Determining the number of hidden neurons is critical in the design of the ANN-based process model. There is no perfect rule which gives us the exact number of neurons in hidden layer for the best performance of neural network. There are some empirically-derived rules-of-thumb, of these, the most commonly relied on is the optimal size of the hidden layer is usually between the size of the input and

size of the output layers [240]. In addition, number of training data is important in ANN construction. To prevent overfitting, the number of training data should be larger than the number of weights in the network [7, 241, 242]. Therefore, comparison various ANNs which are trained by back propagation with different topology (number of hidden neurons) may get a rough idea what architecture of ANN gives best performance.

The trained neural model presents an objective function for a GA, who, by maximizing the objective function finds the optimal parameters setting among all possible solutions in a continual multi-dimensional space. Herein, a possible solution represents an individual (chromosome) which can provide the fitness value of a GA. Genes in the individual (chromosome) are formed by the values of process parameters.

According to Sibalija and Majstorovic recommendations [243], the following required information for GA are provided:

- Set the population size: To provide the wider space that will be examined by a GA, in this approach the population size is equal or larger than dimensionality.
- Set the initial range: According to trained ANN, initial ranges are obtained. This feature of the suggested model is of essential importance because it allows GAs to converge to the global optimum faster and enhance its capability to find the actual global solution in the given number of generations.
- Set the bounds of process parameters.
- The fitness function: First, the selected ANN-based process model is loaded. Then, function variables are defined, as well as specific

constrains or inter relations among them if such exist. Finally, the calculation of fitness function values of individuals in the current population is performed. To speed up the calculation of the fitness, the function is defined in a vector form.

- Set the termination criterion: A pre-set number of generation is used as the stopping criteria.

A GA starts with evaluation of the fitness of individuals in the initial population, searching for the best individuals that will form a new population by using the following operations:

- Scaling: Rank scaling function is adopted in this approach.
- Selection, crossover, and mutation: According to the results of previous analysis by Sibalija and Majstorovic [234], the choice of basic GA's operations selection and crossover functions depends on the application. But, they could conclude that the stochastic uniform selection, the scattered crossover, and Gaussian mutation are adequate for many applications. The Crossover fraction migration fraction set to 0.7 – 0.9. The Elite count is 5 – 20.

Relaying on this and setting GA's parameters as described above, the proposed approach ensures optimal performance of GA to converge to a global rather than a locally optimum solution of an optimization problem [234, 243].

3.9 Particle size distribution (PSD)

Various distribution functions, such as normal distribution, log-normal distribution, Rosin-Rammler distribution, Johnson's distribution, etc. have been

developed to describe PSD. It was reported in the literature that RR distribution could provide a good description of PSD in comminution processes such as grinding, milling, and crushing [244-246]. Since particle attrition in a fluidized bed can be considered a comminution process, RR distribution is used in this study to describe the PSD. A RR probability density function (PDF) of variable X follows the form:

$$F(X, n, d) = \begin{cases} \frac{n}{d} \left(\frac{X}{d} \right)^{n-1} e^{-(X/d)^n} & X \geq 0 \\ 0 & X < 0 \end{cases} \quad (3.4)$$

where n and d are fitted parameters.

n governs the shape of the distribution while d primarily changes the location of the distribution and has a slight impact on the shape of the distribution. When n has value between 0 and 1, the PDF has a monotonic shape which decreases from ∞ as X increases from 0. For n equals to 1, the PDF decreases monotonically from $\frac{1}{d}$ as X increases from 0. When n has value greater than 1, the PDF becomes log-normally distributed. In many applications of RR PSD, the PDF follows the behavior of n value between 0 and 1. The relationship between d and n on RR PDF is shown in Figure 3.4. It can be seen that a decrease in d shifts the distribution to the left and reduces the spread of the distribution. While a reduction in n also shifts the distribution to the left, the spread of distribution increases. A reduction in both d and n is thus able to shift the distribution to the left and maintain the spread of the distribution. In this study, a Matlab toolbox of Rosin-Rammler distribution (RRD) is utilized [247].

Figure 3. 4. Effect of d and n on RR distribution.

Chapter 4 Modeling the change in particle size distribution in a gas-solid fluidized bed due to particle attrition using a hybrid artificial neural network-genetic algorithm approach

4.1 Preface

For more than 70 years, many industries in chemical, petrochemical, food, and pharmaceutical fields have been using fluidized beds due to the versatility and ease of operation. Understanding the behavior of fluidization is crucial to the development of proper control and optimization strategies. Parameters such as fluid velocity, particle and fluid properties, temperature, pressure, height and diameter of reactor, voidage and particle size distribution (PSD) are all important properties required in mathematical model of a fluidized bed [248, 249].

The average diameter of the particles is one of the most important factors affecting fluidized bed behavior. For example, the diameter plays a big role in the Geldart classification of a fluidized bed. Also the important minimum fluidization and terminal velocities of the particles depend largely on the diameter and even relatively small changes in the particle sizes affect these velocities substantially [250]. However, it is not enough to only consider the average sizes of the particles, since also PSD plays a vital role in the performance and operation of fluidized beds [251]. For example, in circulating fluidized beds it is typical that the largest particles tend to remain near the bottom of the bed in dense suspension while the smaller particles flow more freely in the upper region. If one performs a simulation using only the average diameter for the whole bed, it can be hard to predict the proper solid distribution in the vertical direction.

The shape of the size distribution also affects the pressure drop and minimum fluidization behavior of the bed. With uniform particles the pressure drop increases linearly with the velocity until the minimum fluidization velocity is reached. After that point the pressure drop decreases a bit and then stays almost constant in the bubbling region. On the other hand, with a wide size distribution the minimum fluidization velocity is not that well defined since the smaller particles start to fluidize earlier than the larger particles and the bed can be in a partially fluidized state. In these cases minimum fluidization velocity (U_{mf}) can be defined. In addition to the influence on pressure drop, size distribution can also affect the voidage of the bed. The smaller particles can fit in the holes which are in between the bigger particles and that way reduce the empty space inside the bed [252].

During fluidization, the bed media in the fluidized bed are constantly experiencing mechanical stress due to particle-particle and particle-wall collisions. As a result, the particles size reduces continuously by attrition. Two main modes of attrition in fluidized bed are abrasion and fragmentation [253, 254]. The particles size reduction can change the fluidization performance and cause undesirable behavior. For example, a reduction in PSD in a coal combustion process can lead to a loss of both unburnt fuel and sorbent, causing a decrease in the efficiency of combustion and sulfur retention [253]. In catalytic reactions, attrition was reported to cause a reduction in the catalyst efficiencies [255-258]. One study has also shown that heavy metals (HMs) in incineration bottom ash (IBA) can be concentrated in a specific size range after fluidization of IBA [259]. An understanding of the change in PSD during fluidization can be beneficial in optimizing the recovery or removal of HMs in IBA.

Many computational studies of fluidized beds make use of mathematical relations [260, 261] or computational fluid dynamics (CFD) [262] to model the fluidization process. CFD Intrinsically is a computation-intensive tool for fluidization simulation. Moreover, direct solution of Navier-Stokes equations for finding particle size in fluidization can impose more than several million computations on the simulation process [263-266]. Therefore, a majority of the CFD models developed for fluidization process considered the PSD to be constant throughout the entire process [263, 267-279]. Few researchers used some estimations to implement the particle size distribution parameter into the CFD models. For example, Mathiesen et al. [280], performed CFD simulation for a circulating fluidized bed with just two different particle size distribution or Hamzehei [263], applied a multi-fluid Eulerian model incorporating kinetic theory for solid particles to simulate the unsteady state behavior of fluidized bed reactor. The model predictions of fluidized bed behavior are often prone to discrepancies as particle attrition is typically inevitable. Therefore, a systematic approach is needed to provide information on the change of PSD during fluidization.

Over the past decades, many research efforts were put into understanding particles attrition processes in fluidized beds. Due to a large range of operating conditions involved, it is difficult to rely solely on experimental results [281]. Bulk test results were found to be not suitable to represent large-scale operations [251]. The attrition rate was found to be a function of material properties, initial particle size, time, and fluidization conditions such as temperature, pressure, gas velocity, and fluidized bed geometry [257, 282-300]. The first attrition model was presented by Gwyn [301] that has the form:

$$\frac{m_{elutriated}}{m_{bed,0}} = K_a t^b \quad (4.1)$$

where $m_{elutriated}$ and $m_{bed,0}$ are the elutriated mass and the bed mass at $t = 0$, respectively. K_a is the attrition constant that decreases with the mean particle size and b is a constant that is material dependent. Most subsequent studies aimed to determine the attrition rate constants for different particle size ranges separately [254, 289, 302-307]. However, the constants K_a and b need to be determined separately for each material. In addition, these correlated models were obtained from specific conditions and therefore have limited applicability. An alternative approach is to consider the change of attrition rate as a function of time [255, 308-315]. However, these attrition rate models are not able to translate the results into PSD information. There is also a lack of studies the presence of foreign particles on particle attrition in fluidization, which can be of industrial relevance.

In this chapter, the change in PSD due to particle attrition with and without the use of glass beads as grinding medium in a co-fluidization process is investigated. Since PSD simulation by simulation tools is a complex process, artificial neural network (ANN) is used to model the change in PSD at various operating conditions as a surrogate model. Genetic algorithm (GA) is applied to determine the ANN parameters. A Rosin-Rammler (RR) distribution function is used to describe the PSD. ANN being a tool of surrogate modeling allows to provide an approximate model to use for system behavior prediction with low computation costs when multiple parameters are involved [2, 119, 316] and is widely used for process optimization [6, 13, 27, 56, 146, 235, 317-322].

4.2 Experimental setup

Experiments were conducted in an acrylic fluidized bed. The fluidization section is 1200 mm in height and 30 mm in internal diameter. Perforated plate was used as the gas distributor and to support the solids. Compressed air was used as the fluidizing medium. The gas pressure was controlled by a pressure regulator and the gas flow rate was monitored by a rotameter. The air flow rate was set at 0.78m/s to maintain the fluidization in the bubbling fluidization regime. IBA particles having a particle size up to 1.7 mm were used as the main solid phase for the study of particle attrition. Two different sizes of glass beads 0.212 – 0.6 mm (small) and 1.7 – 2 mm (large) were used as grinding medium for co-fluidization with IBA particles. The schematic diagram of the experimental setup is shown in Figure 4.1. At various fluidization times, the experiments were stopped and the PSD of the IBA particles were measured using a sieving machine (RETSCH vibratory sieve shakers AS200) using 5 sieves of mesh sizes 0.212 mm, 0.6 mm, 1.0 mm, 1.4 mm, and 1.7 mm. While large glass beads can be separated from IBA particles relatively easily by sieving, the separation of small glass beads from IBA particles can be difficult. Therefore, the mass of IBA particles in the size range of 0.212 – 0.6 mm was determined by deducting the total mass of small glass beads used in the experiment from the total mass of particles obtained in this size range.

Figure 4. 1. Fluidized bed experimental setup.

4.3 Data collection

The datasets used in this study were collected experimentally using the setup described in the experimental section. The d and n in RR PSD were considered as the outputs. The factors affecting d and n include residence time (time), initial PSD (d_0 and n_0), mass percentage of IBA, and mass percentage of the two glass beads size ranges. Two ANN models were constructed separately to predict d and n . The input variables and their range of values are given in Table 4.1.

Table 4.1. Input and output variables their range of values

Input variables	Symbols	Range of values
Time (min)	x_1	10.00 – 1200
d_0	x_2	0.50 – 1.00
n_0	x_3	1.00 – 2.00
IBA mass fraction	x_4	0.50 – 1.00
Mass fraction of either small glass beads (0.2-0.6 mm) or large glass beads (1.7-2 mm)	x_5	0.00- 0.50
Output variables		
d	y_1	0.10 – 1.00
n	y_2	0.50 – 2.00

ANN surrogate models were built using 267 training samples and additional 14 samples were used as test sets. The testing samples were selected randomly to cover all the range of input variables.

4.4 Design of the ANN model for prediction of PSD

In this study, a five-layer (namely, one input layer, three hidden layers, and one output layer) network was used. A schematic diagram of this multilayer feed-forward neural network architecture is shown in Figure 4.2. The weights, biases and GA convergence graphs have been shown in Table A1 and Figure A1 in Appendix. The optimum number of hidden layers was found by trial and error following the rule of thumb that the number of hidden layers should be between number of input and output variables [240]. A publicly available Matlab SUMO toolbox (<http://www.sumowiki.intec.ugent.be>) [11, 323, 324] was employed. The toolbox was also used to normalize all the input and output data to scale between -1 and 1. A hybrid of the Gauss-Newton nonlinear regression method and gradient steepest descent method (Levenberg-Marquardt algorithm) were used to build the feed forward ANN. The weights and biases were generated randomly in the range of -1 and 1. The tansig (tan-sigmoid) transfer function was used between the input and hidden layers. The purelin (linear) function was used as the transfer function between the hidden and output layers. The network weights and biases were modified according to the GA to achieve the global minima of deviation error, and optimum weights and biases were obtained simultaneously. The parameters used in GA for ANN optimization are listed in Table 4.2. The weights, biases and GA convergence graph has been shown in Appendix.

Table 4.2. GA parameters for ANN optimization

GA parameters	ANN optimization
Population size	120
Maximum generation	140
Elite count	10
Crossover fraction	0.7
Selection technique	Stochastic uniform
Crossover method	Scattered
Mutation function	Gaussian
Stopping criteria	Maximum generation

Figure 4. 2. Artificial neural network structure for prediction of d (y_1) and n (y_2) as RR distribution parameters..

4.5 Results and Discussion

4.5.1 Application of the Rosin–Rammler model to the IBA particle size distribution analysis in fluidization

The goodness of fit was presented in terms of R-square and RMSE values. Figure 4.3 shows the original IBA PSD as well as the best and the worst experimental data fitted by RR PSD. The histogram represents the experimental result and the curve represents the fitted RR PSD. The original IBA PSD is presented in Figure 4.3a and the PSD is represented by five particle size ranges between 0 and 1.7 mm corresponding to the sieves used in the sieving machine. The original IBA PSD can also be considered to be the starting IBA PSD under all experimental conditions. The best-fitting case by RR PSD is for pure IBA after 30 minutes of fluidization (Figure 4.3b) and the worst one is for co-fluidization of IBA particles with 50% by mass of small glass beads after 300 minutes (Figure 4.3c). Fluidization time has insignificant impact on the goodness of fit for RR PSD. For instance, R-square and RMSE are 0.9912 and 0.0271, respectively at time = 0 and they are 0.9873 and 0.0274, respectively after 1200 minutes for the fluidization of pure IBA particles. Co-fluidization of IBA particles with glass beads reduces the goodness of fit, but the most deviation was found to be under co-fluidization of IBA particles with small glass beads. It is also found that the goodness of fit decreases with an increase in the glass beads mass percentage. For example, R-square and RMSE are 0.9376 and 0.0678, respectively for 50% small glass beads and 0.9655 and 0.0549, respectively for 50% large glass beads after 300 minutes. The highest deviation in the PSD compared to experimental results presented in Figure 4.3c can be found in the ranges of 0.212 – 0.6 mm and 0.6 – 1 mm. When small glass beads were involved in the co-fluidization, measurement of the mass of IBA particles in the size range of 0.212 – 0.6 mm

alone was difficult. As described in the experimental section, the mass of IBA particles in this size range was calculated by deducting the mass of small glass beads from the total mass of particles obtained in this size range. Hence, the mass of IBA particles in the size range of 0.212 – 0.6 mm is always under-estimated since it is assumed that there is no loss of small glass beads in the size range of 0.6 – 1 mm. As a result, it caused the highest deviation between the fitted PSD and experimental data in the size range of 0.212 – 0.6 mm and 0.6 – 1 mm shown in Figure 4.3c. Nonetheless, all the R-square and RMSE values are above 0.93 and below 0.07, respectively for all the conditions. It indicates that RR distribution is a suitable model to describe the PSD in gas-solid fluidized bed. Therefore, the parameters of RR PSD (d and n) will be used as the modeling parameters in the development of surrogate models for the PSD during fluidization.

a) Original IBA PSD at time = 0 min

b) Pure IBA PSD at time = 30 min

c) IBA PSD at time = 300 min with using 50% small glass beads

Figure 4. 3. Fitting of PSD using RR distribution function: a) Original IBA PSD with, b) Pure IBA PSD at time = 30 min, c) IBA PSD at time = 300 min with using 50% small glass beads

4.5.2 Accuracy and validation of surrogate model

In the building of the surrogate models, 267 training points were used 14 testing points were reserved for validation. Leave-one-out cross-validation has also been used as accuracy evaluation for surrogate models. It is to note that two different accuracies were involved in the surrogate model development. One relates to the accuracy of the surrogate models to the fitted RR parameters based on experimental data while the other refers to the accuracy of the surrogate models

to the actual experimental results. A comparison between the model-predicted PSD and experimental results indicate that the largest error is present in the conditions of co-fluidization with small glass beads. Based the results of RR fittings on experimental data in section 4.5.1, the worst fitting case also occurs in the conditions of co-fluidization with small glass beads. Therefore, it can be concluded that the error between the predicted PSD and experimental results are originated from the error involved in the fitting of RR distributions before the development of the surrogate models. On the other hand, a comparison of the model predicted RR parameters (d and n) and the training data (fitted RR parameters) is shown in Figure 4.4. Both d and n were found to have good agreement between the model prediction and the training data. It also can be seen in Figure 4.4 that the prediction on d is slightly better than on n . In addition, the highest predicted errors for d and n are at large d and n values where the number of training data is low. In fact, the number and spread of training data play a vital role on the accuracy of developed surrogate models.

a) Predicted d versus fitted d

b) Predicted n versus fitted n

Figure 4. 4. Parity plots of experimental and predicted RR parameters values calculated by the models for training data

Table 4.3. Validation results of surrogate models

Models	d				n	
	Cross validation	Training points	Testing Points	Cross validation	Training points	Testing Points
R square	0.9643 ± 0.02654	0.9863	0.9717	0.9523 ± 0.0400	0.9675	0.9421
RMSE	0.0047 ± 0.0021	0.0023	0.0049	0.0118 ± 0.0039	0.0081	0.0137

Validation of the surrogate models developed is also performed using 14 testing points and the R-square and RSME results for both training points and testing points are shown in Table 4.3. It can be seen that while the error incurred in the testing points are higher than that in the training points, the low RMSE involved indicates that the surrogate models can predict the values of both d and n reasonably well. Similar to the parity plot shown in Figure 4.4, the predicted values for d are more accurate than n by the ANN models developed. In addition to the examination of R-square and RMSE values of the testing points, the specific testing points having the highest and lowest accuracies are then examined in detail. The fluidization of pure IBA at 240 minute and co-

fluidization of IBA with 25% large glass beads at 120 minute were found to be the conditions having the highest and lowest accuracies, respectively. For comparison, the experimental data, the RR PSD fitted to the experimental data and model predicted RR PSD are presented. Figure 4.5a shows that even the lowest accuracy for the condition of fluidization of pure IBA at 240 min has good agreement between the predicted and fitted PSDs. Figure 4.5b illustrates the medium accuracy condition. While the RR PSD fitted shows low accuracy for co-fluidization with small glass beads (as stated in section 4.5.1), the predicted PSD can forecast it with high accuracy. The highest accuracy condition shown in Figure 4.5c shows almost identical predicted and fitted PSDs. It demonstrates that the ANN-GA model developed in this work is suitable for modeling of the PSD at different fluidization conditions.

a) Pure IBA at time 240 minutes

b) IBA with 50% small glass beads at time 180 minutes

c) IBA with 25% large glass beads at time 120 minutes

Figure 4. 5. Experimental data of IBA particle size with fitted and predicted RR distribution from ANN models

4.5.3 Effect of glass beads on particle attrition

In order to study the inter-relationship between the glass beads size and the glass beads percentage, three-dimensional surface plots were developed based on the ANN models to depict the effect of glass beads on particle attrition. Particle attrition is represented by the change in PSD based on the PSD function

parameters, d and n based on RR distribution. It can be seen in Figures 4.6 and 4.7 that the surface plots can resemble the experimental behaviors rather accurately. In general, it can be seen from Figures 4.6 and 4.7 that both d and n decrease significantly initially and rate of decrease reduces over time. It is reasonable that the continuous collision of particles during fluidization causes IBA particles to constantly decrease with time. Abrasion and fragmentation are two main attrition modes in fluidization [253, 254]. Abrasion produces fine particles by breaking the edge and cracks as well as removing the roughness on the particle surface [255]. In this case, the particle surface smoothens but there is negligible change in the particle size. Fragmentation, on the other hand, breaks particle into smaller pieces through internal defects. As a result, particles size reduction is significant.

Figure 4. 6. Three-dimensional surfaces of ANN models for d as a function of the time and glass beads percentage at $d_0 = 0.9$ and $n_0 = 1.6$ for a) small glass beads; b) large glass beads

As shown in Figures 4.6 and 4.7, both d and n decrease significantly in the beginning and the rate of reduction decreases over time. It indicates that

fragmentation is the governing mechanism in the beginning of fluidization. When most of the particles with defects are fragmented, the abrasion mechanism becomes dominating and the fluidization process tends to generate more fine particles without significant change to the particle size.

Figure 4. 7. Three-dimensional surfaces of ANN models for n as a function of the time and glass beads percentage at $d_0 = 0.9$ and $n_0 = 1.6$ a) small glass beads; b) large glass beads

The comparison of effect of small and large glass beads as grinding medium are also shown in Figures 4.6 and 4.7. It can be seen that the use of an appropriate mass percentage of small glass beads in the co-fluidization process can enhance particle attrition, indicated by the increased rate of reduction of d and n . On the other hand, the use of large glass beads lessened both the rate and extent of particle attrition. While a study reported that attrition slightly increases with an increase in the particle size of grinding medium [325], the range of particle size studied was limited. In reference to a ball milling comminution process, the difference in behavior of co-fluidization with small and large glass beads can be explained by the concept of contact points. Based on statistical method, the

number of contact points is governed by the size of the particles [326, 327]. The number of contact points increases with a reduction in ball size. It was also reported that the number of contact points and grinding effects are at the maximum when the ball size is the same as the feed particle size [326-332]. In this study, the size of the small glass beads used were 0.212 – 0.6 mm and similar to the IBA particles. Therefore, co-fluidization with small glass beads contributes to a higher number of contact points and was observed to be more effective particle attrition compared to co-fluidization with large glass beads that are 1.7 mm – 2 mm in size. However, the enhancement of particle attrition with the co-fluidization with small glass beads diminishes when the small glass beads percentage is higher than 25%. It is believed that, similar to a ball milling process, there is an optimum particle filling ratio that maximizes the collisions between particles and glass beads [333, 334]. When the particle filling is below the optimum filling ratio, glass bead-glass bead collision dominates over particle-glass bead collision. When the particle filling exceeds the optimum filling ratio, particle-particle collision prevails. As particle-glass bead collision provides more effective attrition than particle-particle and glass bead-glass bead collisions, optimal particle attrition performance can be observed when particle-glass bead collision is maximized [333, 334].

It was found in ball milling process that the use of grinding medium can enhance particle attrition. However, this enhancement in particle attrition not observed when large glass beads were used in the co-fluidization process, as illustrated in Figures 4.6b and 4.7b. The difference in behavior for the involvement of grinding medium between ball milling and co-fluidization can be attributed to the differences in mode of operations. The ball milling process rotates the particles

involved regardless of the sizes and densities of the particles. On the other hand, the fluidization requires the use of upward flowing gas to overcome the gravitational force acting on the particles, in which the particle sizes and densities are crucial. While the densities of large glass beads and IBA particles are similar, the involvement of the large particle size of large glass beads increased the minimum fluidization velocity required for the particle mixture. Moreover, there are studies reported that particle segregation can occur when there is a binary mixture of particles of different size distributions [335, 336]. Hence, co-fluidization of large glass beads and IBA particles mixture can yield particle segregation and uneven fluidization that the movement of IBA particles is significantly higher than that of large glass beads. Therefore, it is possible that co-fluidization with large glass beads promotes particle-particle collision rather than particle-glass bead collision and the particle attrition decreases with an increase in large glass beads percentage.

4.6 Summary

Particle size distribution (PSD) is an important parameter in gas-solid fluidized bed. The change in PSD due to particle attrition can affect the long-term performance of fluidized bed. In this study, artificial neural network (ANN) with genetic algorithm (GA) as a surrogate modeling tool was employed to model the change in PSD during fluidization. Experiments were conducted using incineration bottom ash (IBA) as the fluidizing particles and different mass percentage of large and small glass beads were used as the grinding medium. Rosin-Rammler (RR) distribution was used to describe the IBA PSD. The ANN-GA models developed were subsequently used to study the effect of fluidization

time, mass percentage of glass beads and size of glass beads used on the IBA particle attrition during fluidization.

The RR distribution is able to describe the PSD reasonably well. The results show that the developed ANN models can provide good representation of the change in PSD during fluidization. Effect of fluidization time, mass percentage of glass beads and size of glass beads used on the IBA particle attrition during fluidization were also studied using a response surface methodology. In general, both d and n decrease with fluidization time. An increase in the mass percentage of glass beads having similar size range as the IBA particles in the co-fluidization process can enhance the particle attrition until an optimum filling ratio of the space between the glass beads is reached. Beyond this point, an increase in the mass percentage of glass beads reduces the filling ratio and the effect of particle attrition reduces. The use of glass beads larger than the IBA particles can cause uneven fluidization. Collision between IBA particles is enhanced with an increase in the mass percentage of large glass beads and leads to a reduction in IBA particle attrition.

Chapter 5 Modeling of particle size distribution in a gas-solid fluidized bed by planetary ball milling results using a hybrid artificial neural network-genetic algorithm approach

5.1 Preface

As it was mentioned in the previous chapter, PSD in fluidized beds is dependent on many variables such as material properties, initial particles size, time, and fluidization conditions such as temperature, pressure, gas velocity, and fluidized bed geometry. Due to the large number of variables involved, it is not feasible to use a simple mathematical relation to describe the change of PSD during fluidization. ANN was successfully applied to predict the change of PSD during fluidization [337].

However, only one material is involved in this study and the applicability is limited. Therefore, this chapter aims to develop a model for the prediction of the change in PSD during fluidization for a range of materials and gas velocities using ANN. It is believed that material properties such as hardness, density, brittleness, structure, etc. play a crucial role in the particle attrition behavior. Unfortunately, little information on material properties is available, considering the wide variety of materials presence. As a result, planetary ball milling is proposed as a fast assessment technique to identify the properties of different materials and subsequently implement in the ANN models of particle attrition. Planetary ball milling devices are readily available in most laboratories and continuous change in PSD can resemble the particle attrition process during fluidization. A Rosin-Rammler (RR) distribution was used to describe the PSD for both fluidization and ball milling processes.

5.2 Experimental setup

Fluidization experiments were conducted in an acrylic fluidized bed which was reported in our previous chapter [259]. The schematic diagram of the experimental setup is shown in Figure 4.1. The fluidized bed internal diameter and height are 30 mm and 1200 mm, respectively. A perforated plate was installed in the bottom of fluidized bed to serve as the gas distributor and support the solids. Compressed air was used as the fluidizing medium. A pressure regulator controlled the air pressure and the gas velocity was adjusted by a rotameter. The gas velocities were varied from 0.73 m/s to 1.34 m/s. Experimental results using three different materials, including activated graphite, gypsum, and silica were used as training data and experimental results using IBA were used for model validation. Starting particle size of all materials was kept below 1.7 mm for uniform fluidization within the operating gas velocity range.

For ball milling, materials were ground in a Planetary Mill Pulverisette 5 (Fritsch, Germany) consisting of two grinding bowls. Each grinding bowl is 80 mL and five agate grinding balls of 10 mm diameter were used in each bowl. The samples were milled at 170 RPM and a constant fill ratio of 75% was maintained. The materials were milled until no further reduction in size was observed. Typical operating period for graphite, gypsum, silica and IBA were 10 minutes, 70 minutes, 120 minutes, and 180 minutes, respectively.

At various fluidization and ball milling times, the experiments were stopped and the PSD of the materials particles were measured using a sieving machine (RETSCH vibratory sieve shakers AS200) using 5 sieves of mesh sizes 0.212 mm, 0.6 mm, 1.0 mm, 1.4 mm, and 1.7 mm.

5.3 Data collection

The change in PSD from ball milling was first fitted to a RR distribution by least-square method. The d and n obtained was then fitting to the following equations, respectively:

$$\frac{d - d_f}{d_0 - d_f} = \exp(-B_d \times t) \quad (5.1)$$

$$\frac{n - n_f}{n_0 - n_f} = \exp(-B_n \times t) \quad (5.2)$$

Where d_0 and n_0 are RR parameters of initial PSD, d_f and n_f are RR parameters of final PSD, t is time, and d and n are RR parameters at a particular time. B_d and B_n are the fitted parameters that are used to describe the material properties.

The PSDs obtained from fluidization experiments were also fitted to a RR distribution by least-square method and the d and n obtained were used directly in the construction of ANN models.

Table 5.1. Input and output variables and their range of values

Input variables	Symbols	Range of values
time (min)	In_1	10.00 – 1200
U_g	In_2	0.73 – 1.34
B_d or B_n	In_3	0.06517 – 0.5875
d_0	In_4	0.07185 – 0.8146
n_0	In_5	0.52746 – 0.9105
Output variables		
d	Out_1	0.82429 – 1.8146
n	Out_2	0.51265 – 0.89881

The operating conditions such as fluidization time (time), gas velocity (U_g), material properties (B_d and B_n), and initial PSD (d_0 and n_0) were considered as the input variables. PSD at different fluidization time (d and n) were used as the output variables. Two ANN models were constructed separately to determine d and n . The input variables and their range of values are given in Table 5.1.

5.4 Accuracy and validation of surrogate model

Leave-one-out cross-validation has been used as accuracy evaluation for surrogate models. In addition, the accuracy of the constructed ANN models was validated using 80 test data collected from experiments using graphite, gypsum, and silica and 6 test data collected from experiments using IBA. Root-mean-square error (RMSE) and R^2 (R-square) were used for the evaluation of model accuracy and validation.

5.5 Genetic algorithms (GA) design for different purposes

The parameters used in GA for ANN optimization and SR are listed in Table 5.2. The parameters were obtained by trial and error using some known useful defaults as initial values [338].

Table 5.2. GA parameters for variable selection and ANN optimization

GA parameters	ANN optimization	SR
Population size	120	400
Maximum generation	140	200
Elite count	10	15
Crossover fraction	0.7	0.85
Selection technique	Stochastic uniform	Stochastic uniform
Crossover method	Scattered	Scattered
Mutation function	Gaussian	Gaussian
Stopping criteria	Maximum generation	Maximum generation

5.6 Results and Discussion

5.6.1 Application of the Rosin–Rammler model to the particle size distribution analysis in ball milling and fluidization

Figure 5.1 and 5.2 show the experimental PSDs as well as the best and worst fitted RR PSDs for ball milling and fluidization processes, respectively. The histogram represents the experimental result and the curve represents the fitted RR PSD. The goodness of fit was presented in terms of R-square and RMSE values. The histogram of experimental PSD is represented by five particle size ranges between 0 and 1.7 mm corresponding to the sieves used in the sieving machine.

The best-fitting case by RR PSD in ball milling is for silica after 108 minutes of ball milling (Figure 5.1a) and the worst one is for ball milling of gypsum particles after 120 minutes (Figure 5.1b). Ball milling time has insignificant impact on the goodness of fit for RR PSD. For instance, R-square and RMSE are 0.9946 and 0.0392, respectively at time = 0 and they are 0.9897 and 0.0461, respectively after 120 minutes for the ball milling of graphite particles.

The best-fitting case by RR PSD in fluidization is for silica after 240 minutes of fluidization with gas velocity (U_g) 1.3 m/s (Figure 5.2a) and the worst one is for fluidization of gypsum particles after 300 minutes at a gas velocity (U_g) of 0.73 m/s (Figure 5.2b). Similar to ball milling, fluidization time has a negligible effect on the goodness of fit for RR PSD. For example, R-square and RMSE are 0.9963 and 0.0240, respectively at time = 0 and they are 0.9748 and 0.0547, respectively after 1200 minutes for the fluidization of activated carbon particles with gas velocity (U_g) 1.23 m/s. Nonetheless, all the R-square and RMSE values for both ball milling and fluidization are above 0.97 and below 0.06, respectively for all

the conditions. It indicates that RR distribution is an appropriate model to describe the PSD in ball milling and gas-solid fluidized bed.

Figure 5. 1. Fitting of PSD in ball milling using RR distribution function: a) Silica PSD at time = 108 min, b) Gypsum PSD at time = 120 min

Figure 5. 2. Fitting of PSD in fluidization using RR distribution function: a) Silica PSD at time = 240 min and $U_g = 1.3$ m/s, b) Activated carbon PSD at time = 300 min and $U_g = 0.73$ m/s

5.6.2 Determination of attrition related material properties by ball milling

d and n decrease monotonically with ball milling time. Therefore, the fitted constants, B_d and B_n can be considered as the decay constants for d and n , respectively. For example, Figure 5.3 shows the change in d and n in ball milling for gypsum and determination of the fitted constants. The B_d and B_n determined for all the materials used in the study are shown in Table 5.3. Furthermore, Figure 5.4 shows B_d and B_n versus materials hardness. IBA hardness was found roughly according to major chemical content in IBA [259]. As it is seen in Table 5.3 and Figure 5.4, B_d and B_n reduce with an increase in material hardness. For example, while B_d and B_n are 0.5875 and 0.8146 for activated carbon as the softest material, B_d and B_n are 0.06517 and 0.07185 for silica as the hardest material. The decay parameters reduction is reasonable because hard materials tend to shrink less.

Figure 5. 3. a) Change of d and n in ball milling for gypsum, b) determination of B_d and B_n for gypsum

Table 5.3. Calculated materials attrition properties by ball milling

	Activated carbon	Gypsum	Silica	IBA
Materials hardness (Mohs)	0.5	2	7	4.5 - 5.5
B_d	0.5875	0.1836	0.06517	0.1016
B_n	0.8146	0.3175	0.07185	0.1245

Figure 5. 4. Comparison of materials hardness with B_d and B_n

5.6.3 Accuracy and Validation of ANN models

In this study, a four-layer (namely, one input layer, two hidden layers, and one output layer) network was used. The weights, biases and GA convergence graphs have been shown in Table A2 and Figure A2 in Appendix. 400 training points were used in ANN model construction. Furthermore, the accuracy of the constructed ANN models was validated using 80 test data collected from experiments using graphite, gypsum, and silica and 6 test data collected from experiments using IBA as well as leave-one-out cross-validation.

A comparison of the model predicted RR parameters (d and n) and the experimental data is shown in Figure 5.5. The R-square and RSME results for training and testing points are shown in Table 5.4. Both d and n were found to have good agreement between the model prediction and the testing data. It can be seen in both Figure 5.5 and Table 5.4 that the prediction of d is slightly better than that of n for all training and testing points. It can be seen in Table 5.4 that while the testing points have lower R-square and higher RMSE than that in the training points, but according to Alexander et al. [218] model fitness criteria the surrogate models can predict the values of both d and n reasonably well. Alexander et al. [218] have recommended high R-square (e.g. $R^2 > 0.6$) and reasonable RMSE for testing points against training points as criteria for suitable model for predictive purposes.

Figure 5. 5. Parity plots of experimental and predicted RR parameters values calculated by the models for testing data

Table 5.4. Accuracy and validation results of ANN models

Models	d				n			
	Error metrics	Cross validation	Training points	Testing Points	IBA testing Points	Cross validation	Training points	Testing Points
R^2		0.9684 ± 0.0295	0.9811	0.9670	0.9064	0.9589 ± 0.02063	0.9684	0.9509
RMSE		0.0078 ± 0.0046	0.0043	0.0097	0.4045	0.0131 ± 0.0041	0.0107	0.0167

In addition to the examination of R-square and RMSE values of test data collected from experiments using graphite, gypsum, and silica, test data collected from experiments using IBA are used. The specific graphite, gypsum, silica and IBA testing points having the highest, medium, and lowest error is then examined in detail in Figure 5.6 and Figure 5.7. It is to note that two different accuracies were involved in the ANN model development. One relates to the accuracy of the ANN models to the fitted RR parameters (it is shown in Figure 5.6 and Figure 5.7) while the other refers to the accuracy of the ANN models to the experimental results. Among the graphite, gypsum, and silica testing points, the fluidization of silica at 180 minute with gas velocity 1.3 m/s and the fluidization of activated carbon at 60 minute with gas velocity 1.23 m/s were found to be the conditions having the highest and lowest accuracies, respectively. For comparison, the experimental data, the RR PSD fitted to the experimental data and model predicted RR PSD are presented. The highest accuracy condition shown in Figure 5.6a shows almost identical predicted and fitted PSDs. Figure 5.6b presents Gypsum PSD at 1200 minute with gas velocity 0.84 m/s as medium condition in PSD prediction. Figure 5.6c shows that even the lowest accuracy for the condition of fluidization of activated carbon at 60 min with gas velocity 1.23 m/s has good agreement between the predicted and fitted PSDs. The error metrics analyzing shows that the error between the experimental data and fitted RR is

higher than that between the fitted and predicted RR. For example, the R-square and RMSE between the experimental data and fitted RR is 0.9778 and 0.0588 and between fitted and predicted RR is 0.9897 and 0.0038 for fluidization of activated carbon at 60 min with gas velocity 1.23 m/s.

(a)

(b)

(c)

Figure 5. 6. Experimental data of particle size with fitted and predicted RR distribution from ANN models for testing points: a) Silica PSD at time = 180 min and $U_g = 1.3$ m/s as the best prediction, b) Gypsum PSD at time = 1200 min and $U_g = 0.84$ m/s as the medium prediction, c) Activated carbon PSD at time = 60 min and $U_g = 1.23$ m/s as the worst prediction

Figure 5. 7. Experimental data of IBA particle size with fitted and predicted RR distribution from ANN models for IBA testing points: a) IBA PSD at time = 60 min and $U_g = 0.78$ m/s as the best prediction, b) IBA PSD at time = 180 min and $U_g = 0.78$ m/s as the medium prediction c) IBA PSD at time = 1200 min and $U_g = 0.78$ m/s as the worst prediction

IBA fluidization results with gas velocity 0.78 m/s were also used as additional testing points which are not in training data. The fluidization of IBA at 60 minute and 1200 minute were found to have the highest and lowest accuracies among the IBA testing points, respectively. In addition, the accuracy with medium

condition has been found for fluidization of IBA at 180 minute. Similar to graphite, gypsum, and silica testing points, the experimental data, the RR PSD fitted to the experimental data and model predicted RR PSD are presented in Figure 5.7 for IBA. The highest, medium, and lowest accuracy condition are shown in Figure 5.7a, 5.7b, and 5.7c, respectively. According to Alexander et al [218] criteria, both shown R-square and RMSE are acceptable for IBA testing points. It demonstrates that the ANN-GA model developed in this work is suitable for modeling of the PSD at different fluidization conditions.

5.6.4 Symbolic regression (SR) of d and n

Although ANN method can be used to develop an accurate model for prediction of PSD during fluidization, but ANN developed equation is complex and impractical. On the other hand, finding an equation for d and n can be used for rough estimation of PSD during fluidization. Therefore, SR which can find a simple equation for complex problems is used. GA as a powerful SR method is utilized for finding these equations for d and n . The input variables for d and n from experimental data are utilized for building equation by SR. Following equations have been developed by GA for d and n , respectively. The input variables numbers have been marked-up base on Table 5.1. R-square and RMSE are 0.9486 and 0.0608 for d and 0.9112 and 0.0701 for n . A parity plot for comparison of fitted and calculated d and n based on developed equations has been shown in Figure 5.8. This figure indicates that the equations can calculate the trend of the data pretty well. Therefore, these equations can be used for rough computation of d and n for PSD prediction during fluidization.

$$d = \frac{0.06648 \times In_2}{In_5} - \frac{0.5124}{In_4} + 0.2215 \times In_2^5 \times In_3^2 \times In_4^2 \times (In_1 \times 10^{-3} + In_4)^2 \times (In_1 \times 10^{-3} + In_2) - \frac{0.01911 \times In_1 \times In_3 \times In_4 \times (In_1 \times 10^{-3} - 6.321) \times (2 \times In_3 - 9.209 \times In_4 + In_5 + In_2 \times In_3 \times In_5)}{In_2} + 1.411 \quad (5.3)$$

$$n = 0.3671 \times In_2 - 0.1836 \times (In_3 \times In_5 \times In_2^3 + In_3 + In_1 \times 10^{-3} + \exp(In_4)) - 0.3671 \times In_4 - \frac{2.649 \times In_4}{In_5 \times (In_4 + In_5)} + 2.588 \quad (5.4)$$

a) Calculated d versus fitted d

b) Calculated n versus fitted n

Figure 5. 8. Parity plots of real and calculated d and n by the developed equations by GA SR

5.7 Summary

Only one material was involved in the previous chapter and the applicability was limited. Therefore, in this chapter, experiments were conducted using three different materials data, including activated carbon (graphite), gypsum, and silica as training data and IBA was used as validating data. Planetary ball milling method has been proposed as an approach to determine the material properties. Rosin-Rammler (RR) distribution was used to describe the material PSD. The ANN-GA models developed were subsequently used to study the effect of gas

velocity, material property, and initial PSD on the material particle attrition during fluidization. Finally, symbolic regression (SR) was applied to find a simple equation to determine RR parameters during fluidization.

The RR distribution is able to describe the PSD in ball milling and fluidization reasonably well. The fitted constants, B_d and B_n , which was obtained from ball milling results, can be considered as the decay constants for d and n , respectively. B_d and B_n reduce with an increase in material hardness. The results of cross validation and testing points show that both d and n were found to have good agreement between the model prediction and the experimental data. IBA fluidization results with gas velocity 0.78 m/s were also used as validating points, which are not in training and testing data. The results show that the developed ANN-GA model in this work is valid for modeling of the PSD at different fluidization conditions.

Chapter 6 Evaluation of hydroxyapatite size and morphology in dry powder inhalation for carrier-based pulmonary delivery formulations by response surface methodology

6.1 Preface

To present application of surrogate modeling in biomedical processes, surrogate modeling of dry powder inhalation (DPI) are investigated in this chapter. One of the aerosol-generating devices used for inhaled drug delivery is DPI. DPI consists of micronized drug particles, which are inhaled as a cloud with or without the aid of carrier particles [339-341]. Drug particles having aerodynamic diameters between 2–4 μm were found to ideal for deposition in the peripheral airway [342, 343]. However, particles of this size range are cohesive in nature that exhibits to poor dispersibility [344] and flowability [345]. A mixture of large carrier particles 20–200 μm in size with the drug particles was able to drug aggregation problem and improve the drug delivery efficiency [346-352].

Similar to gas-solid fluidized bed, dry powder inhalation (DPI) is a process between gas and solid phases (carrier with the drug is solid phase in DPI). In simulation of DPI by physical equations, there is a similar issue with simulation of PSD in fluidized beds. Due to a large number of particle-particle interactions (there is an equation for each interaction), study the fluidization process of a powder bed is computationally expensive.

In vitro aerosolization performance and drug delivery efficiency of a DPI is generally represented by emitted dose (ED) and fine particle fraction (FPF) [349, 353]. ED represents the powder flow behavior while FPF characterizes the aerosolization behavior of the drug particles. Certain factors such as inhalation

flow rate and carrier-to-drug mixing ratio that affect ED and FPF are typically well studied and reported in the literature. An increase in the inhalation flow rate increases both the ED and FPF [354, 355]. An increase in carrier-to-drug mixing ratio was reported to decrease the FPF until a carrier-to-drug mixing ratio of 20:1, beyond which a further increase in carrier-to-drug mixing ratio increases the FPF [356-359].

The physical properties of the carriers such as particle size, shape, surface roughness, etc and other chemical properties of the carrier material have a significant impact on ED and FPF of carrier-based DPI [355, 360, 361]. However, findings reported in the literature for these factors are not conclusive. For example, FPF was reported to increase with a decrease in carrier particle size [31, 355-357, 362-369]. On the other hand, there are also findings showing FPF to increase with an increase in carrier particle size [370, 371]. Similarly, good FPF were demonstrated in carriers that have smooth surfaces [372, 373] and carriers that have rough surfaces [374]. The relationship among the carrier properties, ED, and FPF, are highly complex and often interdependent. In a recent study, an empirical model was proposed to predict the FPF for carrier-based DPI by taking into account the contribution of carrier surface roughness in addition to [52]. However, the understanding of the actual physical relationship between the carrier properties and FPF remains unclear. On the other hand, finding the effect of variables interaction on the efficiency of DPI by experiments is not possible because usually change in one variable, will change other variables inevitably. For example, change in carrier particle size will change the carrier surface roughness.

This chapter aims to develop artificial neural network (ANN) models to enhance

the understanding of the relationship between various carrier properties with ED and FPF. ANN serves as an excellent tool for surrogate modeling that allows a better understanding of behaviors among inter-related parameters [2, 6, 13, 27, 56, 119, 146, 235, 316-322]. In addition to standard parameters such as carrier average particle size, bulk and tap densities, inhalation flow rate, and carrier-to-drug mixing ratio, parameters such as the dispersion of the carrier particle size distribution, fine carrier fractions, and various descriptive parameters for carrier surface roughness were also used as input in the development of the ANN models. Information regarding the carrier particle size distribution is believed to play an important role in the behavior of ED and FPF. It was reported that the addition of a suitable amount of fine particles in the formulation is effective in enhancing FPF [31, 366, 375-377]. The impact of carrier surface roughness not only limits to the adhesion forces between drug and carrier particles [370, 378], the scale of the carrier surface roughness can also affect the entrapment of drug particles on the carrier surface [379, 380]. Therefore, it is necessary to explore the suitability of different parameters to describe the carrier surface roughness [374, 381, 382]. Genetic algorithm (GA), an evolutionary algorithm, is applied to select variables important for ANN models and to determine the ANN parameters. Sensitivity analysis was also carried out to determine the relative importance of different parameters to ED and FPF.

6.2 Dataset

All the carrier formulations and aerosolization properties used in this study were obtained from published literature using the hydroxyapatite as carrier material [359, 371, 383, 384].

Table 6.1. Complete list of input and output variables considered in the study

Variable number	Description of variables	Variable group
Input variables		
I1	Carrier particle average size	Carrier size distribution
I2	Carrier particle size standard deviation	
I3	Span of volume distribution	
I4	Particle diameter at 10% of the volume distribution (D10)	
I5	Particle diameter at 50% of the volume distribution (D50)	
I6	Particle diameter at 90% of the volume distribution (D90)	
I7	Volume fraction of carrier particles having a particle size smaller than 5 μm	
I8	Volume fraction of carrier particles having a particle size smaller than 10 μm	
I9	Volume fraction of carrier particles having a particle size smaller than 15 μm	
I10	Volume fraction of carrier particles having a particle size smaller than 20 μm	
I11	Tap density	Carrier density
I12	Bulk density	
I13	Carr's Compressibility Index (CI)	
I14	Flow rate	Operating conditions
I15	Carrier-to-drug ratio	
I16	Root mean square deviation (Rq)	Carrier surface characteristics
I17	Arithmetical mean deviation (Ra)	
I18	Skewness of the assessed profile (Rsk)	
I19	Kurtosis of the assessed profile (Rku)	
I20	Lowest valley (Rv)	
I21	Highest peak (Rp)	
I22	Total height of the profile (Rt)	
I23	Average height of an unleveled surface (Rc)	
I24	Mean polar facet orientation (FPO)	
I25	Variation of the polar facet orientation (MFOV)	
I26	Direction of azimuthal facets (FAD)	
I27	Mean resultant vector (MRV)	
I28	Surface area (SA)	
Output variables		
O1	Fine particle fraction (FPF)	
O2	Emitted dose (ED)	

Parameters extracted include ED, FPF, flow rate, carrier-to-drug ratio, tap density, bulk density, CI, and carrier particle size distribution, represented by mean and variance of a fitted Gaussian distribution function. In addition, some studies show that the presence of carrier particles can improve DPI efficiency [354]. Hence, the volume fractions of particles less than a certain size are selected for investigation. In addition, various roughness parameters [349] of the carriers are also used since carrier surface roughness has both microscopic and macroscopic influence on the performance of DPI [380]. There are a total of 30 variables considered in the study and a complete list of input and output variables is shown in Table 6.1. ED and FPF are the only two output variables while the rest of the other 28 are input variables containing information about the formulation such as carrier size, density, roughness variables, flowability, drug content, and inhalation flow rate.

6.3 Surface and shape analysis

The carrier surface and shape analysis were performed based on image analysis of scanning electron microscopy (SEM) images using ImageJ® software (v2.33) [385, 386] with SurfCharJ (v.1q) plugin [387, 388]. SEM images of carrier particles were rescaled to a standardized resolution of 12.4 pixels per micron. A square area of 200 pixels x 200 pixels in the center region of each carrier particle in the image was cropped for surface characterization. More than three cropped images were used to obtain an average value for each surface variable for each carrier particle. Thirteen roughness variables describing the surface properties were calculated with SurfCharJ (v.1q) plugin. Surface roughness variables include root mean square deviation (R_q), arithmetical mean deviation (R_a), skewness of the assessed profile (R_{sk}), kurtosis of the assessed profile (R_{ku}),

lowest valley (R_v), highest peak (R_p), total height of the profile (R_t), average height of an unlevelled surface (R_c), and surface orientation variables including mean polar facet orientation (FPO), variation of the polar facet orientation (MFOV), direction of azimuthal facets (FAD), mean resultant vector (MRV), and surface area (SA).

6.4 Design of ANN

In this chapter, the ANN models were built using 20 training samples and additional 5 samples were used as test sets. The testing samples were selected randomly to cover the entire range of input variables. In addition, leave-one-out cross-validation is used to evaluate developed ANN models accuracy. A four-layer (namely, one input layer, two hidden layers, and one output layer) network was used. The optimum number of hidden layers was found by trial and error following the general concept that the number of hidden layers should be between the number of input and output variables [240]. A publicly available Matlab SUMO toolbox (<http://www.sumowiki.intec.ugent.be>) [11, 323, 324] was employed. The toolbox was also used to normalize all the input and output data to scale between -1 and 1. A hybrid of the Gauss-Newton nonlinear regression method and gradient steepest descent method (Levenberg-Marquardt algorithm) were used to build the feed forward ANN. The weights and biases were generated randomly in the range of -1 and 1. The tansig (tan-sigmoid) transfer function was used between the input and hidden layers. The purelin (linear) function was used as the transfer function between the hidden and output layers. The network weights and biases were modified according to GA to achieve the global minima of deviation error, and optimum weights and biases were obtained simultaneously.

6.4 Genetic algorithms (GA) parameters

The parameters and stopping criteria used in GA for both variable selection and ANN optimization are listed in Table 6.2. The chosen parameters and stopping criteria were obtained by using some known useful defaults as initial values [338] and refined by trial and error. The weights, biases and GA convergence graphs have been shown in Figure A3 and Table A3 in Appendix.

Table 6.2. GA parameters for variable selection and ANN optimization

GA parameters	Variable selection	ANN optimization
Population size	50	112
Maximum generation	100	140
Elite count	5	10
Crossover fraction	0.8	0.7
Selection technique	Stochastic uniform	Stochastic uniform
Crossover method	Scattered	Scattered
Mutation function	Gaussian	Gaussian
Stopping criteria	Maximum generation	Maximum generation

6.5 Results and discussion

6.5.1 Analysis of surface roughness

A sample workflow on the analysis of surface roughness based on SEM images of carrier particles is shown in Figure 6.1. Based on the gray scale level of individual pixel of a cropped image, a three dimensions surface plot and various surface roughness variables can be determined using ImageJ® software (v2.33) with SurfCharJ (v.1q) plugin. It can be seen in Figure 6.1 that certain variation is present among different cropped images of the same particle. Therefore, an average value from a triplicate of cropped images for the same particle formulation was used to represent each roughness variable in the construction of ANN models.

Figure 6. 1. A sample workflow of surface roughness analysis: SEM image, 2D, 3D surface plots, and surface properties

6.5.2 Selection of important variables

Variable selection is a method for input variables reduction. The results of the PLS toolbox were used for selection of the most relevant variables. Since ED and FPF are two independent outputs in this study, variable selection was carried out separately for ED, and FPF and two different ANN models were developed. Figure 6.2 shows the variable selection results for ED and FPF. The horizontal axis is the variable number based on Table 6.1, and the vertical axis indicates the frequency of inclusion of each variable in the population. Since the populations having high fitness values were used to produce the next generation in GA, the variables that have high frequency of inclusion in the generations can be considered significant in surrogate modeling. For more confidence, variables included in at least 10 percent of generations were used in the development of surrogate models.

Among the input variables, Figure 6.2a demonstrates that the carrier particle average size (I1), carrier particles size standard deviation (I2), volume fraction of carrier particles less than 5 micron (I7), tap density (I11), flow rate (I14), carrier-to-drug ratio (I15) are significant in the prediction of ED. It is also found that Rq (I16), Ra (I17), Rp (I21), Rc (I23), MFOV (I25), MRV (I27), and SA (I28) are the more descriptive variables to represent the impact of carrier surface characteristics on ED. For FPF, Figure 6.2b indicates that carrier particle average size (I1), carrier particles size standard deviation (I2), the volume fraction of carrier particles less than 10 micron (I3), tap density (I11), flow rate (I14), carrier-to-drug ratio (I15), Rq (I16), Ra (I17), Rsk (I18), Rku (I19), FPO (I24), and SA (I28) have the most contribution in the model development. Carrier particle average size (I1) has always been considered one of the most important variables

affecting ED and FPF [349, 389-391]. In addition, the selection of the carrier particles size standard deviation (I2) and volume fraction of particles less than 5 (I7) and 10 (I8) micron are in agreement with literature findings that the presence of fine carrier particles can improve FPF [354]. Fine carrier particles can have different impact on ED and FPF. Particles less than 5 micron (I7) in size are cohesive and reduce the lost of drug particles in inhalers. However, the cohesive nature of the fine carrier particles also demonstrates poor dispersability and flowability that are not favorable for FPF. Therefore, the fine carrier particle size selection criteria can be different between ED and FPF. In agreement with previous studies [359, 389, 392], the result of variables selection by GA shows that tap density (I11) as well as flow rate (I14) and carrier-to-drug ratio (I15) have significant effects on ED and FPF.

While Rq (I16) and Ra (I17) are often considered as important variables to describe surface roughness [349], the use of other carrier surface variables in literature are not common. Rp (I21) and Rc (I23), which show maximum peak height and mean height of roughness, respectively, were found to be significant for ED. On the other than, the symmetry of peaks and valleys (Rsk) (I18) and the sharpness of the height distribution (Rku) (I19) appear to be significant for FPF. In terms of surface roughness orientation analysis, MFOV (I25) and MRV (I27) which express variation of the polar facet orientation and mean resultant vector are important for ED. FPO (I24) that is mean polar facet orientation was selected for FPF. The SA (I28) that closely associates to the active carrier surface plays an essential role in DPI process would also be significant for both ED and FPF.

Figure 6. 2. Frequency of variable usage in models; a) ED, b) FPF

6.5.3 Design of the ANN model for prediction of FPF and ED

In the building of the surrogate models, 20 training points were used and 5 randomly selected testing points were reserved for validation. Leave-one-out cross-validation have been used as accuracy evaluation for surrogate models. A comparison of the model predicted ED and FPF and the respective training data is shown in Figure 6.3. Both ED and FPF were found to have good agreement between the model prediction and the training data. It also can be seen in Figure 6.3 that the prediction on FPF is better than on ED. In addition, the highest predicted errors for ED and FPF are at low ED and high FPF values where the number of training data is low. However, the number and spread of training data are often limited by the feasibility and availability of experiments in the desirable conditions.

Validation of the surrogate models developed is performed using 5 testing points and the R-square and RSME results for both training points and testing points are shown in Table 6.3. $R^2 > 0.6$ and reasonable RMSE were recommended as

criteria for successful modeling [218]. It can be seen that while the error incurred in the testing points are higher than that in the training points, the criteria for successful modeling are fulfilled reasonably well for both ED and FPF.

Table 6.3. Validation results of surrogate models

Models	ED			FPF		
Error metrics	Cross validation	Training points	Testing Points	Cross validation	Training points	Testing Points
R square	0.7948 ± 0.0190	0.8355	0.6320	0.8653 ± 0.0128	0.9596	0.6206
RMSE	2.8950 ± 1.0871	2.8123	3.1370	3.2438 ± 1.6723	2.0725	5.077

(a)

(b)

Figure 6. 3. Parity plots of experimental and predicted values based on training data a) ED; b) FPF

6.5.4 The sensitivity analysis of input variables on ED and FPF

Sensitivity analysis is carried out to investigate the output variability against changes in the input variables. It can provide important information about

deterministic variables for output prediction. One thousand Monte Carlo samples are generated for each input variable by considering normally distributed samples having the mean value of the input variable and a standard deviation of 0.2 times of the mean value. The result of sensitivity analysis is presented in Figure 6.4. The columns represent the mean and error bars represent the standard deviation of the output variables based on one thousand Monte Carlo inputs. Key input variables can be considered as those that lead to long error bars. It can be seen in Figure 6.4 that the carrier particle average size (I1), carrier particle size standard deviation (I2), flow rate (I14), and carrier-to-drug ratio (I15) are significant for both ED and FPF. The results also demonstrate that ED and FPF are not sensitive to tap density (I11). Although there are reports declared that the presence of fine carrier particles can improve ED and FPF for carriers such as lactose, glucose, sorbitol, and mannitol [356, 367, 375, 393-395], sensitivity analysis shows that fine particles (I7 and I8) are not effective on ED and FPF for HA. Carrier surface roughness is an important characteristic affecting the performance of DPI. Among the carrier surface roughness variables, Rq (I16), Ra (I17), MFOV (I25) and SA (I28) are found to be the most critical surface variables in relation to ED (Figure 6.4a), while Rq (I16), Ra (I17), FPO (I24) and SA (I28) are the most sensitive surface roughness variables for FPF (Figure 6.4b).

Figure 6. 4. Sensitivity analysis of input variables in prediction of (a) ED and (b) FPF

6.5.5 Effects and interactions of various factors on ED and FPF

Based on the findings of sensitivity analysis, three-dimensional surface plots were developed based on the ANN models to depict the relationship among the most important factors on ED and FPF, and the results are shown in Figures 6.5–6.7.

6.5.5.1 Effect of particle average size and size distribution on ED and FPF

As shown in Figure 6.5a, both particle size standard deviation and particle average size have a linear relationship with ED. An increase in particle size standard deviation and particle average size can increase ED. On the other hand, Figure 5b indicates that only particle average size has a linear relationship with FPF even though increase of both particle size standard deviation and particle average size can increase FPF. However, the effect of particle size standard deviation on FPF becomes negligible when the standard deviation is higher than about 5 μ m. The effect of particle average size on FPF is also diminished as particle size standard deviation is increased.

An increase in particle average size was reported to increase the accessible surface area and the number of active sites for drug attachment. Moreover, an increase in particle average size also improves the particle flowability and aerosolization properties of particles. As a result, increasing the carrier particle size can lead to an increase in both ED and FPF as reported in a number of literature studies [371, 396, 397]. The HA particles size standard deviation is a representation of the homogeneity of the size distribution of the samples. An increase in the particle size standard deviation increases the spread of particle size distribution. It implies that the amount as well as the size of particles larger than the average particle size is also increased with an increase in the particle size standard deviation. It is to note that while an increase in the particle size standard deviation increases the presence of large carrier particles, it is also accompanied by an increase in the presence of small carrier particles for the same particle average size. The increase in both ED and FPF with an increase in the particle size indicates that the presence of large particles is the dominating factor to improve both ED and FPF. It is also observed that the effect of particle size standard deviation on FPF becomes negligible at high particle size standard deviation values. Since the effect of particle size standard deviation is still observable on ED, it is possible that when the amount of small carrier particles reached a certain level, the aerosolization properties are decreased while the flowability and surface area/active sites are not affected.

Figure 6. 5. Three-dimensional surfaces of ANN models for ED and FPF as a function of the HA particle average size and size standard deviation; a) ED; b) FPF

6.5.5.2 Effect of flow rate and carrier-to-drug ratio on ED and FPF

It can be seen in Figure 6.6 that both flow rate and carrier-to-drug ratio have positive and relatively linear relationships with ED and FPF. The effects of flow rate and carrier-to-drug ratio are particularly significant on FPF compared to ED. It is well accepted that an increase in flow rate increases both the turbulence and shear stress. It allows the pickup of drugs that are not attached on carriers and lead to an increase in ED. In addition, the high turbulence and shear stress promote the aerosolization of the drug particles and facilitates the drug detachment after aerosolization. As a result, high flow rate is more impactful to FPF than ED. An increase in the carrier-to-drug ratio increases the number of active sites for drug attachment and reduces drug agglomeration. Therefore, the flowability of the formulation as well as the ED and FPF are improved [398]. It is also to note from Figure 6.6b that high flow rates enhance the effect of high carrier-to-drug ratios on FPF and similarly high carrier-to-drug ratios simulate the effect of high flow rates on FPF. It is believed that high flow rates help to

overcome the potential drug liberation problem at high carrier-to-drug ratios to provide a synergistic effect on FPF.

Figure 6. 6. Three-dimensional surfaces of ANN models for ED and FPF as a function of the flow rate and carrier-to-drug ratio; a) ED; b) FPF

6.5.5.3 Effect of surface morphology on ED and FPF

Based on the sensitivity analysis presented in Section 6.5.4, Ra, Rq, SA and MFOV are the most important carrier surface roughness variables for ED while Ra, Rq, SA and FPO are the most sensitive surface roughness variables for FPF. Ra is the arithmetic mean of the vertical heights of the surface while Rq is the root mean square of the vertical heights. Both Ra and Rq increase with an increase in surface roughness. It is also to note that because of the difference in definition, a single large peak or trough can affect Rq more than Ra. In addition, both Ra and Rq cannot differentiate peaks from troughs on a smooth surface. SA is the surface area determined based on the gradient analysis among the vertical heights [388]. MFOV is variation of the polar facet orientation, which describes the variations of the orientation of the surface. FPO is mean polar facet orientation and it represents the average polar orientation of the surface.

Therefore, MFOV and FPO can used describe the orientation of the surface features and potentially the distance between surface features. Three-dimensional surface plots of the effects of Ra, Rq, SA and MFOV on ED and the effects of Ra, Rq, SA and FPO on FPF are shown in Figure 6.7.

Figure 6. 7. Three-dimensional surfaces of ANN models for ED and FPF as a function of the HA surface roughness variables; a) ED as a function of Ra and Rq; b) ED as a function of SA and MFOV; c) FPF as a function of Ra and Rq; c) FPF as a function of SA and FPO

Figures 6.7a and 6.7b show that an increase in Ra, Rq, SA and MFOV all would lead to an increase in ED even though the degree of improvement is diminishing at high Ra, Rq, SA and MFOV values. It can be concluded ED is the most

concerning to the attachment of drugs, which link directly to the surface area of carrier particles. Surface area is known to increase with an increase in Ra, Rq, SA and MFOV. It was reported that an increase in the active surface area on carrier particles can increase ED [381, 399-402]. A comparison between Figure 6.7a and Figure 6.7b indicates that the effect of SA and MFOV on ED is more significant than that of Ra and Rq on ED. The results show that orientation of the surface features and active surface area have an important role in drug attachment compared to just the peaks and troughs of the surface.

Different from the requirement of surface roughness for high ED, which relies primarily on high surface roughness for increased attachment of drug particles on carrier particles, high FPF demands an appropriate degree of adhesion between the drug and carrier particles. It can be seen in Figures 7c and 7d that there are optimal Ra, Rq, SA and FPO values exist for FPF. This finding can help to support the literature findings that FPF were observed to increase with an increase in carrier surface roughness [400, 403, 404] as well as FPF can decrease with an increase in carrier surface roughness [31, 381]. As demonstrated in the effects of Ra, Rq, and SA on ED, an increase in their values increases the active surface area and lead to an increase in drug attachment. It is believed that the increase in FPF at low Ra, Rq, and SA values is a result of the increase in ED. When the drug attachment becomes too high, the drug particles can be deposited pre-maturely in the initial airways while remained attached to the large carrier particles and therefore lowers FPF. A recent study analyzed the contact modes between different carrier surfaces and drug and found that the optimal surface roughness should take drug particle size into consideration [380]. Since FPO can be view as a measure of the orientation of the surface features, it should be linked

indirectly the distance between peaks or troughs on the carrier surface. Table 6.4 shows the cropped carrier surface images obtained from the literature scaled to the same magnification and arranged in increasing order of FPO values. It can be seen that the distance between the surface features decreases with an increase in FPO. An attempt is thus made to correlate the FPO values with a quantitative measure of distance between peaks on different carrier surfaces. Arrows shown in each of the image presented in Table 6.4 represents generally the distance between surface features and a plot of the distance against FPO is shown in Figure 6.8. A positive relationship can be observed from the figure, indicating that FPO can be a good representation of the feature distance. Based on Figure 6.7d, a maximum FPF is observed at FPO between 60°–70°, which corresponds to about 1.6 μ m for distance between surface features on Figure 6.8. Since the drug particles used in all these studies are the same and have an average size of 2.5 $\pm$ 1.1 μ m, the observed optimal distance between surface features corresponds well with the ideal drug-carrier adhesion that provides the highest FPF [380]. Therefore, for optimal carrier performance, the FPO of the carrier particle is proposed to have a value that corresponds to a peak distance that is 60-70% smaller than the mean drug particle size.

Table 6.4. Analysis of roughness parameters based on cropped carrier surface images

Cropped carrier surface image	Roughness Parameters	References
$Ra = 32.4962 \pm 3.1596$ $Rq = 40.1153 \pm 4.0613$ $SA = 5.9302 \pm 0.6911$ $FPO = 63.8071 \pm 1.6038$	HA-1 in [383]
$Ra = 18.2921 \pm 0.7486$ $Rq = 23.0766 \pm 0.7141$ $SA = 4.1700 \pm 0.1945$ $FPO = 64.3406 \pm 1.0777$	HA-3 in [371]
$Ra = 27.6103 \pm 1.6566$ $Rq = 35.3137 \pm 2.1092$ $SA = 7.1895 \pm 1.4682$ $FPO = 73.3797 \pm 4.7532$	HA-2 in [383]
$Ra = 37.1002 \pm 1.7807$ $Rq = 47.1868 \pm 1.7231$ $SA = 9.1026 \pm 0.2035$ $FPO = 75.5205 \pm 0.5583$	HA-50 in [384]
$Ra = 48.3568 \pm 5.1986$ $Rq = 59.1084 \pm 5.8301$ $SA = 9.6587 \pm 0.6613$ $FPO = 76.9019 \pm 0.8934$	HA-40 in [384]
$Ra = 38.7639 \pm 3.4593$ $Rq = 48.8284 \pm 4.8804$ $SA = 12.1537 \pm 1.5129$ $FPO = 79.7067 \pm 1.0341$	HA-original in [384]

Figure 6. 8. Relationship between FPO and peak distance.

6.6 Summary

In this chapter, artificial neural network (ANN) was applied to study the performance of carrier-based dry powder inhalation based on experimental data were obtained from the literature. Carrier surface characteristics were evaluated by image analysis of carrier SEM images. Twenty eight variables describing the carrier size distribution, density, surface characteristics and operating conditions of dry powder inhaler were used as the input variables and emitted dose (ED) and fine particle fraction (FPF) were used as the output variables for the construction of ANN models. Before surrogate modeling, 13 input variables were selected for the construction of ANN models for ED and FPF by Variable selection using generic algorithm (GA). Sensitivity analysis by Monte Carlo simulation was also performed to determine the key variables affecting ED and FPF. Among the variable groups that describe the carrier particle size distribution, carrier density and operating conditions, carrier particle average size, carrier particle size

standard deviation, flow rate and carrier-to-drug ratio were found to be significant for both ED and FPF. Carrier surface roughness is an important characteristic affecting the performance of DPI. For carrier surface characteristic, Rq, Ra, SA and MFOV are found to be the most critical surface variables in relation to ED, while Rq, Ra, SA and FPO are the most sensitive surface roughness variables for FPF.

Three-dimensional surface plots were generated to study the relationship of the important input variables with ED and FPF. It was found that an increase in the carrier particle average size and size distribution can enhance both ED and FPF and this increase is primarily dominated by the presence of large particles. However, the effect of particle size standard deviation on FPF becomes negligible at high particle size standard deviation values. Both flow rate and carrier-to-drug ratio were found to have positive relationship with ED and FPF though the influences on FPF are more significant than that on ED. The effects of Ra, Rq, SA and MFOV on ED are mostly concerning the attachment of drugs, which link directly to the surface area of carrier particles. SA and MFOV have more significant influence on ED than Ra and Rq, indicating that the orientation of the surface features and active surface area than just the peaks and troughs of the surface. High FPF relies on a sufficiently high drug attachment to maximize ED but low enough to allow drug liberation and minimize pre-mature deposition. Therefore, optimal Ra, Rq, SA and FPO values exist for FPF. In addition, FPO was found to be closely related to distance between peaks and can be used in conjunction with drug particle size for the determination of optimal distance between surface features.

Chapter 7 Modeling of emitted dose and fine particle fraction in dry powder inhalation for carrier-based pulmonary delivery formulations by using neural networks and genetic algorithms

7.1 Preface

In the previous chapter is tried to unfold the relationship between HA carrier properties and DPI efficiency [405] by surrogate modeling. However, since DPI efficiency depends on carrier materials and contradictory results have been reported for same carrier in different operating conditions such as carrier size, carrier surface roughness, flow rate, and carrier-to-drug ratio, development of a comprehensive model that includes all the conditions, is necessary. In this work a hybrid ANN – GA is used for the modeling of the ED and FPF in a DPI process for three carriers, lactose (LA), hydroxyapatite (HA), and mannitol (MN). Due to the highly complex and nonlinear behavior of ED and FPF, the use of surrogate models is preferred over physical models. ANN is a powerful technique used to create surrogate models and it has been successfully applied to various complex systems such as [2, 6, 13, 27, 56, 119, 146, 235, 316-322, 406]. GA is applied for variable selection, symbolic regression and ANN hyper-parameters optimization. Sensitivity analysis is also performed to determine the importance of various input variables on ED and FPF. Design strategies to enhance the effectiveness of carrier based DPI are also proposed.

7.1 Dataset

All the carrier formulations and aerosolization properties used in this study were obtained from published literature that involves LA, HA, and MN as carrier materials [359, 362, 371, 374, 383, 384, 407-409].

Table 7.1. General description of created database

Variables	Information description	
Input variables		
I1	Carrier material	Lactose (LA)
I2		Hydroxyapatite (HA)
I3		Mannitol (MN)
I4	Carrier particle size distribution	Carrier particles average size
I5		Carrier particles size standard deviation
I6		Volume fraction of carrier particles having a particle size smaller than 5 μm
I7		Volume fraction of carrier particles having a particle size smaller than 10 μm
I8		Volume fraction of carrier particles having a particle size smaller than 15 μm
I9		Volume fraction of carrier particles having a particle size smaller than 20 μm
I10	Drug particles size	Drug particles average size
I11	Carrier density	Tap density
I12	Flow rate	
I13	Carrier- to- drug ratio	
I14	Carrier morphology and surface characteristics	Root mean square deviation (Rq)
I15		Arithmetical mean deviation (Ra)
I16		Skewness of the assessed profile (Rsk)
I17		Kurtosis of the assessed profile (Rku)
I18		Lowest valley (Rv)
I19		Highest peak (Rp)
I20		Total height of the profile (Rt)
I21		Average height of an unleveled surface (Rc)
I22		Mean polar facet orientation (FPO)
I23		Variation of the polar facet orientation (MFOV)
I24		Direction of azimuthal facets (FAD)
I25		Mean resultant vector (MRV)
I26		Surface area (SA)
Output variables		
O1	Fine particle fraction (FPF)	
O2	Emitted dose (ED)	

Variables such as ED, FPF, flow rate, carrier-to-drug ratio, tap density, drug particle size, carrier material, and carrier particle size are obtained directly from the references. All carrier particle size distributions are normalized to a Gaussian distribution by least-square method. Since some studies showed that the presence of fine carrier particles can improve DPI efficiency [354, 410], volume fractions of carrier particles less than 5, 10, 15, and 20 μm are also extracted based on the fitted Gaussian distribution. A list of all the input and output variables used in the

construction of ANN models is shown in Table 7.1. A total of 65 datasets obtained from 9 references, of which 50 datasets are used as training data while 15 datasets are used as validation data.

7.2 Surface and shape analysis

Surface properties of carrier particles are obtained through image analysis of SEM images using ImageJ® software (v2.33) [385, 386] with SurfCharJ (v.1q) plugin [349, 380]. For standardization, all SEM images of carrier particles were rescaled to resolution of 12.4 pixels per micron [387, 388]. A square area of 200 pixels x 200 pixels in the different regions of each carrier particle in the image was cropped for surface characterization. Finally, the average value of each variable has been selected for training data. Thirteen surface roughness size variables include root mean square deviation (Rq), arithmetical mean deviation (Ra), skewness of the assessed profile (Rsk), kurtosis of the assessed profile (Rku), lowest valley (Rv), highest peak (Rp), total height of the profile (Rt), average height of an unleveled surface (Rc), and surface roughness orientation variables are mean polar facet orientation (FPO), variation of the polar facet orientation (MFOV), direction of azimuthal facets (FAD), mean resultant vector (MRV), and surface area (SA) which describe the surface properties were determined.

7.3 Design of ANN

ANN surrogate models were built using 50 training samples and additional 15 samples were used as test sets. Leave-one-out cross-validation has been also used as accuracy evaluation for surrogate models. The testing samples were selected

randomly to cover all the range of input variables. In this study, a four-layer (namely, one input layer, two hidden layers, and one output layer) network was used. The optimum number of hidden layers was found by trial and error following the general concept that the number of hidden layers should be between number of input and output variables [240]. A publicly available Matlab SUMO toolbox (<http://www.sumowiki.intec.ugent.be>) [11, 323, 324] was employed. The toolbox was also used to normalize all the input and output data to scale between -1 and 1. A hybrid of Gauss-Newton nonlinear regression method and gradient steepest descent method (Levenberg-Marquardt algorithm) were used to build the feed forward ANN. The weights and biases were generated randomly in the range of -1 and 1. Tansig (tan-sigmoid) transfer function was used between the input and hidden layers. Purelin (linear) function was used as the transfer function between the hidden and output layers. The network weights and biases were obtained using GA to achieve the global minima of deviation error, and optimum weights and biases were obtained simultaneously.

7.4 Genetic algorithms parameters

The parameters and techniques used in GA for variable selection, ANN optimization, and SR are listed in Table 7.2. The parameters obtained by trial and error and some known useful defaults [338] were used as initial values. The weights, biases and GA convergence graphs have been shown in Figure A4 and Table A4 in Appendix.

Table 7.2. GA parameters for variable selection and ANN optimization

GA parameters	Variable selection	ANN optimization	SR
Population size	50	112	400
Maximum generation	100	140	200
Elite count	5	10	15
Crossover fraction	0.8	0.7	0.85
Selection technique	Stochastic uniform	Stochastic uniform	Stochastic uniform
Crossover method	Scattered	Scattered	Scattered
Mutation function	Gaussian	Gaussian	Gaussian
Stopping criteria	Maximum generation	Maximum generation	Maximum generation

7.5 Results and discussion

7.5.1 Analysis of surface roughness

A sample workflow on the analysis of surface properties of carrier particle based on SEM images is shown in Figure 7.1. It can be seen that certain variation is present for the same particle. Therefore, average values from a triplicate of cropped images for the same particle formulation was used in the construction of ANN models.

Figure 7. 1. A sample workflow of surface property analysis.

7.5.2 Selection of important variables

Figure 7.2 shows the variable selection results for ED and FPF. The description of the variable numbers shown in the horizontal axis can be found in Table 7.1. The vertical axis indicates the frequency of inclusion of each variable in a generation of GA. and the higher the frequency of inclusion, the more likely the variable is going to be useful in the surrogate models. In this study, variables that have more than 10 percent of inclusion were selected in the development of ANN models. It is shown in Figure 7.2a that carrier material, carrier particle average size, carrier particles size standard deviation, drug particle average size, tap density, flow rate, carrier-to-drug ratio are significant in the prediction of ED. It is also found that Rq, Ra, Rp, Rc, MFOV, MRV, and SA are the more descriptive variables to describe the impact of carrier surface characteristics on ED. For FPF, Figure 7.2b indicates that carrier material, carrier particle average size, carrier particles size standard deviation, drug particle average size, tap density, flow rate, carrier-to-drug ratio, Rq, Ra., Rsk, Rku, FPO, and SA have the most contribution in the model development. Paclawski et al. have confirmed the deterministic role of carrier material in DPI [52]. Carrier particles size parameters (carrier particle average size and standard deviation) have always been considered one of the most important variables affecting ED and FPF [349, 389-391].

Although there are many reports declared that the presence of fine carrier particles can improve ED and FPF for carriers such as lactose, glucose, sorbitol, and mannitol [356, 367, 375, 393-395], variable selection result shows that fine particles are not effective on ED and FPF which confirms by results of previous chapter.

The tap density has the most important role in particle flowability. As well as previous studies [359, 389, 392], the result of variables selection by GA shows that flow rate and carrier to drug ratio have the most effect on ED and FPF. Rq and Ra which are the root mean squared and the arithmetical mean of the absolute values of roughness, were used always as the most important variables of surface roughness [349], so they were selected for both ED and FPF. There is no specific information for other carrier surface variables in literature. Rp and Rc which show maximum peak height and mean height of roughness, respectively, were selected for ED, while Rsk and Rku that express the symmetry of peaks and valleys using the average line as the center and the sharpness of the height distribution, respectively, were choose for FPF. In terms of surface roughness orientation analysis, MFOV and MRV which express variation of the polar facet orientation and mean resultant vector are important for ED and FPO that is mean polar facet orientation was selected for FPF. Since active carrier surface has an essential role in DPI process, SA as surface area is significant variable for both ED and FPF.

Figure 7. 2. Frequency of variable usage in models; a) ED, b) FPF.

7.5.3 Design of the ANN model for prediction of FPF and ED

In the building of ANN models, 50 training points were used and 15 testing points were selected randomly for validation. A comparison of the model predicted ED and FPF and the respective training data is shown in Figure 7.3. Both ED and FPF were found to have good agreement between the model prediction and the training data.

Validation of the ANN models developed is also performed using 15 testing points and the R-square and RSME results for both training points and testing points are shown in Table 7.3. It can be seen that while the error incurred in the testing points are higher than that in the training points, an $R^2 > 0.6$ is considered reasonable [218].

Table 7.3. Validation results of surrogate models

Models		ED			FPF	
Error metrics	Cross validation	Training points	Testing Points	Cross validation	Training points	Testing Points
R square	0.7546 ± 0.0207	0.8717	0.6724	0.8674 ± 0.0130	0.9550	0.6928
RMSE	4.5601 ± 1.8347	3.8467	5.7076	3.1124 ± 1.9072	2.1714	5.1150

a) Predicted ED versus fitted ED

b) Predicted FPF versus fitted FPF

Figure 7. 3. Parity plots of experimental and predicted ED and FPF values calculated by the models for training data.

7.5.4 Sensitivity analysis of input variables

Sensitivity analysis is carried out to determine the key input variables that have the most impact on ED and FPF. For each carrier material (LA, HA, MN), one thousand Monte Carlo samples are generated based on inputs normally distributed around the mean of each input variable and a standard deviation 0.2 times of each input variable mean. The result of sensitivity analysis is presented in Figure 7.4. The bars and error bars present the mean and standard deviation values, respectively of ED and FPF for three thousand inputs in Figure 7.4. It can be seen in Figure 7.4 that carrier particles average size, carrier particles size standard deviation, drug particle average size, tap density, flow rate, and carrier-to-drug ratio are significant for both ED and FPF. Among the carrier surface roughness variables, Ra and SA are found to be the most critical in the prediction of ED (Figure 7.4a), while Rq, Ra, FPO, and SA are the most sensitive for FPF (Figure 7.4b).

Figure 7. 4. Sensitivity analysis of input variables in prediction of (a) ED and (b) FPF.

7.5.5 Effects of carrier materials and interactions of various factors on ED and FPF

7.5.5.1 Carrier materials

In order to understand the interactions of various factors on ED and FPF and their relationship with different carrier materials, it is important to identify the chemical structure and properties of the carrier materials. The chemical structure and properties of carrier materials are shown in Table 7.4.

Carrier structure has an important effect in DPI performance. The surface adhesion energy for amorphous part is higher than crystalline surfaces. The surface adhesion energy increase causes to decrease deaggregation of drug particles. Therefore, the presence of amorphous material can cause to reduce drug dispersion. It was reported that the crystallized form of mannitol shows better DPI performance than different polymorphs of mannitol [411]. There are two isometric forms, α and β , and one amorphous form for LA. LA is the most common and frequently used carrier in DPIs formulations. Because of clinical

considerations lactose or other sugars cannot be used for drug delivery to diabetic patients. MN, a hexahydric alcohol, can be used as an alternative carrier. Mannitol is less hygroscopic than lactose, so it has good stability and lowest particle cohesion than other sugars. HA is a calcium phosphate which can be used as non-sugar carrier in DPI formulation. HA has a crystallized form and are observed to exhibit two types of morphology, namely petallike and needlelike.

Table 7.4. Chemical structure and properties of carrier materials

Carrier materials	Lactose (LA)	Hydroxyapatite (HA)	Mannitol (MN)
Chemical formula	C ₁₂ H ₂₂ O ₁₁	Ca ₅ (PO ₄) ₃ (OH)	C ₆ H ₁₄ O ₆
Molecular weight (g/mol)	342.3	502.3	182.2
Crystallinity	α -Lactose monohydrate	Crystallized HA	Three crystallized form α , β , δ mannitol coexistence
Chemical Structure

According to the sensitivity analysis results, three-dimensional surface plots for ED and FPF versus the most important factors are shown in Figures 7.5 – 7.8 for HA, LA, and MN.

7.5.5.2 Effect of carrier particle average size and size distribution

As shown in Figure 7.5a, particle size standard deviation has negligible effect on ED, except for LA at large average size only. For LA at large average size, an increase in the particle size standard deviation is found to enhance ED.

(a)

(b)

Figure 7. 5. Three-dimensional surfaces of ANN models for ED and FPF as a function of the carrier particle average size and size standard deviation; a) ED; b) FPF.

The effect of average size on ED, on the other hand, has opposite effect on LA from HA and MN. It can be seen in Figure 5a that ED increases with an increase

in particle average size for HA and MN, while ED decreases with an increase in particle average size for LA. For FPF, Figure 5b indicates that particle size standard deviation only affects low average size for HA and high average size for LA only. It can also be seen that the effect of average size is opposite between LA and HA/MN. It gives a standardized indication of the literature reported data that indeed an increase in the average size of LA decreases the FPF while an increase in the average size of HA and MN increases the FPF when the effect of carrier particle size distribution, carrier particle density, and surface properties do not come into play.

7.5.5.2 Effect of carrier-to-drug ratio and drug particles average size

It can be seen in Figure 7.6a that both carrier-to-drug ratio and drug particles average size have positive and relatively linear relationships with ED. The carrier-to-drug ratio is particularly significant on ED compared to drug particles average size. Figure 7.6b shows while FPF increases with carrier-to-drug ratio increase, larger drug particles average size reduces FPF. The change of ED and FPF with carrier-to-drug ratio and drug particles average size for LA is more meaningful than other carriers. The number of active sites on carrier's surface grows with increase in the carrier-to-drug ratio. In this situation, there is more space for drug attachment and it prevents from build-up of the pure drug as aggregates which can reduce DPI efficiency. Therefore, ED and FPF are improved by carrier-to-drug ratio increase [357, 398].

(a)

(b)

Figure 7. 6. Three-dimensional surfaces of ANN models for ED and FPF as a function of the carrier-to-drug ratio and drug particles average size; a) ED; b) FPF.

Drug particles average size has an important effect on drug deposition. Although better flowability of larger drug particles can help to reach higher ED, but these particles are usually deposited in the oral cavities and pharynx. Hence, the drug particles cannot efficiently absorb in the target areas and FPF decreases [412-415]. It is also to note from Figure 7.6b that LA has the better performance in comparison with MN in the high carrier-to-drug ratio. The drug attachment is probably improved more in LA than MN with carrier-to-drug ratio increase, so more drugs can be carried by LA particle to the target areas in lung.

7.5.5.3 Effect of flow rate and carrier tap density

It can be seen in Figure 7.7 that while increasing the flow rate increases ED and FPF, but the tap density has little effect on ED and FPF. The effect of flow rate is particularly significant on FPF compared to ED. The turbulence and shear stress increase with flow rate increase, so even drug particles that are not attached on carriers can disperse and lead to an increase in ED. Moreover, the drug particles aerosolization improves in the high turbulence and shear stress and helps to drug detachment in the target area in lung. Accordingly, high flow rate is more effective on FPF than ED. Although carrier tap density has insignificant effect on ED and FPF, but it is more impactful to LA and MN than HA. On the other hand, it is believed that high flow rates help to overcome the potential flowability problem at high tap density [355]. Therefore, in Figure 7.7b FPF is almost same for all carriers in the high flow rates, but FPF reduces with tap density increase in the low flow rates.

(a)

(b)

Figure 7. 7. Three-dimensional surfaces of ANN models for ED and FPF as a function of the flow rate and tap density; a) ED; b) FPF.

7.5.5.4 Effect of carrier surfaces morphology

Based on the sensitivity analysis presented in section 7.5.4, Ra and SA are the most important carrier surface roughness variables for ED while Rq, Ra, SA and FPO are the most sensitive surface roughness variables for FPF. Rq and Ra which are arithmetical mean and root mean square deviation of the roughness, respectively, have a direct relationship with surface roughness height. Another necessary analysis on surface roughness is gradient analysis. An outward surface normal vector can be used in gradient analysis of each surface roughness. Then the direction of this vector is utilized to find the orientation of gradient which can be determined by polar and azimuthal angles in the polar coordinate system. The angle between the normal vector and zenith direction is polar angle. FPO is mean of surface roughness polar angles. Surface area, SA, which can be calculated by gradient analysis, is a practical variable in DPI. Previous studies on the surface roughness show that FPO and SA increase with surface roughness growth [388].

Figure 7.8a shows that ED increases with Ra and SA. A surface with high Ra has deep valleys and high peaks. The drug particles adhere and stay on the holes and peaks of surfaces, therefore ED increases. ED improvement with carrier surface roughness increase has also been reported in some previous studies [400, 403, 404].

Moreover, the drug-carrying capacity raises with SA increase. In this situation carrier tends to carry more amount of drug particles which can improve ED. In this regard, many studies show that high ED can be provided by carrier with larger active surface [381, 399-402]. On the other hand, Figure 7.8a illustrates that SA increase improves ED more than Ra increase for HA and MN. It is

believed that existence of active sites on the carrier surface is more impactful on ED than surface height of crests and crevices. Furthermore, high ED is provided when Ra and SA increase together.

Figure 7.8b demonstrates that FPF increases with Rq and Ra increase to achieve a maximum. After that FPF reduces with Rq and Ra increase. Although it can be seen this behavior for HA in all Rq and Ra values, but for LA and MN, FPF always increase with Rq in small Ra and vice versa. As it was mentioned, both Rq and Ra are mean functions of surface peaks and valleys, but their values are different. For example, although Ra shows a small value for a surface with just one high peak, but Rq can be larger than Ra for the same surface. Therefore, the existence of a surface with high Rq and low Ra is possible. Generally, due to weak adhesion force between drug particles and carrier surface at low Rq and Ra, dispersion is high in the first steps, so ED is low. Therefore FPF decreases in this situation too. Although at high Rq and Ra, there is a strong adhesion between drug particles and carrier surface which can increase ED, but detachment of drug particles becomes difficult and as a result FPF decreases. Immobility of drug particles at high Rq and Ra has been reported by previous studies [400, 403, 404]. In this way, as it was discussed in previous chapter, the reason for contradictory results about FPF against Rq and Ra in the previous studies is existence of an optimum Rq and Ra which can provide maximum FPF.

FPF change versus FPO and SA is presented in Figure 7.8c. FPO increase at low SA increases FPF, but at high SA, FPF reduces with FPO increase. Increasing the facet orientation angle to the surface plane at low SA which active sites on carrier surface is limited, can be caused to reinforce adhesion between drug particles and carrier, so drug particles dispersion reduces and ED increases. As a

result, FPF improves too. But at high SA which provides more active sites on carrier surface, facet orientation angle of roughness to the surface plane increase can create strong force in adhesion of drug particles to carrier. So, although ED improves with FPO increase, but according to complicated drug particles detachment FPF reduces with FPO increase. In addition, Figure 7.8c indicates that first FPF increases slightly with SA, then after a maximum, FPF decreases with SA growth. As it was mentioned, SA has a relationship with amount of active sites on carrier surface, so SA growth shows active sites increase. Accordingly, carrier can carry more drug particles and ED increases (Figure 7.8a) and in this way FPF also increases. After that, although SA growth is accompanied with active site increase on the carrier surface, but since carrier size is constant, SA growth is possible just by surface roughness increase. In this situation, drug particles are entrapped in carrier surface roughness, so ED increases. But drug particles entrapment makes difficult drug detachment in the target area of lung, so FPF reduces.

In overall, all Figures 7.5, 7.6, 7.7, and 7.8 demonstrate that carrier material plays an important role in DPI efficiency. As it can be seen in these figures, HA is a better choice in comparison with LA and MN in the same conditions. Therefore carrier material type should be considered in ED and FPF computation.

(a)

(b)

(c)

Figure 7. 8. Three-dimensional surfaces of ANN models for ED and FPF as a function of the carrier surface roughness variables; a) ED as a function of Ra and SA; b) FPF as a function of Rq and Ra; c) FPF as a function of FPO and SA.

In general, according to findings from three-dimensional surface plots, the optimum design parameters for carrier particles that provide the highest FPF was found to be HA carriers having average particle size = 231 μm , particle size standard deviation = 10.2 μm , carrier-to-drug ratio = 67.5, Ra = 27.7, Rq = 33.2, FPO = 62.8, and SA = 5.1. The maximum FPF achievable under such condition is 47.0 while the maximum reported FPF in the literature was 42.3 and of course, the most reported FPF was less than 30.

7.5.6 Symbolic regression (SR)

ANN allows the development of accurate models. However, the equations

developed by ANN are complex and difficult to use. Therefore, symbolic regression (SR), which can find a simple equation for complex problems such as DPI, has been used. GA as a powerful SR method is utilized for finding these equations for ED and FPF. The important variables for ED and FPF that were determined by sensitivity analysis of developed ANN model as well as carrier material type are used for building equation by SR. The following equations have been developed by GA for ED and FPF, respectively. The input variables numbers are defined as shown in Table 7.1.

$$\begin{aligned}
 ED = & 3.515 \times 10^{-5} \times (I_{11} + I_{12}) \times (I_5 + I_{11} + I_{12} + \frac{I_3}{I_{10}}) \times (I_5 + I_{13} + 9.121 \times I_2 \times I_{26}) \\
 & - 5.905 \times (I_2 + \ln(I_{12})) + \frac{\ln(I_{12})}{I_{15}} + \frac{I_{15}}{I_4} \\
 & - 11.81 \times (I_1 + I_{11}) + 118
 \end{aligned}
 \tag{7.1}$$

$$\begin{aligned}
 FPF = & 0.1135 \times I_{12} \\
 & + 1.6 \times 10^{-6} \times I_{12} \times I_{13} \times \exp(I_2) \times \ln(I_5) \times ((\frac{I_{10} \times I_{11}}{I_{14}}) + 2 \times I_3 + I_{15} - I_{22} + I_{26})^2 \\
 & + 0.05677 \times (\frac{I_3 \times I_{15} \times I_{26}^2 \times (I_4 - 18.33)}{I_{22}} - \exp(I_1)) + 6.722
 \end{aligned}
 \tag{7.2}$$

RMSE and R^2 are 5.9336 and 0.6372, respectively for ED and 5.8171 and 0.6535, respectively for FPF. A parity plot for comparison of real and calculated ED and FPF based on the following equations have been shown in Figure 7.9. This figure indicates that the equations can calculate the trend of the data pretty well.

Therefore, these equations can be used for rough computation of ED and FPF in DPI system.

Figure 7. 9. Parity plots of real and calculated ED and FPF by the developed equations by GA SR, a) calculated ED versus experimental ED b) calculated FPF versus experimental FPF

7.6 Summary

The performance of different carrier materials on DPI was investigated in this chapter. Different carriers have been proposed which more of them are sugar such as Lactose and mannitol and recently non-sugar carriers such as HA were suggested as the better choice. Drug delivery efficiency of a DPI usually is measured by ED and FPF. Hence, in this chapter, after variable selection, ANN with GA as a surrogate modeling tool was employed to model ED and FPF for three carriers, LA, HA, and MN on selected variables. The ANN-GA models developed were subsequently used to investigate sensitivity analysis to determine the most important variables in DPI formulation. Then, the effect of carrier properties, carrier surfaces morphology, flow rate, drug particles size and carrier-

to-drug ratio on ED and FPF was studied. Finally, SR was applied to find a simple equation to find ED and FPF for DPI formulation for all carrier materials.

Sensitivity analysis shows that carrier particles average size, carrier particles size standard deviation, drug particle average size, tap density, flow rate, and carrier-to-drug ratio are significant for both ED and FPF. Among the carrier surface roughness variables, Ra and SA are found to be the most critical in the prediction of ED, while Rq, Ra, FPO, and SA are the most sensitive for FPF. The effect of average size is opposite between LA and HA/MN. ED increases with an increase in particle average size for HA and MN, while ED decreases with an increase in particle average size for LA. An increase in the average size of LA decreases the FPF while an increase in the average size of HA and MN increases the FPF. Particle size standard deviation only has effect at low average size for HA and at high average size for LA only. FPF increases with carrier-to-drug ratio increase; larger drug particles average size reduces FPF. The change of ED and FPF with carrier-to-drug ratio and drug particles average size for LA is more significant than other carriers. ED and FPF are improved by carrier-to-drug ratio increase. About carrier surface roughness parameters, ED increases with Ra and SA. FPF increases with Rq and Ra increase to achieve a maximum. After that FPF reduces with Rq and Ra increase. FPO increase at low SA increases FPF, but at high SA, FPF reduces with FPO increase. In general, HA is a better carrier for drug delivery in comparison with LA and MN in the same conditions. Finally, it was found that the highest FPF can be provided by HA carriers having average particle size = 231 μm , particle size standard deviation = 10.2 μm , carrier-to-drug ratio = 67.5, Ra = 27.7, Rq = 33.2, FPO = 62.8, and SA = 5.1. The maximum FPF achievable under optimum condition is 47.0.

Chapter 8 Conclusion and outlook

8.1 Conclusions

This work aimed to apply surrogate modeling to predict particle size distribution (PSD) in gas-solid fluidized beds and drug delivery efficiency in carrier-based dry powder inhalation (DPI) system.

Firstly, an artificial neural network (ANN) combined with a genetic algorithm (GA) was applied to predict PSD in a gas-solid fluidized bed. Ball milling was proposed to derive descriptive parameters to account for the effect of material properties in the particle attrition process. It found that the derived descriptive parameters from ball milling have a relationship with material properties such as hardness. The results showed that the Rosin-Rammler (RR) distribution is able to describe the PSD reasonably well. Two ANN-GA models were developed based on the RR parameters (d and n) which can estimate PSD during fluidization reasonably well. The developed surrogate model were also applied to study effect of fluidization time, mass percentage and size of glass beads used as foreign particles, and gas velocity on the particle attrition during fluidization. The particle attrition increases with an increase in the mass percentage of glass beads having similar size range as the particles in the co-fluidization process. There is an optimum filling ratio of the space between the glass beads and particles. When the filling ratio increases beyond the optimum, collisions between glass beads and particles reduces and particle attrition decreases with an increase in the mass percentage of glass beads. Uneven fluidization can occur when the glass beads used are larger than the particles. An increase in the mass percentage of large glass beads increase the collision between particles, so particle attrition reduces.

The proposed material attrition properties, B_d and B_n , reduce with an increase in material hardness. Particle attrition reduces with B_d and B_n decrease. In addition, it is found that there is an optimum gas velocity exists, which can provide the maximum particle attrition in gas-solid fluidized bed.

Similar to gas-solid fluidized bed, carrier-based DPI is a process between gas and solid phases. Due to a large number of particle-particle interactions, simulation of the DPI process is computationally expensive. Therefore, ANN-GA surrogate models were constructed to predict emitted dose (ED) and fine particle fraction (FPF) as drug delivery efficiency in DPI with reasonable accuracy. The sensitivity analysis based on developed surrogate models demonstrated that the carrier particle average size, carrier particle size standard deviation, average drug particle size, tap density, flow rate, and carrier-to-drug ratio were significant for both ED and FPF. Among the carrier surface roughness variables, arithmetical mean deviation (R_a) and surface area (SA) were found to be the most critical surface variables in the prediction of ED, while root mean square deviation (R_q), R_a , mean polar facet orientation (FPO), and SA were the most sensitive surface roughness variables for FPF. The relationship of the important input variables with ED and FPF were studied with three-dimensional surface. An increase in the carrier particle average size and size distribution can enhance both ED and FPF for hydroxyapatite (HA), and mannitol (MN). However, at high particle size standard deviation values for HA and MN, the effect of particle size standard deviation on FPF becomes negligible. Nevertheless, ED and FPF decrease with an increase in average particle size for lactose (LA). Both flow rate and carrier-to-drug ratio have a positive relationship with ED and FPF. However, the effect of flow rate and carrier-to-drug ratio on FPF are more significant than that on

ED. The tap density of carriers has negligible effect on ED and FPF. Drug particles average size has a positive and relatively linear relationship with ED. But, FPF reduces with drug particles average size increase. The effects of Ra and SA in ED are mainly related to drug attachment that they depend on the surface area of carrier particles. It is found that SA has more significant effect on ED than Ra. It shows that the orientation of the surface features and active surface area are more important than just the peaks and troughs of the surface. Moreover, it was found that FPO is related to distance between peaks. Hence, FPO can be used to link optimal distance between surface features with drug particle size. The highest FPF reported in the literature was 42.3 while a majority other studies have FPF less than 30. It demonstrates the potential of the developed surrogate model can be used as a tool to provide better formulation for DPI.

On the whole, surrogate modeling is believed to have the potential to use as a simple, cheap, and accurate approach to modeling various complex chemical and biomedical processes. The developed surrogate model provides a cheap-to-compute tool for designers for design, optimization, control, and sensitivity analysis of processes in less cycle time.

8.2 Outlooks

In this work, surrogate modeling was applied in the modeling of complex chemical (PSD during fluidization in gas-solid fluidized bed) and biomedical (efficiency of carrier-based DPI system) processes. To expand the developed surrogate models applications, future work is needed to complement this research topic. Recommendations for future work are as follows:

Gas-solid fluidized bed study

- In the case of gas-solid fluidized bed study, all experiments have been done in a fluidized bed. Therefore, the fluidized bed geometry effect on particle attrition was not considered in the experiments. Hence, experiments can be conducted in different fluidized beds. Then, the results can be applied to the improvement of the developed surrogate models.
- The method used to determine the material attrition property using a planetary ball milling machine is versatile and can be used for other materials. Hence, to improve the developed surrogate model accuracy, other materials can be tested, and their results can be applied as surrogate model training and testing data.
- This study showed that the developed surrogate model can predict PSD during fluidization with high accuracy. Combining this surrogate model with CFD codes can be an interesting combination to obtain reliable and accurate results for fluidization process simulation.
- Incineration is an important municipal solid waste management process in Singapore. The IBA generated may contain heavy metals. Certain heavy metals were found to present in high concentrations at particular particle size ranges. The developed ANN models can be used to reach these particular particle size ranges by optimization of glass beads size, percentage and fluidization time.
- Basically, the fluidization experiments are difficult, expensive and time-consuming. Hence, this study proposes planetary ball milling with the

help of the developed ANN surrogate models to predict PSD during fluidization as cheap and fast alternative for fluidization experiments.

Dry powder inhalation study

- There are more carriers which are used in DPI. These carriers' data also can be added as training data in surrogate modeling of DPI systems to construct more accurate surrogate model.
- There are different methods in production of carrier for DPI such as Spray drying and freeze-drying. The developed surrogate model can be used to compare the carriers produced by these methods in term of their performance in DPI to find the best method.
- There are more accurate methods of surface roughness measurement than SEM imaging, such as contact profilometer, optical microscopy, and confocal microscopy. However, each requires more information than SEM images, which was not provided in this study. Therefore, a new experiment can be conducted on carrier surface roughness management with these methods to use the results as training data.
- As this study proves, the surrogate modeling as an accurate, fast and cheap-to-compute method can be applied for other complex chemical and biomedical processes.

List of publication

Journal

Amir Abbas Kazemzadeh Farizhandi, Han Zhao, Raymond Lau. Modeling the change in particle size distribution in a gas-solid fluidized bed due to particle attrition using a hybrid artificial neural network-genetic algorithm approach. Chemical Engineering Science, volume 155, 22 November 2016, Pages 210–220.

Amir Abbas Kazemzadeh Farizhandi, Han Zhao, Theodore Chen, and Raymond Lau. Using ball milling results for particles attrition to predict particle attrition behavior in the gas solid fluidized bed by response surface methodology. Submitted to Chemical Engineering Science journal

Amir Abbas Kazemzadeh Farizhandi, Ngoc Nguyen, Adam Paćławski, Jakub Szlęk, Aleksander Mendyk, Yu-Hsuan Shao, Raymond Lau. Modeling of emitted dose and fine particle fraction in dry powder inhalation for carrier-based pulmonary delivery formulations by using neural networks and genetic algorithms. Submitted to International Journal of Engineering Science.

Amir Abbas Kazemzadeh Farizhandi, Ngoc Nguyen, Adam Paćławski, Jakub Szlęk, Aleksander Mendyk, Yu-Hsuan Shao, Raymond Lau. Evaluation of hydroxyapatite size and morphology in dry powder inhalation for carrier-based pulmonary delivery formulations by response surface methodology. Submitted to International Journal of Engineering Science.

Conference

Amir Abbas Kazemzadeh Farizhandi, Fengying Zhang and Raymond Lau. « Study of DPI Carriers Design Parameters for Dry Powder Inhalation through

Artificial Neural Network Modeling », School of Chemical and Biomedical Engineering, Nanyang Technological University, Singapore, Singapore, 2016 AICHE Annual Meeting, November 13-18, 2016, Hilton San Francisco Union Square, San Francisco, CA, Parc 55 San Francisco, Hotel Nikko San Francisco.

Amir Abbas Kazemzadeh Farizhandi and Raymond Lau. « Prediction of particle size distribution in gas-solid fluidized bed by using a hybrid artificial neural network–genetic algorithm approach », School of Chemical and Biomedical Engineering, Nanyang Technological University, Singapore, The 2016 AOTULE annual meeting and conference, NTU, Singapore.

Han Zhao, Amir Abbas Kazemzadeh Farizhandi and Raymond Lau. « Attrition of Incineration Bottom Ash Particles in a Fluidized Bed », School of Chemical and Biomedical Engineering, Nanyang Technological University, Singapore, Singapore, 2015 AICHE Annual Meeting, November 8-13, 2015, Salt Palace Convention Center, Salt Lake Marriott Downtown at City Creek, Hilton Salt Lake City Center.

References

1. Wang, G.G. and S. Shan, Review of Metamodeling Techniques in Support of Engineering Design Optimization. *Journal of Mechanical Design*, 2006. 129(4): p. 370-380.
2. Li, Y.F., et al., A systematic comparison of metamodeling techniques for simulation optimization in Decision Support Systems. *Applied Soft Computing*, 2010. 10(4): p. 1257-1273.
3. Zhao, D. and D. Xue, A comparative study of metamodeling methods considering sample quality merits. *Structural and Multidisciplinary Optimization*, 2010. 42(6): p. 923-938.
4. Jin, R., W. Chen, and T.W. Simpson, Comparative studies of metamodeling techniques under multiple modelling criteria. *Structural and multidisciplinary optimization*, 2001. 23(1): p. 1-13.
5. Kleijnen, J.P.C., Kriging metamodeling in simulation: A review. *European Journal of Operational Research*, 2009. 192(3): p. 707-716.
6. Chi, G., et al., Response surface methodology with prediction uncertainty: A multi-objective optimisation approach. *Chemical Engineering Research and Design*, 2012. 90(9): p. 1235-1244.
7. Abu-Mostafa, Y.S., M. Magdon-Ismael, and H.-T. Lin, *Learning from data*. 2012: AMLBook Berlin, Germany.
8. Tang, B., *Latin Hypercube Designs*, in *Encyclopedia of Statistics in Quality and Reliability*. 2008, John Wiley & Sons, Ltd.
9. Creswell, J.W., *Educational research: Planning, conducting, and evaluating quantitative*. New Jersey: Upper Saddle River, 2002.
10. Boukouvala, F. and M.G. Ierapetritou, Feasibility analysis of black-box processes using an adaptive sampling Kriging-based method. *Computers & Chemical Engineering*, 2012. 36(0): p. 358-368.

11. Behrisch, M., et al. Sumo-simulation of urban mobility-an overview. in SIMUL 2011, The Third International Conference on Advances in System Simulation. 2011.
12. Berry, M.J. and G. Linoff, Data mining techniques: for marketing, sales, and customer support. 1997: John Wiley & Sons, Inc.
13. Hadjmohammadi, M.R. and K. Kamel, Response surface methodology and support vector machine for the optimization of separation in linear gradient elution. *Journal of Separation Science*, 2008. 31(22): p. 3864-3870.
14. Wang, G.G. and S. Shan, Review of metamodeling techniques in support of engineering design optimization. *Journal of Mechanical design*, 2007. 129(4): p. 370-380.
15. Weihong, Z. and X. Shunqing, Optimization of BP neural network classifier using genetic algorithm, in *Intelligence Computation and Evolutionary Computation*. 2013, Springer. p. 599-605.
16. Gay, D.H. and W.H. Ray, Identification and control of distributed parameter systems by means of the singular value decomposition. *Chemical Engineering Science*, 1995. 50(10): p. 1519-1539.
17. Hoo, K.A. and D. Zheng, Low-order control-relevant models for a class of distributed parameter systems. *Chemical Engineering Science*, 2001. 56(23): p. 6683-6710.
18. Shvartsman, S.Y., et al., Order reduction for nonlinear dynamic models of distributed reacting systems. *Journal of Process Control*, 2000. 10(2): p. 177-184.
19. Romijn, R., et al., A grey-box modeling approach for the reduction of nonlinear systems. *Journal of Process Control*, 2008. 18(9): p. 906-914.
20. Kennedy, M.C. and A. O'Hagan, Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2001. 63(3): p. 425-464.

21. Sacks, J., et al., Design and Analysis of Computer Experiments. *Statist. Sci.*, 1989. 4(4): p. 409-423.
22. Dutournié, P., P. Salagnac, and P. Glouannec, Optimization of Radiant-Convective Drying of a Porous Medium by Design of Experiment Methodology. *Drying Technology*, 2006. 24(8): p. 953-963.
23. Caballero, J.A. and I.E. Grossmann, Rigorous flowsheet optimization using process simulators and surrogate models. *Computer Aided Chemical Engineering*, 2008. 25: p. 551-556.
24. Caballero, J.A. and I.E. Grossmann, An algorithm for the use of surrogate models in modular flowsheet optimization. *AIChE Journal*, 2008. 54(10): p. 2633-2650.
25. Palmer, K. and M. Realff, Metamodeling Approach to Optimization of Steady-State Flowsheet Simulations. *Chemical Engineering Research and Design*, 2002. 80(7): p. 760-772.
26. Chen, T., J. Morris, and E. Martin, Gaussian process regression for multivariate spectroscopic calibration. *Chemometrics and Intelligent Laboratory Systems*, 2007. 87(1): p. 59-71.
27. Yan, W., et al., Development of high performance catalysts for CO oxidation using data-based modeling. *Catalysis Today*, 2011. 174(1): p. 127-134.
28. Gomes, M.V.C., et al., Using kriging models for real-time process optimisation. *Computer Aided Chemical Engineering*, 2008. 25: p. 361-366.
29. Gu, L. A comparison of polynomial based regression models in vehicle safety analysis. in *ASME Design Engineering Technical Conferences*, ASME Paper No.: DETC/DAC-21083. 2001.
30. Simpson, T.W., et al., Metamodels for computer-based engineering design: survey and recommendations. *Engineering with computers*, 2001. 17(2): p. 129-150.

31. Zeng, X.M., K.H. Pandhal, and G.P. Martin, The influence of lactose carrier on the content homogeneity and dispersibility of beclomethasone dipropionate from dry powder aerosols. *International Journal of Pharmaceutics*, 2000. 197(1–2): p. 41-52.
32. Haftka, R.T., E.P. Scott, and J.R. Cruz, Optimization and experiments: a survey. *Applied Mechanics Reviews*, 1998. 51(7): p. 435-448.
33. Simpson, T.W., et al., Approximation methods in multidisciplinary analysis and optimization: a panel discussion. *Structural and multidisciplinary optimization*, 2004. 27(5): p. 302-313.
34. Koch, P.N., et al., Statistical approximations for multidisciplinary design optimization: the problem of size. *Journal of Aircraft*, 1999. 36(1): p. 275-286.
35. Barthelemy, J.-F. and R.T. Haftka, Approximation concepts for optimum structural design—a review. *Structural optimization*, 1993. 5(3): p. 129-144.
36. Simpson, T.W., et al., On the use of statistics in design and the implications for deterministic computer experiments. *Design Theory and Methodology-DTM'97*, 1997: p. 14-17.
37. Barthelemy, J.-F. and R.T. Haftka, Approximation concepts for optimum structural design—a review. *Structural and Multidisciplinary Optimization*, 1993. 5(3): p. 129-144.
38. Sobieszczanski-Sobieski, J. and R.T. Haftka, Multidisciplinary aerospace design optimization: survey of recent developments. *Structural optimization*, 1997. 14(1): p. 1-23.
39. Ullman, D.G., Toward the ideal mechanical engineering design support system. *Research in Engineering Design*, 2002. 13(2): p. 55-64.
40. Kahng, A.B., B. Lin, and S. Nath. High-dimensional metamodeling for prediction of clock tree synthesis outcomes. in 2013 ACM/IEEE

International Workshop on System Level Interconnect Prediction (SLIP). 2013.

41. Kumar, B. and A.R. Rao, Metamodeling approach to predict friction factor of alluvial channel. *Computers and Electronics in Agriculture*, 2010. 70(1): p. 144-150.
42. Aflalo, B. and D. Ferrari, METAMODELING TECHNIQUES APPLIED TO AIRCRAFT NOISE PREDICTION. *Blucher Mechanical Engineering Proceedings*, 2012. 1(1): p. 5018-5033.
43. Geyer, P. and A. Schlüter, Automated metamodel generation for Design Space Exploration and decision-making – A novel method supporting performance-oriented building design and retrofitting. *Applied Energy*, 2014. 119: p. 537-556.
44. Hongwei, W., et al. An accurate ACOSSO metamodeling technique for processor architecture design space exploration. in *The 20th Asia and South Pacific Design Automation Conference*. 2015.
45. Montgomery, D.C., *Design and Analysis of Experiments*, John Wiley & Sons. New York, 2001: p. 64-65.
46. Wong, P.C. and R.D. Bergeron. 30 years of multidimensional multivariate visualization. in *Scientific Visualization*. 1994.
47. Keim, D.A. and H.-P. Kriegel, Visualization techniques for mining large databases: A comparison. *IEEE Transactions on knowledge and data engineering*, 1996. 8(6): p. 923-938.
48. Winer, E. and C. Bloebaum, Development of visual design steering as an aid in large-scale multidisciplinary design optimization. Part II: method validation. *Structural and Multidisciplinary Optimization*, 2002. 23(6): p. 425-435.
49. Winer, E. and C. Bloebaum, Development of visual design steering as an aid in large-scale multidisciplinary design optimization. Part I: method

- development. *Structural and Multidisciplinary Optimization*, 2002. 23(6): p. 412-424.
50. Eddy, J. and K. Lewis. Visualization of multi-dimensional design and optimization data using cloud visualization. in *ASME Design Technical Conferences, Design Automation Conference*. 2002.
 51. Kodiyalam, S., R. Yang, and L. Gu, High-performance computing and surrogate modeling for rapid visualization with multidisciplinary optimization. *AIAA journal*, 2004. 42(11).
 52. Paćławski, A., et al., Empirical modeling of the fine particle fraction for carrier-based pulmonary delivery formulations. *International Journal of Nanomedicine*, 2015. 10: p. 801-810.
 53. Inamdar, S., S.K. Gupta, and D. Saraf, Multi-objective optimization of an industrial crude distillation unit using the elitist non-dominated sorting genetic algorithm. *Chemical Engineering Research and Design*, 2004. 82(5): p. 611-623.
 54. López, D.-C., et al., Optimization model of a system of crude oil distillation units with heat integration and metamodeling. *CT&F-Ciencia, Tecnología y Futuro*, 2009. 3(5): p. 159-173.
 55. López C, D.C., et al., Optimization Model of Crude Oil Distillation Units for Optimal Crude Oil Blending and Operating Conditions. *Industrial & Engineering Chemistry Research*, 2013. 52(36): p. 12993-13005.
 56. Palmer, K. and M. Realff, Metamodeling Approach to Optimization of Steady-State Flowsheet Simulations: Model Generation. *Chemical Engineering Research and Design*, 2002. 80(7): p. 760-772.
 57. Myers, R.H., D.C. Montgomery, and C.M. Anderson-Cook, *Response surface methodology: process and product optimization using designed experiments*. 2016: John Wiley & Sons.
 58. Chen, W., *A robust concept exploration method for configuring complex systems*. 1995.

59. Mitchell, T.J., An algorithm for the construction of “D-optimal” experimental designs. *Technometrics*, 1974. 16(2): p. 203-210.
60. Giunta, A.A., et al., Multidisciplinary optimization of a supersonic transport using design of experiments theory and response surface modeling. 1997.
61. Sacks, J., et al., Design and analysis of computer experiments. *Statistical science*, 1989: p. 409-423.
62. Koehler, J. and A. Owen, 9 Computer experiments. *Handbook of statistics*, 1996. 13: p. 261-308.
63. Currin, C., et al., Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments. *Journal of the American Statistical Association*, 1991. 86(416): p. 953-963.
64. Johnson, M.E., L.M. Moore, and D. Ylvisaker, Minimax and maximin distance designs. *Journal of statistical planning and inference*, 1990. 26(2): p. 131-148.
65. McKay, M.D., R.J. Beckman, and W.J. Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 2000. 42(1): p. 55-61.
66. Iman, R.L. and W. Conover, Small sample sensitivity analysis techniques for computer models. with an application to risk assessment. *Communications in statistics-theory and methods*, 1980. 9(17): p. 1749-1842.
67. Tang, B., Orthogonal array-based Latin hypercubes. *Journal of the American statistical association*, 1993. 88(424): p. 1392-1397.
68. Park, J.-S., Optimal Latin-hypercube designs for computer experiments. *Journal of statistical planning and inference*, 1994. 39(1): p. 95-111.
69. Kenny, Q.Y., W. Li, and A. Sudjianto, Algorithmic construction of optimal symmetric Latin hypercube designs. *Journal of statistical planning and inference*, 2000. 90(1): p. 145-159.

70. Taguchi, G., Taguchi methods: design of experiments. American Supplier Institute. Inc., MI, 1993.
71. Owen, A.B., Orthogonal arrays for computer experiments, integration and visualization. *Statistica Sinica*, 1992: p. 439-452.
72. Hedayat, A.S., N.J.A. Sloane, and J. Stufken, Orthogonal arrays: theory and applications. 2012: Springer Science & Business Media.
73. Kalagnanam, J.R. and U.M. Diwekar, An efficient sampling technique for off-line quality control. *Technometrics*, 1997. 39(3): p. 308-319.
74. Meckesheimer, M., et al., Computationally inexpensive metamodel assessment strategies. *AIAA journal*, 2002. 40(10): p. 2053-2060.
75. Fang, K.-T., et al., Uniform design: theory and application. *Technometrics*, 2000. 42(3): p. 237-248.
76. Chen, V.C., et al., A review on design, modeling and applications of computer experiments. *IIE transactions*, 2006. 38(4): p. 273-291.
77. Simpson, T.W., D.K. Lin, and W. Chen, Sampling strategies for computer experiments: design and analysis. *International Journal of Reliability and Applications*, 2001. 2(3): p. 209-240.
78. Au, S. and J.L. Beck, A new adaptive importance sampling scheme for reliability calculations. *Structural safety*, 1999. 21(2): p. 135-158.
79. Zou, T., et al. An indicator response surface-based Monte Carlo method for efficient component and system reliability analysis. in *ASME 2003 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. 2003. American Society of Mechanical Engineers.
80. Zou, T., et al., Reliability analysis of automotive body-door subsystem. *Reliability Engineering & System Safety*, 2002. 78(3): p. 315-324.

81. Kloess, A., Z. Mourelatos, and P. Meernik, Probabilistic analysis of an automotive body-door system. *International journal of vehicle design*, 2004. 34(2): p. 101-125.
82. Walker, J. Practical application of variance reduction techniques in probabilistic assessments. in the *Second International Conference on Radioactive Waste Management*. 1986.
83. Ditlevsen, O., R. Olesen, and G. Mohr, Solution of a class of load combination problems by directional simulation. *Structural Safety*, 1986. 4(2): p. 95-109.
84. Nie, J. and B.R. Ellingwood, Finite element-based structural reliability assessment using efficient directional simulation. *Journal of engineering mechanics*, 2005. 131(3): p. 259-267.
85. Wang, L., S. Shan, and G.G. Wang, Mode-pursuing sampling method for global optimization on expensive black-box functions. *Engineering Optimization*, 2004. 36(4): p. 419-438.
86. Shan, S. and G.G. Wang, An efficient Pareto set identification approach for multiobjective optimization on black-box functions. *Journal of Mechanical Design*, 2005. 127(5): p. 866-874.
87. Wang, G.G., L. Wang, and S. Shan, Reliability assessment using discriminative sampling and metamodeling. 2005, SAE Technical Paper.
88. Fu, J.C. and L. Wang, A random-discretization based Monte Carlo sampling method and its applications. *Methodology and Computing in Applied Probability*, 2002. 4(1): p. 5-25.
89. Lin, Y., An efficient robust concept exploration method and sequential exploratory experimental design. 2004.
90. Jin, R., W. Chen, and A. Sudjianto, An efficient algorithm for constructing optimal design of computer experiments. *Journal of Statistical Planning and Inference*, 2005. 134(1): p. 268-287.

91. Sasena, M., et al. Adaptive Experimental Design Applied to an Ergonomics Testing Procedure DAC-34091. in Proc. of the 2002 ASME DETC Conferences. 2002.
92. Wang, G.G., Adaptive response surface method using inherited latin hypercube design points. *Journal of Mechanical Design*, 2003. 125(2): p. 210-220.
93. Wang, G.G. and T. Simpson, Fuzzy clustering based hierarchical metamodeling for design space reduction and optimization. *Engineering Optimization*, 2004. 36(3): p. 313-335.
94. Jin, R., W. Chen, and A. Sudjianto. On sequential sampling for global metamodeling in engineering design. in ASME 2002 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. 2002. American Society of Mechanical Engineers.
95. Kajero, O.T., et al., Meta-modelling in chemical process system engineering. *Journal of the Taiwan Institute of Chemical Engineers*, 2017. 73: p. 135-145.
96. Sacks, J., S.B. Schiller, and W.J. Welch, Designs for computer experiments. *Technometrics*, 1989. 31(1): p. 41-47.
97. Cressie, N., Spatial prediction and ordinary kriging. *Mathematical geology*, 1988. 20(4): p. 405-421.
98. Papadrakakis, M., N.D. Lagaros, and Y. Tsompanakis, Structural optimization using evolution strategies and neural networks. *Computer methods in applied mechanics and engineering*, 1998. 156(1): p. 309-333.
99. Dyn, N., D. Levin, and S. Rippa, Numerical procedures for surface fitting of scattered data by radial functions. *SIAM Journal on Scientific and Statistical Computing*, 1986. 7(2): p. 639-659.

100. Fang, H. and M.F. Horstemeyer, Global response approximation with radial basis functions. *Engineering Optimization*, 2006. 38(04): p. 407-424.
101. Friedman, J.H., Multivariate adaptive regression splines. *The annals of statistics*, 1991: p. 1-67.
102. De Boor, C. and A. Ron, On multivariate polynomial interpolation. *Constructive Approximation*, 1990. 6(3): p. 287-302.
103. Langley, P. and H.A. Simon, Applications of machine learning and rule induction. *Communications of the ACM*, 1995. 38(11): p. 54-64.
104. Varadarajan, S., W. CHEN*, and C.J. Pelka, Robust concept exploration of propulsion systems with enhanced model approximation capabilities. *Engineering Optimization+ A35*, 2000. 32(3): p. 309-334.
105. Mullur, A.A. and A. Messac, Extended radial basis functions: more flexible and effective metamodeling. *AIAA journal*, 2005. 43(6): p. 1306-1315.
106. Turner, C.J. and R.H. Crawford. Selecting an appropriate metamodel: The case for NURBs metamodels. in *ASME 2005 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. 2005. American Society of Mechanical Engineers.
107. Barton, R.R. Metamodels for simulation input-output relations. in *Proceedings of the 24th conference on Winter simulation*. 1992. ACM.
108. Wilson, B., et al., Efficient Pareto frontier exploration using surrogate approximations. *Optimization and Engineering*, 2001. 2(1): p. 31-50.
109. Giunta, A.A. and L.T. Watson, A comparison of approximation modeling techniques: polynomial versus interpolating models. 1998.
110. Simpson, T., et al., Kriging Meta-Models for Global Approximation in Simulation-Based Design. *Eng. Comput.[CrossRef]*, 2001. 17(2): p. 129-150.

111. Wang, L., et al. Gaussian process meta-models for efficient probabilistic design in complex engineering design spaces. in ASME 2005 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. 2005. American Society of Mechanical Engineers.
112. Qian, Z., et al., Building surrogate models based on detailed and approximate simulations. *Journal of Mechanical Design*, 2006. 128(4): p. 668-677.
113. Martin, J.D. and T.W. Simpson, Use of kriging models to approximate deterministic computer models. *AIAA journal*, 2005. 43(4): p. 853-863.
114. Li, R. and A. Sudjianto. Penalized likelihood kriging model for analysis of computer experiments. in ASME 2003 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. 2003. American Society of Mechanical Engineers.
115. Van Beers, W.C. and J.P. Kleijnen, Kriging for interpolation in random simulation. 2001: Tilburg University.
116. Daberkow, D.D. and D.N. Mavris. An investigation of metamodeling techniques for complex systems design. in Proceedings 9th AIAA/USFA/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization. 2002.
117. Lophaven, S.N., H.B. Nielsen, and J. Søndergaard, DACE-A Matlab Kriging toolbox, version 2.0. 2002.
118. Clarke, S.M., J.H. Griebisch, and T.W. Simpson, Analysis of support vector regression for approximation of complex engineering analyses. *Journal of mechanical design*, 2005. 127(6): p. 1077-1087.
119. Yan, W., et al., Model-aided optimization and analysis of multi-component catalysts: Application to selective hydrogenation of cinnamaldehyde. *Chemical Engineering Science*, 2012. 76(0): p. 26-36.

120. Morris, M.D., T.J. Mitchell, and D. Ylvisaker, Bayesian design and analysis of computer experiments: use of derivatives in surface prediction. *Technometrics*, 1993. 35(3): p. 243-255.
121. Koehler, J., Estimating the response, derivatives, and transmitted variance using computer experiments. *Computing science and statistics*, 1998: p. 25-32.
122. Toropov, V., A. Filatov, and A. Polynkin, Multiparameter structural optimization using FEM and multipoint explicit approximations. *Structural optimization*, 1993. 6(1): p. 7-14.
123. Wang, L., R.V. Grandhi, and R.A. Canfield, Multivariate Hermite approximation for design optimization. *International Journal for Numerical Methods in Engineering*, 1996. 39(5): p. 787-803.
124. Rasmussen, J., Nonlinear programming by cumulative approximation refinement. *Structural optimization*, 1998. 15(1): p. 1-7.
125. Shin, Y. and R. Grandhi, A global structural optimization technique using an interval method. *Structural and Multidisciplinary Optimization*, 2001. 22(5): p. 351-363.
126. Huber, K.-P., M.R. Berthold, and H. Szczerbicka, Analysis of simulation models with fuzzy graph based metamodeling. *Performance Evaluation*, 1996. 27: p. 473-490.
127. Madu, C., A fuzzy theoretic approach to simulation metamodeling. *Applied mathematics letters*, 1995. 8(6): p. 35-41.
128. Kleijnen, J.P., Design and analysis of Monte Carlo experiments. 2004.
129. Giunta, A.A., et al. Noisy aerodynamic response and smooth approximations in HSCT design. in *Proc. 5-th AIAA/USAF/NASA/ISSMO Symp. on Multidisciplinary and Structural Optimization*. 1994.
130. Madsen, J.I., W. Shyy, and R.T. Haftka, Response surface techniques for diffuser shape optimization. *AIAA journal*, 2000. 38(9): p. 1512-1518.

131. Jin, R., X. Du, and W. Chen, The use of metamodeling techniques for optimization under uncertainty. *Structural and Multidisciplinary Optimization*, 2003. 25(2): p. 99-116.
132. Van Beers, W.C. and J.P. Kleijnen. Kriging interpolation in simulation: a survey. in *Simulation Conference*, 2004. *Proceedings of the 2004 Winter*. 2004. IEEE.
133. Villa-Vialaneix, N., et al., A comparison of eight metamodeling techniques for the simulation of N₂O fluxes and N leaching from corn crops. *Environmental Modelling & Software*, 2012. 34: p. 51-66.
134. El Tabach, E., et al., Use of artificial neural network simulation metamodeling to assess groundwater contamination in a road project. *Mathematical and Computer Modelling*, 2007. 45(7): p. 766-776.
135. Hurriion, R.D. and S. Birgil, A comparison of factorial and random experimental design methods for the development of regression and neural network simulation metamodels. *Journal of the Operational Research Society*, 1999. 50(10): p. 1018-1033.
136. Bourquin, J., et al., Advantages of Artificial Neural Networks (ANNs) as alternative modelling technique for data sets showing non-linear relationships using data from a galenical study on a solid dosage form. *European Journal of Pharmaceutical Sciences*, 1998. 7(1): p. 5-16.
137. Oberkampf, W.L. and T.G. Trucano, Validation methodology in computational fluid dynamics. *AIAA paper*, 2000. 2549: p. 19-22.
138. Trucano, T. and D. Post, Verification and validation in computational science and engineering. *Computing in Science and Engineering*, 2004. 6(5): p. 8-9.
139. Meckesheimer, M., A framework for metamodel-based design: subsystem metamodel assessment and implementation issues. 2001, The Pennsylvania State University.

140. Curry, B. and P. Morgan, Neural networks: a need for caution. *Omega*, 1997. 25(1): p. 123-133.
141. Gupta, J.N. and R.S. Sexton, Comparing backpropagation with a genetic algorithm for neural network training. *Omega*, 1999. 27(6): p. 679-684.
142. Porto, V.W., D.B. Fogel, and L.J. Fogel, Alternative neural network training methods. *IEEE Expert: Intelligent Systems and Their Applications*, 1995. 10(3): p. 16-22.
143. Fogel, D.B., *Evolutionary computation: toward a new philosophy of machine intelligence*. Vol. 1. 2006: John Wiley & Sons.
144. Dorsey, R.E. and W.J. Mayer, Genetic algorithms for estimation problems with multiple optima, nondifferentiability, and other irregular features. *Journal of Business & Economic Statistics*, 1995. 13(1): p. 53-66.
145. Mitchell, T.J. and M.D. Morris, Bayesian design and analysis of computer experiments: two examples. *Statistica Sinica*, 1992: p. 359-379.
146. Palmer, K. and M. Realff, Optimization and Validation of Steady-State Flowsheet Simulation Metamodels. *Chemical Engineering Research and Design*, 2002. 80(7): p. 773-782.
147. Henao, C.A. and C.T. Maravelias, Surrogate-based superstructure optimization framework. *AIChE Journal*, 2011. 57(5): p. 1216-1232.
148. Wright, M.M., Y. Román-Leshkov, and W.H. Green, Investigating the techno-economic trade-offs of hydrogen source using a response surface model of drop-in biofuel production via bio-oil upgrading. *Biofuels, Bioproducts and Biorefining*, 2012. 6(5): p. 503-520.
149. Fahmi, I. and S. Cremaschi, Process synthesis of biodiesel production plant using artificial neural networks as the surrogate models. *Computers & Chemical Engineering*, 2012. 46: p. 105-123.
150. Ochoa-Estopier, L.M., M. Jobson, and R. Smith, The use of reduced models for design and optimisation of heat-integrated crude oil distillation systems. *Energy*, 2014. 75: p. 5-13.

151. Chu, J.-Z., et al., Constrained optimization of combustion in a simulated coal-fired boiler using artificial neural network model and information analysis☆. *Fuel*, 2003. 82(6): p. 693-703.
152. Kusiak, A. and Z. Song, Combustion efficiency optimization and virtual testing: A data-mining approach. *IEEE Transactions on Industrial Informatics*, 2006. 2(3): p. 176-184.
153. Zheng, L., et al., Combining support vector regression and ant colony optimization to reduce NO_x emissions in coal-fired utility boilers. *Energy & Fuels*, 2008. 22(2): p. 1034-1040.
154. Li, S., et al., Using surrogate models for efficient optimization of simulated moving bed chromatography. *Computers & Chemical Engineering*, 2014. 67: p. 121-132.
155. Beck, J., et al., Multi-objective optimisation using surrogate models for the design of VPSA systems. *Computers & Chemical Engineering*, 2015. 82: p. 318-329.
156. Gutiérrez-Antonio, C. and A. Briones-Ramírez, Multiobjective Stochastic Optimization of Dividing-wall Distillation Columns Using a Surrogate Model Based on Neural Networks. *Chemical and biochemical engineering quarterly*, 2016. 29(4): p. 491-504.
157. Nuchitprasittichai, A. and S. Cremaschi, An algorithm to determine sample sizes for optimization with artificial neural networks. *AIChE Journal*, 2013. 59(3): p. 805-812.
158. Wiltowski, T., et al., Neural network approximation of iron oxide reduction process. *Chemical Engineering and Processing: Process Intensification*, 2005. 44(7): p. 775-783.
159. Chang, J.-S., Y.-P. Lee, and R.-C. Wang, Optimization of nanosized silver particle synthesis via experimental design. *Industrial & engineering chemistry research*, 2007. 46(17): p. 5591-5599.

160. Chang, J.-S., J.-T. Lee, and A.-C. Chang, Neural-network rate-function modeling of submerged cultivation of *Monascus anka*. *Biochemical Engineering Journal*, 2006. 32(2): p. 119-126.
161. Chen, J., et al., Optimal design of filament winding using neural network experimental design scheme. *Journal of Composite Materials*, 1999. 33(24): p. 2281-2300.
162. Chen, J., et al., Optimal design using neural network and information analysis in plasma etching. *Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena*, 1999. 17(1): p. 145-153.
163. Liao, L.-K., et al., Process modeling and optimization of PECVD silicon nitride coated on silicon solar cell using neural networks. *Solar energy materials and solar cells*, 2002. 71(2): p. 169-179.
164. Chuang, Y.-C. and C.-T. Chen, Mathematical modeling and optimal design of an MOCVD reactor for GaAs film growth. *Journal of the Taiwan Institute of Chemical Engineers*, 2014. 45(1): p. 254-267.
165. Zhang, G.P., Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 2003. 50: p. 159-175.
166. Chen, S., Nonlinear time series modelling and prediction using Gaussian RBF networks with enhanced clustering and RLS learning. *Electronics letters*, 1995. 31(2): p. 117-118.
167. Sapankevych, N.I. and R. Sankar, Time series prediction using support vector machines: a survey. *IEEE Computational Intelligence Magazine*, 2009. 4(2).
168. Wang, J., A. Hertzmann, and D.M. Blei. Gaussian process dynamical models. in *Advances in neural information processing systems*. 2006.
169. Joseph, B. and F.W. Hanratty, Predictive control of quality in a batch manufacturing process using artificial neural network models. *Industrial & engineering chemistry research*, 1993. 32(9): p. 1951-1961.

170. Desai, K., et al., Soft-sensor development for fed-batch bioreactors using support vector regression. *Biochemical Engineering Journal*, 2006. 27(3): p. 225-239.
171. Gonzaga, J., et al., ANN-based soft-sensor for real-time process monitoring and control of an industrial polymerization process. *Computers & Chemical Engineering*, 2009. 33(1): p. 43-49.
172. Liu, G., et al., Model optimization of SVM for a fermentation soft sensor. *Expert Systems with Applications*, 2010. 37(4): p. 2708-2713.
173. Ge, Z., T. Chen, and Z. Song, Quality prediction for polypropylene production process based on CLGPR model. *Control Engineering Practice*, 2011. 19(5): p. 423-432.
174. Yu, J., Online quality prediction of nonlinear and non-Gaussian chemical processes with shifting dynamics using finite mixture model based Gaussian process regression approach. *Chemical Engineering Science*, 2012. 82: p. 22-30.
175. Nahas, E., M. Henson, and D. Seborg, Nonlinear internal model control strategy for neural network models. *Computers & Chemical Engineering*, 1992. 16(12): p. 1039-1057.
176. Willis, M.J., et al., Artificial neural networks in process estimation and control. *Automatica*, 1992. 28(6): p. 1181-1187.
177. Kocijan, J., et al. Gaussian process model based predictive control. in *American Control Conference*, 2004. Proceedings of the 2004. 2004. IEEE.
178. Tsen, A.Y.D., et al., Predictive control of quality in batch polymerization using hybrid ANN models. *AIChE Journal*, 1996. 42(2): p. 455-465.
179. Ng, C. and M. Hussain, Hybrid neural network—prior knowledge model in temperature control of a semi-batch polymerization process. *Chemical Engineering and Processing: Process Intensification*, 2004. 43(4): p. 559-570.

180. Fullana, M., F. Trabelsi, and F. Recasens, Use of neural net computing for statistical and kinetic modelling and simulation of supercritical fluid extractors. *Chemical Engineering Science*, 2000. 55(1): p. 79-95.
181. Yu, C.J., et al., Core network identification using parametric sensitivity and multi-way principal component analysis in NFkB signaling network. *Journal of the Taiwan Institute of Chemical Engineers*, 2013. 44(5): p. 724-733.
182. Sobol, I.M., Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation*, 2001. 55(1): p. 271-280.
183. Cukier, R., et al., Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. *The Journal of chemical physics*, 1973. 59(8): p. 3873-3878.
184. McRae, G.J., J.W. Tilden, and J.H. Seinfeld, Global sensitivity analysis—a computational implementation of the Fourier amplitude sensitivity test (FAST). *Computers & Chemical Engineering*, 1982. 6(1): p. 15-25.
185. Sobol, I.M., Theorems and examples on high dimensional model representation. *Reliability Engineering & System Safety*, 2003. 79(2): p. 187-193.
186. Xiu, D. and G.E. Karniadakis, Modeling uncertainty in flow simulations via generalized polynomial chaos. *Journal of computational physics*, 2003. 187(1): p. 137-167.
187. Crestaux, T., O. Le Maître, and J.-M. Martinez, Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering & System Safety*, 2009. 94(7): p. 1161-1172.
188. Martin, J.D. and T.W. Simpson, Use of kriging models to approximate deterministic computer models. *AIAA J*, 2005. 43(4): p. 853-863.

189. Marrel, A., et al., Calculations of sobol indices for the gaussian process metamodel. *Reliability Engineering & System Safety*, 2009. 94(3): p. 742-751.
190. Saltelli, A., et al., Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, 2010. 181(2): p. 259-270.
191. Janon, A., et al., Asymptotic normality and efficiency of two sobol index estimators. *ESAIM: Probability and Statistics*, 2014. 18: p. 342-364.
192. Koda, M., G.J. Mcrae, and J.H. Seinfeld, Automatic sensitivity analysis of kinetic mechanisms. *International Journal of Chemical Kinetics*, 1979. 11(4): p. 427-444.
193. Saltelli, A., et al., Sensitivity analysis for chemical models. *Chemical reviews*, 2005. 105(7): p. 2811-2828.
194. Kiparissides, A., et al., Global sensitivity analysis challenges in biological systems modeling. *Industrial & Engineering Chemistry Research*, 2009. 48(15): p. 7168-7180.
195. Lucay, F., L. Cisternas, and E. Gálvez, Global sensitivity analysis for identifying critical process design decisions. *Chemical Engineering Research and Design*, 2015. 103: p. 74-83.
196. Carrero, E., et al., Global sensitivity analysis of Alkali–Surfactant–Polymer enhanced oil recovery processes. *Journal of Petroleum Science and Engineering*, 2007. 58(1): p. 30-42.
197. Wang, K., et al., Meta-modelling for fast analysis of CFD-simulated vapour cloud dispersion processes. *Computers & Chemical Engineering*, 2014. 69: p. 89-97.
198. Natrella, M., *NIST/SEMATECH e-handbook of statistical methods*. 2010.
199. Fonseca, D.J. and D. Navarrese, Artificial neural networks for job shop simulation. *Advanced Engineering Informatics*, 2002. 16(4): p. 241-246.

200. Rumelhart, D.E., G.E. Hinton, and R.J. Williams, Learning representations by back-propagating errors. *Nature*, 1986. 323(6088): p. 533-536.
201. Fogel, D.B., Theoretical and Empirical Properties of Evolutionary Computation, in *Evolutionary Computation*. 2005, John Wiley & Sons, Inc. p. 105-181.
202. Greenwood, G.W., Training partially recurrent neural networks using evolutionary strategies. *Speech and Audio Processing, IEEE Transactions on*, 1997. 5(2): p. 192-194.
203. Goldberg, D.E. and J.H. Holland, Genetic algorithms and machine learning. *Machine learning*, 1988. 3(2): p. 95-99.
204. Broadhurst, D., et al., Genetic algorithms as a method for variable selection in multiple linear regression and partial least squares regression, with applications to pyrolysis mass spectrometry. *Analytica Chimica Acta*, 1997. 348(1-3): p. 71-86.
205. Paterlini, S. and T. Minerva. Regression model selection using genetic algorithms. in *Proceedings of the 11th WSEAS International Conference on Neural Networks and Evolutionary Computing and Fuzzy Systems*. 2010.
206. Eigenvector, PLS Toolbox 4.0. 2006.
207. Massaad, A., et al., Repeatability and validation of Gait Deviation Index in children: Typically developing and cerebral palsy. *Gait & Posture*, 2014. 39(1): p. 354-358.
208. Harrison, R.L. Introduction to monte carlo simulation. in *AIP conference proceedings*. 2010. AIP.
209. McGrath, E.J., et al., Techniques for efficient Monte Carlo simulation. Volume 1. Selecting probability distributions. 1975, Oak Ridge National Lab., Tenn.(USA).

210. Allen, T.T., Probability Theory and Monte Carlo, in Introduction to Discrete Event Simulation and Agent-based Modeling. 2011, Springer. p. 9-27.
211. Stijven, S., et al., Prime-Time: Symbolic Regression Takes Its Place in the Real World, in Genetic Programming Theory and Practice XIII, R. Riolo, et al., Editors. 2016, Springer International Publishing: Cham. p. 241-260.
212. Affenzeller, M., et al., Gaining deeper insights in symbolic regression, in Genetic Programming Theory and Practice XI. 2014, Springer. p. 175-190.
213. Poli, R., W. Langdon, and N. McPhee, A field guide to genetic programming (With contributions by JR Koza)(2008). Published via <http://lulu.com>, 2008.
214. Koza, J.R. and R. Poli, Genetic programming, in Search Methodologies. 2005, Springer. p. 127-164.
215. Koza, J.R., et al., Genetic programming: Biologically inspired computation that creatively solves non-trivial problems, in Evolution as computation. 2002, Springer. p. 95-124.
216. Searson, D.P., GPTIPS 2: an open-source software platform for symbolic data mining, in Handbook of genetic programming applications. 2015, Springer. p. 551-573.
217. Sammut, C. and G.I. Webb, Leave-One-Out Cross-Validation, in Encyclopedia of Machine Learning, C. Sammut and G.I. Webb, Editors. 2010, Springer US: Boston, MA. p. 600-601.
218. Alexander, D.L.J., A. Tropsha, and D.A. Winkler, Beware of R²: Simple, Unambiguous Assessment of the Prediction Accuracy of QSAR and QSPR Models. *Journal of Chemical Information and Modeling*, 2015. 55(7): p. 1316-1322.

219. Hsieh, K.-L., Parameter optimization of a multi-response process for lead frame manufacturing by employing artificial neural networks. *The International Journal of Advanced Manufacturing Technology*, 2006. 28(5-6): p. 584-591.
220. Malakooti, B. and V. Raman, An interactive multi-objective artificial neural network approach for machine setup optimization. *Journal of Intelligent Manufacturing*, 2000. 11(1): p. 41-50.
221. Markopoulos, A.P., D.E. Manolakos, and N.M. Vaxevanidis, Artificial neural network models for the prediction of surface roughness in electrical discharge machining. *Journal of Intelligent Manufacturing*, 2008. 19(3): p. 283-292.
222. Hou, T.-H., et al., An integrated system for setting the optimal parameters in IC chip-package wire bonding processes. *The International Journal of Advanced Manufacturing Technology*, 2006. 30(3): p. 247-253.
223. Su, C.-T. and T.-L. Chiang, Optimizing the IC wire bonding process using a neural networks/genetic algorithms approach. *Journal of Intelligent Manufacturing*, 2003. 14(2): p. 229-238.
224. Tong, K., C. Kwong, and K. Yu, Intelligent process design system for the transfer moulding of electronic packages. *International Journal of Production Research*, 2004. 42(10): p. 1911-1931.
225. Chang, H.-H. and Y.-K. Chen, Neuro-genetic approach to optimize parameter design of dynamic multiresponse experiments. *Applied Soft Computing*, 2011. 11(1): p. 436-442.
226. Noorossana, R., S.D. Tajbakhsh, and A. Saghaei, An artificial neural network approach to multiple-response optimization. *The International Journal of Advanced Manufacturing Technology*, 2009. 40(11-12): p. 1227-1238.
227. Mok, S., C. Kwong, and W. Lau, An intelligent hybrid system for initial process parameter setting of injection moulding. *International Journal of Production Research*, 2000. 38(17): p. 4565-4576.

228. Lau, H.C., et al., Design and implementation of a process optimizer: a case study on monitoring molding operations. *Expert Systems*, 2005. 22(1): p. 12-21.
229. Wang, G., et al., Process optimization of the serial-parallel hybrid polishing machine tool based on artificial neural network and genetic algorithm. *Journal of Intelligent manufacturing*, 2012. 23(3): p. 365-374.
230. Huang, C.-C. and T.-T. Tang, Parameter optimization in melt spinning by neural networks and genetic algorithms. *The International Journal of Advanced Manufacturing Technology*, 2006. 27(11-12): p. 1113-1118.
231. Sedighi, M. and D. Afshari, Creep feed grinding optimization by an integrated GA-NN system. *Journal of Intelligent manufacturing*, 2010. 21(6): p. 657-663.
232. Holden, T. and M. Serearuno, A hybrid artificial intelligence approach for improving yield in precious stone manufacturing. *Journal of Intelligent manufacturing*, 2005. 16(1): p. 21-38.
233. Sibalija, T., Development of an intelligent designer of experiment model for application of Taguchi method. 2009, PhD thesis, Faculty of Mechanical Engineering, University of Belgrade, Belgrade (in Serbian).
234. Sibalija, T.V. and V.D. Majstorovic, An integrated approach to optimise parameter design of multi-response processes based on Taguchi method and artificial intelligence. *Journal of Intelligent manufacturing*, 2012. 23(5): p. 1511-1528.
235. Agatonovic-Kustrin, S., et al., Application of neural networks for response surface modeling in HPLC optimization. *Analytica Chimica Acta*, 1998. 364(1-3): p. 265-273.
236. Dorofki, M., et al., Comparison of artificial neural network transfer functions abilities to simulate extreme runoff data. *International Proceedings of Chemical, Biological and Environmental Engineering*, 2012. 33: p. 39-44.

237. Duch, W. and N. Jankowski, Survey of neural transfer functions. *Neural Computing Surveys*, 1999. 2(1): p. 163-212.
238. Duch, W. and N. Jankowski. Transfer functions: hidden possibilities for better neural networks. in *ESANN*. 2001.
239. Maca, P., P. Pech, and J. Pavlasek, Comparing the selected transfer functions and local optimization methods for neural network flood runoff forecast. *Mathematical Problems in Engineering*, 2014. 2014.
240. Heaton, J., *Introduction to Neural Networks for Java*, 2nd Edition. 2008: Heaton Research, Inc. 440.
241. Lawrence, S., C.L. Giles, and A.C. Tsoi, What size neural network gives optimal generalization? Convergence properties of backpropagation. 1998.
242. Lawrence, S., C.L. Giles, and A.C. Tsoi. Lessons in neural network training: Overfitting may be harder than expected. in *AAAI/IAAI*. 1997.
243. Šibalića, T.V. and V.D. Majstorović, *Advanced Multiresponse Process Optimisation*. 2016.
244. Alderliesten, M., Mean Particle Diameters. Part VII. The Rosin-Rammler Size Distribution: Physical and Mathematical Properties and Relationships to Moment-Ratio Defined Mean Particle Diameters. *Particle & Particle Systems Characterization*, 2013. 30(3): p. 244-257.
245. Macías-García, A., E.M. Cuerda-Correa, and M.A. Díaz-Díez, Application of the Rosin–Rammler and Gates–Gaudin–Schuhmann models to the particle size distribution analysis of agglomerated cork. *Materials Characterization*, 2004. 52(2): p. 159-164.
246. Taşdemir, A. and T. Taşdemir, A Comparative Study on PSD Models for Chromite Ores Comminuted by Different Devices. *Particle & Particle Systems Characterization*, 2009. 26(1-2): p. 69-79.

247. Brezáni, I. and F. Zelenák, Improving the effectivity of work with Rosin-Rammler diagram by using MATLAB R GUI tool. *Act Montanist Slovaca*, 2010. 15: p. 152-157.
248. Cornelissen, J.T., et al., CFD modelling of a liquid–solid fluidized bed. *Chemical Engineering Science*, 2007. 62(22): p. 6334-6348.
249. Grace, J.R. and G. Sun, Influence of particle size distribution on the performance of fluidized bed reactors. *The Canadian Journal of Chemical Engineering*, 1991. 69(5): p. 1126-1134.
250. Geldart, D., Types of gas fluidization. *Powder technology*, 1973. 7(5): p. 285-292.
251. Boerefijn, R., N.J. Gudde, and M. Ghadiri, A review of attrition of fluid cracking catalyst particles. *Advanced Powder Technology*, 2000. 11(2): p. 145-174.
252. Kunii, D. and O. Levenspiel, *Fluidization engineering*. 2013: Elsevier.
253. Scala, F., R. Chirone, and P. Salatino, *Attrition phenomena relevant to fluidized bed combustion and gasification systems*. 2013.
254. Werther, J. and E.-U. Hartge, Modeling of Industrial Fluidized-Bed Reactors. *Industrial & Engineering Chemistry Research*, 2004. 43(18): p. 5593-5604.
255. Klett, C., E.-U. Hartge, and J. Werther, Time-dependent behavior of a catalyst in a fluidized bed/cyclone circulation system. *AIChE Journal*, 2007. 53(4): p. 769-779.
256. Liu, X., et al., Zeolite Y synthesized with FCC spent catalyst fines: particle size effect on catalytic reactions. *Journal of Porous Materials*, 2012. 19(1): p. 133-139.
257. Rousset, P., et al., Change in particle size distribution of Torrefied biomass during cold fluidization. *Energy*, 2013. 51(0): p. 71-77.

258. Zacca, J.J., J.A. Debling, and W.H. Ray, Reactor residence time distribution effects on the multistage polymerization of olefins—I. Basic principles and illustrative examples, polypropylene. *Chemical Engineering Science*, 1996. 51(21): p. 4859-4886.
259. Yao, Q., et al., Mobility of heavy metals and rare earth elements in incineration bottom ash through particle size reduction. *Chemical Engineering Science*, 2014. 118: p. 214-220.
260. Farag, H., et al., Modeling of fluidized bed reactor for ethylene polymerization: effect of parameters on the single-pass ethylene conversion. *International Journal of Industrial Chemistry*, 2013. 4(1): p. 1-10.
261. Kaushal, P., J. Abedi, and N. Mahinpey, A comprehensive mathematical model for biomass gasification in a bubbling fluidized bed reactor. *Fuel*, 2010. 89(12): p. 3650-3661.
262. Deen, N.G., et al., Review of discrete particle modeling of fluidized beds. *Chemical Engineering Science*, 2007. 62(1–2): p. 28-44.
263. Hamzehei, M., CFD Modeling and Simulation of Hydrodynamics in a Fluidized Bed Dryer with Experimental Validation. *ISRN Mechanical Engineering*, 2011. 2011.
264. Patankar, S., Numerical heat transfer and fluid flow. 1980: CRC press.
265. Grace, J.R. and F. Taghipour, Verification and validation of CFD models and dynamic similarity for fluidized beds. *Powder technology*, 2004. 139(2): p. 99-110.
266. Ranade, V., Computational Flow Modeling for Chemical Dryer Engineering. 2002, Academic Press, New York, NY, USA.
267. Ahsan, M., Computational fluid dynamics (CFD) prediction of mass fraction profiles of gas oil and gasoline in fluid catalytic cracking (FCC) riser. *Ain Shams Engineering Journal*, 2012. 3(4): p. 403-409.

268. Chalermssinsuwan, B., P. Kuchonthara, and P. Piumsomboon, CFD modeling of tapered circulating fluidized bed reactor risers: Hydrodynamic descriptions and chemical reaction responses. *Chemical Engineering and Processing: Process Intensification*, 2010. 49(11): p. 1144-1160.
269. Chalermssinsuwan, B., D. Gidaspo, and P. Piumsomboon, Two- and three-dimensional CFD modeling of Geldart A particles in a thin bubbling fluidized bed: Comparison of turbulence and dispersion coefficients. *Chemical Engineering Journal*, 2011. 171(1): p. 301-313.
270. Chalermssinsuwan, B., et al., Computational fluid dynamics of circulating fluidized bed downer: Study of modeling parameters and system hydrodynamic characteristics. *Chemical Engineering Journal*, 2012. 189–190(0): p. 314-335.
271. Chen, T., et al., Efficient meta-modelling of complex process simulations with time–space-dependent outputs. *Computers & Chemical Engineering*, 2011. 35(3): p. 502-509.
272. Dehnavi, M.A., et al., CFD simulation of hydrodynamics and heat transfer in gas phase ethylene polymerization reactors. *International Communications in Heat and Mass Transfer*, 2010. 37(4): p. 437-442.
273. Gao, J., et al., CFD modeling and validation of the turbulent fluidized bed of FCC particles. *AIChE Journal*, 2009. 55(7): p. 1680-1694.
274. Gómez-Barea, A. and B. Leckner, Modeling of biomass gasification in fluidized bed. *Progress in Energy and Combustion Science*, 2010. 36(4): p. 444-509.
275. Jiang, L., X. Fang-Zhi, and L. Zheng-Hong, A CFD modeling of the gas–solid two-phase flow in an FCC riser under the electrostatic conditions. *Asia-Pacific Journal of Chemical Engineering*, 2014. 9(5): p. 645-655.
276. Rokkam, R.G., R.O. Fox, and M.E. Muhle, Computational fluid dynamics and electrostatic modeling of polymerization fluidized-bed reactors. *Powder Technology*, 2010. 203(2): p. 109-124.

277. Singh, R.I., A. Brink, and M. Hupa, CFD modeling to study fluidized bed combustion and gasification. *Applied Thermal Engineering*, 2013. 52(2): p. 585-614.
278. van der Hoef, M.A., et al., Numerical Simulation of Dense Gas-Solid Fluidized Beds: A Multiscale Modeling Strategy. *Annual Review of Fluid Mechanics*, 2008. 40(1): p. 47-70.
279. Xiao, Z., et al., CFD Modeling and Simulation of Superheated Steam Fluidized Bed Drying Process, in *Computer and Computing Technologies in Agriculture VI*, D. Li and Y. Chen, Editors. 2013, Springer Berlin Heidelberg. p. 141-149.
280. Mathiesen, V., T. Solberg, and B.H. Hjertager, Predictions of gas/particle flow with an Eulerian model including a realistic particle size distribution. *Powder Technology*, 2000. 112(1): p. 34-45.
281. Bemrose, C.R. and J. Bridgwater, A review of attrition and attrition test methods. *Powder Technology*, 1987. 49(2): p. 97-126.
282. Ammendola, P. and F. Scala, Attrition of lignite char during fluidized bed gasification. *Experimental Thermal and Fluid Science*, 2012. 43: p. 9-12.
283. Nawaz, Z., et al., Attrition behavior of fine particles in a fluidized bed with bimodal particles: Influence of particle density and size ratio. *Korean Journal of Chemical Engineering*, 2010. 27(5): p. 1606-1612.
284. Chen, Z., J.R. Grace, and C. Jim Lim, Limestone particle attrition and size distribution in a small circulating fluidized bed. *Fuel*, 2008. 87(7): p. 1360-1371.
285. Chirone, R., et al., Char attrition during the batch fluidized bed combustion of a coal. *AIChE Journal*, 1985. 31(5): p. 812-820.
286. Ghadiri, M. and Z. Zhang, Impact attrition of particulate solids. Part 1: A theoretical model of chipping. *Chemical Engineering Science*, 2002. 57(17): p. 3659-3669.

287. Johnsen, K. and J.R. Grace, High-temperature attrition of sorbents and a catalyst for sorption-enhanced steam methane reforming in a fluidized bed environment. *Powder Technology*, 2007. 173(3): p. 200-202.
288. Lee, J.M., J.S. Kim, and J.J. Kim, Comminution characteristics of Korean anthracite in a CFB reactor. *Fuel*, 2003. 82(11): p. 1349-1357.
289. Lin, C.-L. and M.-Y. Wey, Influence of hydrodynamic parameters on particle attrition during fluidization at high temperature. *Korean Journal of Chemical Engineering*, 2005. 22(1): p. 154-160.
290. Lupiáñez, C., et al., Primary fragmentation of limestone under oxy-firing conditions in a bubbling fluidized bed. *Fuel Processing Technology*, 2011. 92(8): p. 1449-1456.
291. Nawaz, Z., et al., Experimental investigation of attrition resistance of zeolite catalysts in two particle gas-solid-solid fluidization system. *Pakistan Journal of Scientific and Industrial Research*, 2010. 53(3): p. 169-174.
292. Rydén, M., et al., Measuring attrition resistance of oxygen carrier particles for chemical looping combustion with a customized jet cup. *Powder Technology*, 2014. 256: p. 75-86.
293. Subero-Couroyer, C., et al., Weibull Analysis of Quasi-Static Crushing Strength of Catalyst Particles. *Chemical Engineering Research and Design*, 2003. 81(8): p. 953-962.
294. Thon, A. and J. Werther, Attrition resistance of a VPO catalyst. *Applied Catalysis A: General*, 2010. 376(1–2): p. 56-65.
295. Werther, J. and J. Reppenhagen, Catalyst attrition in fluidized-bed systems. *AIChE Journal*, 1999. 45(9): p. 2001-2010.
296. Werther, J. and W. Xi, Jet attrition of catalyst particles in gas fluidized beds. *Powder Technology*, 1993. 76(1): p. 39-46.

297. Wu, D., J. Zhou, and Y. Li, Mechanical strength of solid catalysts: Recent developments and future prospects. *AIChE Journal*, 2007. 53(10): p. 2618-2629.
298. Xiao, G., J.R. Grace, and C.J. Lim, Limestone Particle Attrition in High-Velocity Air Jets. *Industrial & Engineering Chemistry Research*, 2012. 51(1): p. 556-560.
299. Zhang, Q., et al., Jet attrition in a fluidized bed. Part I: Effect of nozzle operating conditions. *Powder Technology*, 2012. 229: p. 162-169.
300. Zhang, Q., et al., Jet attrition in a fluidized bed. Part II: Effect of fluidized bed hydrodynamics. *Powder Technology*, 2012. 229: p. 170-177.
301. Gwyn, J.E., On the particle size distribution function and the attrition of cracking catalysts. *AIChE Journal*, 1969. 15(1): p. 35-39.
302. Cook, J.L., et al., Attrition and changes in particle size distribution of lime sorbents in a circulating fluidized bed absorber. *Powder Technology*, 1996. 89(1): p. 1-8.
303. Lee, S.K., et al., Attrition of lime sorbents during fluidization in a circulating fluidized bed absorber. *Industrial & Engineering Chemistry Research*, 1993. 32(11): p. 2758-2766.
304. McMillan, J., et al., High velocity attrition nozzles in fluidized beds. *Powder Technology*, 2007. 175(3): p. 133-141.
305. Pis, J.J., et al., Attrition of coal ash particles in a fluidized bed. *Powder Technology*, 1991. 66(1): p. 41-46.
306. Wu, S.-Y., J. Baeyens, and C.-Y. Chu, Effect of the grid-velocity on attrition in gas fluidized beds. *The Canadian Journal of Chemical Engineering*, 1999. 77(4): p. 738-744.
307. Zhang, Z. and M. Ghadiri, Impact attrition of particulate solids. Part 2: Experimental work. *Chemical Engineering Science*, 2002. 57(17): p. 3671-3686.

308. Cooper, D.J. and D.E. Clough, Real-time estimation of dynamic particle-size distributions in a fluidized bed: Theoretical foundation. *AIChE Journal*, 1985. 31(7): p. 1202-1212.
309. Feng, Y.Q. and A.B. Yu, Assessment of Model Formulations in the Discrete Particle Simulation of Gas–Solid Flow. *Industrial & Engineering Chemistry Research*, 2004. 43(26): p. 8378-8390.
310. Hatzantonis, H., A. Goulas, and C. Kiparissides, A comprehensive model for the prediction of particle-size distribution in catalyzed olefin polymerization fluidized-bed reactors. *Chemical Engineering Science*, 1998. 53(18): p. 3251-3267.
311. Klett, C., E.U. Hartge, and J. Werther, Time-dependent behavior of the ash particle size distribution in a circulating fluidized bed system. *Proceedings of the Combustion Institute*, 2005. 30(2): p. 2947-2954.
312. Massoudi, M. and E.J. Boyle, Modeling of solid particles in fluidized beds. *Mathematical and Computer Modelling*, 1990. 14: p. 785-789.
313. Milioli, F.E. and P.J. Foster, A model for particle size distribution and elutriation in fluidized beds. *Powder Technology*, 1995. 83(3): p. 265-280.
314. Närvänen, T., O. Antikainen, and J. Yliruusi, Predicting Particle Size During Fluid Bed Granulation Using Process Measurement Data. *AAPS PharmSciTech*, 2009. 10(4): p. 1268-1275.
315. Thornton, C., K. Yin, and M. Adams, Numerical simulation of the impact fracture and fragmentation of agglomerates. *Journal of Physics D: Applied Physics*, 1996. 29(2): p. 424.
316. Nakayama, H., M. Arakawa, and R. Sasaki, Simulation-Based Optimization Using Computational Intelligence. *Optimization and Engineering*, 2002. 3(2): p. 201-214.

317. Baumes, L., et al., Using Artificial Neural Networks to Boost High-throughput Discovery in Heterogeneous Catalysis. *QSAR & Combinatorial Science*, 2004. 23(9): p. 767-778.
318. Korany, M.A., et al., Application of artificial neural networks for response surface modelling in HPLC method development. *Journal of Advanced Research*, 2012. 3(1): p. 53-63.
319. López C, D.C., et al., Optimization Model of Crude Oil Distillation Units for Optimal Crude Oil Blending and Operating Conditions. *Industrial & Engineering Chemistry Research*, 2013. 52(36): p. 12993-13005.
320. Shao, P., S.T. Jiang, and Y.J. Ying, Optimization of Molecular Distillation for Recovery of Tocopherol from Rapeseed Oil Deodorizer Distillate Using Response Surface and Artificial Neural Network Models. *Food and Bioproducts Processing*, 2007. 85(2): p. 85-92.
321. Tang, Q., et al., Response surface methodology using Gaussian processes: Towards optimizing the trans-stilbene epoxidation over $\text{Co}_2\text{+NaX}$ catalysts. *Chemical Engineering Journal*, 2010. 156(2): p. 423-431.
322. Yan, W., et al., Bayesian migration of Gaussian process regression for rapid process modeling and optimization. *Chemical Engineering Journal*, 2011. 166(3): p. 1095-1103.
323. Krajzewicz, D., et al., Recent Development and Applications of SUMO—Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements*, 2012. 5(3 and 4): p. 128-138.
324. Krajzewicz, D., et al. Sumo (simulation of urban mobility). in *Proc. of the 4th Middle East Symposium on Simulation and Modelling*. 2002.
325. Arena, U., et al., Carbon attrition in the fluidized combustion of a metallurgical coke. *Combustion and Flame*, 1984. 57(1): p. 123-125.
326. Olejnik, T., Kinetics of grinding ceramic bulk considering grinding media contact points. *Physicochemical Problems of Mineral Processing*, 2010. 44: p. 187-194.

327. Olejnik, T.P., Milling kinetics of chosen rock materials under dry conditions considering strength and statistical properties of bed. *Physicochemical Problems of Mineral Processing*, 2011. 46: p. 145-154.
328. Austin, L., K. Shoji, and P.T. Luckie, The effect of ball size on mill performance. *Powder Technology*, 1976. 14(1): p. 71-79.
329. Fuerstenau, D., J. Lutch, and A. De, The effect of ball size on the energy efficiency of hybrid high-pressure roll mill/ball mill grinding. *Powder Technology*, 1999. 105(1): p. 199-204.
330. Heim, A.a.O., T.P. and Pawlak, A. Effect of the number of grinding media contact points on ceramite body grinding rate. in *Proceedings of 16 th International Congress of Chemical and Process Engineering*. 2004.
331. Kotake, N., et al., Experimental investigation on a grinding rate constant of solid materials by a ball mill—effect of ball diameter and feed size. *Powder Technology*, 2004. 143: p. 196-203.
332. Magdalinovic, N., M. Trumic, and L. Andric, The optimal ball diameter in a mill. *Physicochemical Problems of Mineral Processing*, 2012. 48(2): p. 329-339.
333. Deniz, V. and T. Onur, Investigation of the breakage kinetics of pumice samples as dependent on powder filling in a ball mill. *International journal of mineral processing*, 2002. 67(1): p. 71-78.
334. Shoji, N., Two-particle inclusive experiments. *Nuclear Physics A*, 1983. 400: p. 399-414.
335. Chew, J.W., et al., Species segregation of binary mixtures and a continuous size distribution of Group B particles in riser flow. *Chemical Engineering Science*, 2011. 66(20): p. 4595-4604.
336. Chew, J.W., et al., Cluster characteristics of Geldart group B particles in a pilot-scale CFB riser. II. Polydisperse systems. *Chemical Engineering Science*, 2012. 68(1): p. 82-93.

337. Farizhandi, A.A.K., H. Zhao, and R. Lau, Modeling the change in particle size distribution in a gas-solid fluidized bed due to particle attrition using a hybrid artificial neural network-genetic algorithm approach. *Chemical Engineering Science*, 2016. 155: p. 210-220.
338. Holland, J.H., *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. 1992: MIT Press.
339. Adi, H., et al., Role of agglomeration in the dispersion of salmeterol xinafoate from mixtures for inhalation with differing drug to fine lactose ratios. *Journal of Pharmaceutical Sciences*, 2008. 97(8): p. 3140-3152.
340. Martini, A., L. Muggetti, and M.P. Warchol, Nasal and pulmonary drug delivery systems. *Expert Opinion on Therapeutic Patents*, 2000. 10(3): p. 315-323.
341. Young, P.M., et al., The influence of drug loading on formulation structure and aerosol performance in carrier based dry powder inhalers. *International Journal of Pharmaceutics*, 2011. 416(1): p. 129-135.
342. Mitchell JP, N.M., Particle size analysis of aerosols from medicinal inhalers. *KONA*, 2004. 22: p. 32–65.
343. Stephenson, P.L. and W.J. Thiel, The effect of humidity on the production of ordered mixtures. *Powder Technology*, 1980. 25(1): p. 115-119.
344. Hickey, A.J., Concessio, N.M., Van Oort, M.M., & Platz, R.M. , Factors influencing the dispersion of dry powders as aerosols. *Pharmaceutical Technology*, 1994. 18(8): p. 58-64.
345. Byron, P.R., *Some Future Perspectives for Unit Dose Inhalation Aerosols*. *Drug Development and Industrial Pharmacy*, 1986. 12(7): p. 993-1015.
346. El-Gendy, N., et al., Development of budesonide nanocluster dry powder aerosols: Formulation and stability. *Journal of Pharmaceutical Sciences*, 2012. 101(9): p. 3445-3455.

347. Larhrib, H., et al., The use of different grades of lactose as a carrier for aerosolised salbutamol sulphate. *International Journal of Pharmaceutics*, 1999. 191(1): p. 1-14.
348. Le, V.N.P., et al., Dry Powder Inhalers: Study of the Parameters Influencing Adhesion and Dispersion of Fluticasone Propionate. *AAPS PharmSciTech*, 2012. 13(2): p. 477-484.
349. Tan, M.J.B., et al., Particle Surface Roughness – Its Characterisation and Impact on Dry Powder Inhaler Performance, in *Pulmonary Drug Delivery*. 2015, John Wiley & Sons, Ltd. p. 199-222.
350. Rahimpour, Y. and H. Hamishehkar, Lactose Engineering for Better Performance in Dry Powder Inhalers. *Advanced Pharmaceutical Bulletin*, 2012. 2(2): p. 183-187.
351. Timsina, M.P., et al., Drug delivery to the respiratory tract using dry powder inhalers. *International Journal of Pharmaceutics*, 1994. 101(1): p. 1-13.
352. Zhou, Q. and D.A.V. Morton, Drug–lactose binding aspects in adhesive mixtures: Controlling performance in dry powder inhaler formulations by altering lactose carrier surfaces. *Advanced Drug Delivery Reviews*, 2012. 64(3): p. 275-284.
353. Kumon, M., et al., Application and Mechanism of Inhalation Profile Improvement of DPI Formulations by Mechanofusion with Magnesium Stearate. *Chemical and Pharmaceutical Bulletin*, 2008. 56(5): p. 617-625.
354. Zeng, X.M., et al., The Effects of Carrier Size and Morphology on the Dispersion of Salbutamol Sulphate after Aerosolization at Different Flow Rates. *Journal of Pharmacy and Pharmacology*, 2000. 52(10): p. 1211-1221.
355. Kou, X., et al., Physico-chemical aspects of lactose for inhalation. *Advanced Drug Delivery Reviews*, 2012. 64(3): p. 220-232.

356. Adi, H., I. Larson, and P.J. Stewart, Adhesion and redistribution of salmeterol xinafoate particles in sugar-based mixtures for inhalation. *International Journal of Pharmaceutics*, 2007. 337(1–2): p. 229-238.
357. Steckel, H. and B.W. Müller, In vitro evaluation of dry powder inhalers II: influence of carrier particle size and concentration on in vitro deposition. *International Journal of Pharmaceutics*, 1997. 154(1): p. 31-37.
358. Harjunen, P., et al., Effects of carriers and storage of formulation on the lung deposition of a hydrophobic and hydrophilic drug from a DPI. *International Journal of Pharmaceutics*, 2003. 263(1–2): p. 151-163.
359. Hassan, M.S. and R. Lau, Inhalation performance of pollen-shape carrier in dry powder formulation with different drug mixing ratios: Comparison with lactose carrier. *International Journal of Pharmaceutics*, 2010. 386(1–2): p. 6-14.
360. Hickey, A.J. and N.M. Concessio, Descriptors of irregular particle morphology and powder properties. *Advanced Drug Delivery Reviews*, 1997. 26(1): p. 29-40.
361. Concessio, N., et al., Pharmaceutical Dry Powder Aerosols: Correlation of Powder Properties with Dose Delivery and Implications for Pharmacodynamic Effect. *Pharmaceutical Research*, 1999. 16(6): p. 828-834.
362. Donovan, M.J. and H.D.C. Smyth, Influence of size and surface roughness of large lactose carrier particles in dry powder inhaler formulations. *International Journal of Pharmaceutics*, 2010. 402(1–2): p. 1-9.
363. Glover, W., et al., Effect of particle size of dry powder mannitol on the lung deposition in healthy volunteers. *International Journal of Pharmaceutics*, 2008. 349(1–2): p. 314-322.

364. Guenette, E., et al., Understanding the effect of lactose particle size on the properties of DPI formulations using experimental design. *International Journal of Pharmaceutics*, 2009. 380(1–2): p. 80-88.
365. Karhu, M., et al., Pulmonary deposition of lactose carriers used in inhalation powders. *International Journal of Pharmaceutics*, 2000. 196(1): p. 95-103.
366. Louey, M. and P. Stewart, Particle Interactions Involved in Aerosol Dispersion of Ternary Interactive Mixtures. *Pharmaceutical Research*, 2002. 19(10): p. 1524-1531.
367. Louey, M.D., S. Razia, and P.J. Stewart, Influence of physico-chemical carrier properties on the in vitro aerosol deposition from interactive mixtures. *International Journal of Pharmaceutics*, 2003. 252(1–2): p. 87-98.
368. Podczek, F., The Influence of Particle Size Distribution and Surface Roughness of Carrier Particles on the in vitro Properties of Dry Powder Inhalations. *Aerosol Science and Technology*, 1999. 31(4): p. 301-321.
369. Young, P.M., et al., The influence of mechanical processing of dry powder inhaler carriers on drug aerosolization performance. *Journal of Pharmaceutical Sciences*, 2007. 96(5): p. 1331-1341.
370. Podczek, F., Adhesion forces in interactive powder mixtures of a micronized drug and carrier particles of various particle size distributions. *Journal of Adhesion Science and Technology*, 1998. 12(12): p. 1323-1339.
371. Hassan, M.S. and R. Lau, Inhalation performance of pollen-shape carrier in dry powder formulation: Effect of size and surface morphology. *International Journal of Pharmaceutics*, 2011. 413(1–2): p. 93-102.
372. Kaialy, W. and A. Nokhodchi, Freeze-Dried Mannitol for Superior Pulmonary Drug Delivery via Dry Powder Inhaler. *Pharmaceutical Research*, 2013. 30(2): p. 458-477.

373. Maas, S.G., et al., Tailoring dry powder inhaler performance by modifying carrier surface topography by spray drying. *Atomization and Sprays*, 2010. 20(9).
374. Littringer, E.M., et al., Spray dried mannitol carrier particles with tailored surface properties – The influence of carrier surface roughness and shape. *European Journal of Pharmaceutics and Biopharmaceutics*, 2012. 82(1): p. 194-204.
375. Zeng, X.M., et al., The role of fine particle lactose on the dispersion and deaggregation of salbutamol sulphate in an air stream in vitro. *International Journal of Pharmaceutics*, 1998. 176(1): p. 99-110.
376. Guchardi, R., et al., Influence of fine lactose and magnesium stearate on low dose dry powder inhaler formulations. *International journal of pharmaceutics*, 2008. 348(1): p. 10-17.
377. Handoko, A., L. Ian, and S.J. Peter, Influence of the polydispersity of the added fine lactose on the dispersion of salmeterol xinafoate from mixtures for inhalation. *European Journal of Pharmaceutical Sciences*, 2009. 36(2–3): p. 265-274.
378. Podczec, F., Optimization of the operation conditions of an Andersen-Cascade impactor and the relationship to centrifugal adhesion measurements to aid the development of dry powder inhalations. *International Journal of Pharmaceutics*, 1997. 149(1): p. 51-61.
379. Boshhiha, A.M. and N.A. Urbanetz, Influence of carrier surface fines on dry powder inhalation formulations. *Drug Development and Industrial Pharmacy*, 2009. 35(8): p. 904-916.
380. Zhang, F., et al., Roughness-Controlled Self-Assembly of Mannitol/LB Agar Microparticles by Polymorphic Transformation for Pulmonary Drug Delivery. *Molecular Pharmaceutics*, 2015. 12(1): p. 223-231.
381. Kawashima, Y., et al., Effect of surface morphology of carrier lactose on dry powder inhalation property of pranlukast hydrate. *International Journal of Pharmaceutics*, 1998. 172(1–2): p. 179-188.

382. P.W.S. Heng, L.W.C., Liang Theng LIM, Quantification of the Surface Morphologies of Lactose Carriers and Their Effect on the in Vitro Deposition of Salbutamol Sulphate. *CHEMICAL & Pharmaceutical Bulletin*, 2000. 48(3): p. 393-398.
383. Hassan, M.S. and R. Lau, Feasibility study of pollen-shape drug carriers in dry powder inhalation. *Journal of pharmaceutical sciences*, 2010. 99(3): p. 1309-1321.
384. Nguyen, T.Q.N., et al., Surface Modification of Pollen-Shape Carriers for Dry Powder Inhalation through Surface Etching. *Industrial & Engineering Chemistry Research*, 2014. 53(51): p. 19943-19950.
385. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat methods*, 2012. 9(7): p. 671-675.
386. Abràmoff, M.D., P.J. Magalhães, and S.J. Ram, Image processing with ImageJ. *Biophotonics international*, 2004. 11(7): p. 36-42.
387. Chinga, G., SurfCharJ. 2003: Available at: <http://www.gcscsca.net>.
388. Chinga, G., et al., Quantification of the 3D microstructure of SC surfaces. *Journal of Microscopy*, 2007. 227(3): p. 254-265.
389. Ganderton, D. and N. Kassem, Dry powder inhalers. *Advances in Pharmaceutical Sciences* Academic Press, 1992: p. 165–191.
390. Kassem, N., K.K.L. Ho, and D. Ganderton, The effect of air flow and carrier size on the characteristics of an inspirable cloud. *Journal Of Pharmacy & Pharmacology*, 1989. 41: p. 14.
391. Kassem, N., D. Ganderton, and M. Yianneskis, The use of laser doppler anemometry to assess the sampling of a particulate cloud. *Journal Of Pharmacy & Pharmacology*, 1989. 41: p. 13.
392. Schreier, H., et al., Formulation and In-Vitro Performance of Liposome Powder Aerosols. *Stp Pharma Sciences*, 1994. 4(1): p. 38-44.

393. Lucas, P., K. Anderson, and J.N. Staniforth, Protein deposition from dry powder inhalers: fine particle multiplets as performance modifiers. *Pharmaceutical Research*, 1998. 15(4): p. 562-569.
394. Islam, N., et al., Lactose surface modification by decantation: are drug-fine lactose ratios the key to better dispersion of salmeterol xinafoate from lactose-interactive mixtures? *Pharmaceutical Research*, 2004. 21(3): p. 492-499.
395. French, D.L., D.A. Edwards, and R.W. Niven, The influence of formulation on emission, deaggregation and deposition of dry powders for inhalation. *Journal of Aerosol Science*, 1996. 27(5): p. 769-783.
396. de Boer, A.H., et al., A critical evaluation of the relevant parameters for drug redispersion from adhesive mixtures during inhalation. *International Journal of Pharmaceutics*, 2005. 294(1–2): p. 173-184.
397. Hamishehkar, H., et al., Influence of carrier particle size, carrier ratio and addition of fine ternary particles on the dry powder inhalation performance of insulin-loaded PLGA microcapsules. *Powder Technology*, 2010. 201(3): p. 289-295.
398. Du, P., J. Du, and H.D.C. Smyth, Evaluation of Granulated Lactose as a Carrier for DPI Formulations 1: Effect of Granule Size. *AAPS PharmSciTech*, 2014. 15(6): p. 1417-1428.
399. Heng, P.W.S., L.W. Chan, and L.T. Lim, Quantification of the surface morphologies of lactose carriers and their effect on the in vitro deposition of salbutamol sulphate. *Chemical and pharmaceutical bulletin*, 2000. 48(3): p. 393-398.
400. Iida, K., et al., Effect of Surface Covering of Lactose Carrier Particles on Dry Powder Inhalation Properties of Salbutamol Sulfate. *Chemical and Pharmaceutical Bulletin*, 2003. 51(12): p. 1455-1457.
401. Iida, K., et al., Effect of Surface Layering Time of Lactose Carrier Particles on Dry Powder Inhalation Properties of Salbutamol Sulfate. *Chemical and Pharmaceutical Bulletin*, 2004. 52(3): p. 350-353.

402. Staniforth, J.N., et al., Interparticle forces in binary and ternary ordered powder mixes. *Journal of Pharmacy and Pharmacology*, 1982. 34(3): p. 141-145.
403. Chan, L.W., L.T. Lim, and P.W.S. Heng, Immobilization of fine particles on lactose carrier by precision coating and its effect on the performance of dry powder formulations. *Journal of Pharmaceutical Sciences*, 2003. 92(5): p. 975-984.
404. Iida, K., et al., Preparation of Dry Powder Inhalation with Lactose Carrier Particles Surface-Coated Using a Wurster Fluidized Bed. *Chemical and Pharmaceutical Bulletin*, 2005. 53(4): p. 431-434.
405. Farizhandi, A.A.K. and R. Lau, Evaluation of hydroxyapatite size and morphology in dry powder inhalation for carrier-based pulmonary delivery formulations by response surface methodology. *International Journal of Engineering Science*, 2016.
406. Kazemzadeh Farizhandi, A.A.K., H. Zhao, and R. Lau, Modeling the change in particle size distribution in a gas-solid fluidized bed due to particle attrition using a hybrid artificial neural network-genetic algorithm approach. *Chemical Engineering Science*, 2016.
407. Kho, K. and K. Hadinoto, Dry powder inhaler delivery of amorphous drug nanoparticles: Effects of the lactose carrier particle shape and size. *Powder Technology*, 2013. 233: p. 303-311.
408. Mansouri, S., et al., Precipitating smooth amorphous or pollen structured lactose microparticles. *Chemical Engineering Journal*, 2013. 226: p. 312-318.
409. Kaialy, W., et al., Influence of lactose carrier particle size on the aerosol performance of budesonide from a dry powder inhaler. *Powder Technology*, 2012. 227: p. 74-85.
410. Kinnunen, H., et al., An Investigation into the Effect of Fine Lactose Particles on the Fluidization Behaviour and Aerosolization Performance

of Carrier-Based Dry Powder Inhaler Formulations. AAPS PharmSciTech, 2014. 15(4): p. 898-909.

- 411. Saint-Lorant, G., et al., Influence of carrier on the performance of dry powder inhalers. International Journal of Pharmaceutics, 2007. 334(1–2): p. 85-91.
- 412. Zanen, P., L.T. Go, and J.W. Lammers, Optimal particle size for beta 2 agonist and anticholinergic aerosols in patients with severe airflow obstruction. Thorax, 1996. 51(10): p. 977-980.
- 413. Zanen, P., L.T. Go, and J.-W.J. Lammers, The optimal particle size for β -adrenergic aerosols in mild asthmatics. International Journal of Pharmaceutics, 1994. 107(3): p. 211-217.
- 414. Zanen, P., L.T. Go, and J.-W.J. Lammers, The optimal particle size for parasympatholytic aerosols in mild asthmatics. International Journal of Pharmaceutics, 1995. 114(1): p. 111-115.
- 415. Usmani, O.S., et al., Effects of bronchodilator particle size in asthmatic patients using monodisperse aerosols. Journal of Applied Physiology, 2003. 95(5): p. 2106.

Appendix

The genetic algorithm generation versus fitness value have been shown in Figures A1, A2, A3, and A4.

Figure A1: Generation versus fitness value for ANN-GA in IBA fluidization (chapter 4)

Figure A2: Generation versus fitness value for ANN-GA in all materials fluidization (chapter 5)

Figure A3: Generation versus fitness value for ANN-GA in HA carrier DPI (chapter 6)

Figure A4: Generation versus fitness value for ANN-GA in all carriers (chapter 7)

The ANN weights and biases have been presented in Tables A1, A2, A3, A4, A5, A6, A7 and A8.

Table A1: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of d

(a) Input weights

	neuron 1 input	neuron 2 input	neuron 3 input	neuron 4 input	neuron 5 input
neuron 1 hidden layer 1	-0.3370	0.1741	-0.0909	-0.9555	0.5226
neuron 2 hidden layer 1	0.6307	13.6711	1.7769	-1.8821	-3.5506
neuron 3 hidden layer 1	-1.1344	-0.0495	0.2307	-0.1376	2.5178

(b) Hidden layers weights

	neuron 1 hidden layer 1	neuron 2 hidden layer 1	neuron 3 hidden layer 1
neuron 1 hidden layer 2	-1.1593	2.0645	-0.6566
neuron 2 hidden layer 2	-1.7896	2.6109	1.1730
neuron 3 hidden layer 2	-3.8412	4.1662	-3.2017
	neuron 1 hidden layer 2	neuron 2 hidden layer 2	neuron 3 hidden layer 2
neuron 1 hidden layer 3	-0.3800	-0.1167	-0.0303
neuron 2 hidden layer 3	-0.6134	2.1024	0.1680
neuron 3 hidden layer 3	-3.7770	1.8765	-3.5651

(c) Output weights

	neuron 1 hidden layer 3	neuron 2 hidden layer 3	neuron 3 hidden layer 3
neuron 1 output layer	0.7461	0.0789	0.1228

(d) Biases

	hidden layer 1	hidden layer 2	hidden layer 3	output layer
neuron 1	1.9225	-0.1551	1.2900	0.1570
neuron 2	-0.0913	-5.3937	-0.0232	-
neuron 3	-1.9616	-0.8302	-0.8589	-

Table A2: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of n

(a) Input weights

	neuron 1 input	neuron 2 input	neuron 3 input	neuron 4 input	neuron 5 input
neuron 1 hidden layer 1	1.1403	0.1060	-0.1472	-0.2462	-0.1225
neuron 2 hidden layer 1	-2.2627	1.3211	-9.6859	-0.3647	1.6512
neuron 3 hidden layer 1	-0.3866	0.2266	-0.3143	0.5057	0.8141

(b) Hidden layers weights

	neuron 1 hidden layer 1	neuron 2 hidden layer 1	neuron 3 hidden layer 1
neuron 1 hidden layer 2	0.9185	1.7290	8.8853
neuron 2 hidden layer 2	-3.1151	-0.8032	2.8539
neuron 3 hidden layer 2	2.7986	-0.8777	-10.0284
	neuron 1 hidden layer 2	neuron 2 hidden layer 2	neuron 3 hidden layer 2
neuron 1 hidden layer 3	2.4337	0.7941	9.1822
neuron 2 hidden layer 3	-2.6873	0.3398	3.9537
neuron 3 hidden layer 3	3.2286	-0.5907	-6.5708

(c) Output weights

	neuron 1 hidden layer 3	neuron 2 hidden layer 3	neuron 3 hidden layer 3
neuron 1 output layer	-10.3659	-14.0701	-3.6522

(d) *Biases*

	hidden layer 1	hidden layer 2	hidden layer 3	output layer
neuron 1	2.3528	1.0254	-8.7613	0.3367
neuron 2	2.4677	1.3851	-3.3931	
neuron 3	0.2873	0.3925	6.0332	

Table A3: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of d

(e) Input weights

	neuron 1 input	neuron 2 input	neuron 3 input	neuron 4 input	neuron 5 input
neuron 1 hidden layer 1	10.5267	0.7259	-0.8424	2.6215	12.8088
neuron 2 hidden layer 1	2.6981	3.3696	2.7826	3.9570	-4.8417
neuron 3 hidden layer 1	5.0463	-3.9574	2.28078	3.6800	-1.0333
neuron 4 hidden layer 1	8.5172	-5.2775	-2.5466	1.5039	7.2570
neuron 5 hidden layer 1	14.3776	3.4108	1.4583	-3.3240	-2.7038

(f) Hidden layers weights

	neuron 1 hidden layer 1	neuron 2 hidden layer 1	neuron 3 hidden layer 1	neuron 4 hidden layer 1	neuron 5 hidden layer 1
neuron 1 hidden layer 2	0.9256	-1.2409	1.0068	-0.2666	0.1256
neuron 2 hidden layer 2	-0.1885	-1.6845	0.3149	-0.7356	1.8583
neuron 3 hidden layer 2	0.4076	1.0679	-2.2073	-2.3581	1.5022
neuron 4 hidden layer 2	0.6038	-1.17503	-0.3076	1.4484	0.0654
neuron 5 hidden layer 2	1.6334	-0.2833	-0.1940	-0.9154	0.7430
neuron 6 hidden layer 2	-0.3461	-1.3731	0.8734	1.3166	0.6308

(g) *Output weights*

	neuron 1 hidden layer 2	neuron 2 hidden layer 2	neuron 3 hidden layer 2	neuron 4 hidden layer 2	neuron 5 hidden layer 2	neuron 6 hidden layer 2
neuron 1 output layer	-0.3817	-1.6048	-0.6368	1.8390	-0.1862	-0.3414

(h) *Biases*

	hidden layer 1	hidden layer 2	output layer
neuron 1	-21.6909	-2.3448	0.1014
neuron 2	-6.3951	0.3014	-
neuron 3	-4.8690	-0.6675	-
neuron 4	-12.1265	0.6164	-
neuron 5	-6.2254	0.4886	-
neuron 6	-	-1.5221	-

Table A4: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of n

(e) *Input weights*

	neuron 1 input	neuron 2 input	neuron 3 input	neuron 4 input	neuron 5 input
neuron 1 hidden layer 1	-0.2138	-0.1210	-0.0365	-4.1062	-3.8324
neuron 2 hidden layer 1	-0.3943	-0.0912	-0.0474	-0.8641	3.0452
neuron 3 hidden layer 1	0.6764	4.5586	-5.9945	7.1497	-4.2814

(f) *Hidden layers weights*

	neuron 1 hidden layer 1	neuron 2 hidden layer 1	neuron 3 hidden layer 1
neuron 1 hidden layer 2	-0.4439	-1.3136	4.8491
neuron 2 hidden layer 2	0.9205	-1.9276	3.4529
neuron 3 hidden layer 2	4.1820	0.0524	3.1299

(g) *Output weights*

	neuron 1 hidden layer 3	neuron 2 hidden layer 3	neuron 3 hidden layer 3
neuron 1 output layer	-6.1001	2.776980	3.1315

(h) *Biases*

	hidden layer 1	hidden layer 2	output layer
neuron 1	7.7803	-1.4960	1.3965
neuron 2	-1.4018	-3.4233	-
neuron 3	-6.7067	6.8352	-

Table A5: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of ED for HA carrier

(i) *Input weights*

	neuron 1 input	neuron 2 input	neuron 3 input	neuron 4 input	neuron 5 input	neuron 6 input	neuron 7 input	neuron 8 input	neuron 9 input	neuron 10 input	neuron 11 input	neuron 12 input	neuron 13 input
neuron 1 hidden layer 1	0.2636	0.0422	-0.4440	0.1549	0.8702	-0.7672	0.2854	1.1493	-0.2725	0.2522	0.2925	-0.5205	1.0326
neuron 2 hidden layer 1	-0.0800	0.0541	-0.2558	-0.0332	-0.0977	0.0616	0.4040	-0.0840	-0.2786	-0.1773	-0.0910	-0.2428	-0.1119
neuron 3 hidden layer 1	-0.3992	-0.7369	0.2273	-0.5057	1.3173	0.7316	0.5650	-0.6826	0.3058	-0.0816	-0.4157	0.2421	-0.1040
neuron 4 hidden layer 1	-0.5281	-0.6843	-0.2238	0.5198	-0.8066	-1.0280	1.4061	0.1624	-0.3252	-0.5431	-0.6105	-0.1765	-0.0114
neuron 5 hidden layer 1	-0.1090	0.0598	-0.0179	-0.1575	-0.0118	0.3681	-0.7787	-0.4834	-0.0723	-0.2012	-0.1716	0.0190	-0.4015
neuron 6 hidden layer 1	0.1110	-0.0242	0.1276	0.1388	-0.1771	-0.0223	0.6697	0.2959	0.1555	0.1497	0.0972	0.1127	0.2275
neuron 7 hidden layer 1	-0.3580	-0.2557	-0.7948	1.3903	0.7044	-2.1207	-0.4892	1.0330	-0.9407	-0.2528	0.3269	-0.8041	-0.1991

(j) *Hidden layers weights*

	neuron 1 hidden layer 1	neuron 2 hidden layer 1	neuron 3 hidden layer 1	neuron 4 hidden layer 1	neuron 5 hidden layer 1	neuron 6 hidden layer 1	neuron 7 hidden layer 1
neuron 1 hidden layer 2	-0.8966	0.2799	-0.4678	0.1749	0.4651	-0.3389	0.2873
neuron 2 hidden layer 2	0.5167	-0.4976	0.4870	-0.5039	-0.5223	0.4958	-0.5149
neuron 3 hidden layer 2	0.4647	-0.1917	0.5841	-0.5702	-0.3290	0.1231	-0.9269
neuron 4 hidden layer 2	0.5792	-0.4686	0.4870	-0.4672	-0.4141	0.4121	-0.4522
neuron 5 hidden layer 2	0.3871	-0.2130	0.5991	-0.7515	-0.1686	0.1881	-1.0841
neuron 6 hidden layer 2	0.6931	-0.3433	0.4948	-0.4284	-0.3954	0.3283	-0.4227
neuron 7 hidden layer 2	0.3767	-0.1575	0.6199	-0.7284	-0.2182	0.1350	-1.0601
neuron 8 hidden layer 2	-0.4381	0.2478	-0.6047	0.6735	0.2640	-0.1328	1.0282
neuron 9 hidden layer 2	-0.0752	0.0619	-0.7055	0.3803	0.0551	-0.0355	1.9669
neuron 10 hidden layer 2	-0.7567	0.3274	-0.5330	0.3379	0.4149	-0.2984	0.4553
neuron 11 hidden layer 2	0.5097	-0.3245	0.5665	-0.5057	-0.3309	0.1759	-0.8414
neuron 12 hidden layer 2	-0.7031	0.3617	-0.4204	0.3979	0.4049	-0.3833	0.4548

(k) Output weights

	neuron 1 hidden layer 2	neuron 2 hidden layer 2	neuron 3 hidden layer 2	neuron 4 hidden layer 2	neuron 5 hidden layer 2	neuron 6 hidden layer 2	neuron 7 hidden layer 2	neuron 8 hidden layer 2	neuron 9 hidden layer 2	neuron 10 hidden layer 2	neuron 11 hidden layer 2	neuron 12 hidden layer 2
neuron 1 output layer	-6.6361	4.8574	7.6150	7.4575	7.7044	7.0215	7.3971	-7.3155	-7.5779	-7.2799	7.7928	-7.5445

(l) Biases

neuron 1 input	-0.4326	neuron 1 hidden layer 2	-0.3261	neuron 1 output layer	7.8518
neuron 2 input	-0.2635	neuron 2 hidden layer 2	0.5306		
neuron 3 input	0.2445	neuron 3 hidden layer 2	0.3978		
neuron 4 input	-0.2022	neuron 4 hidden layer 2	0.3374		
neuron 5 input	-0.0236	neuron 5 hidden layer 2	0.1825		
neuron 6 input	0.1178	neuron 6 hidden layer 2	0.3715		
neuron 7 input	-0.7808	neuron 7 hidden layer 2	0.2700		
		neuron 8 hidden layer 2	-0.1820		
		neuron 9 hidden layer 2	-0.1998		
		neuron 10 hidden layer 2	-0.3337		
		neuron 11 hidden layer 2	0.3548		
		neuron 12 hidden layer 2	-0.3698		

Table A6: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of FPF for HA carrier

(a) Input weights

	neuron 1 input	neuron 2 input	neuron 3 input	neuron 4 input	neuron 5 input	neuron 6 input	neuron 7 input	neuron 8 input	neuron 9 input	neuron 10 input	neuron 11 input	neuron 12 input
neuron 1 hidden layer 1	0.7263	-3.8132	-1.0733	0.1620	-5.5566	-2.6921	3.7308	-1.4234	4.6127	-2.5376	-12.5131	2.0625
neuron 2 hidden layer 1	0.1901	4.3335	-1.0383	-0.0944	0.3853	-0.8570	0.8849	1.8091	6.3203	-2.3174	8.0672	-7.5413
neuron 3 hidden layer 1	-0.5981	1.6015	-1.3361	-0.5071	2.0212	0.2619	-5.9684	0.2511	-10.1153	1.5073	9.1898	3.1032
neuron 4 hidden layer 1	-0.2272	2.9697	-0.1173	-0.3897	-6.3415	2.5916	-6.2736	1.0385	-7.0655	0.0607	-7.3281	6.9789
neuron 5 hidden layer 1	0.4331	1.7141	2.0112	0.8102	-0.1265	0.8417	-2.3295	0.2910	6.1851	3.3231	-7.6098	4.2078
neuron 6 hidden layer 1	1.4906	0.0038	0.9620	0.6817	-7.8543	1.9870	-7.7244	2.7608	-1.0681	0.5039	8.2984	-4.3153
neuron 7 hidden layer 1	-0.6425	-3.6365	-1.2105	0.6709	1.3288	1.9338	6.3975	-1.1473	-4.5559	2.0397	13.5195	0.2053
neuron 8 hidden layer 1	-1.8280	3.1611	-0.0033	-1.6974	-2.7044	-1.3400	-7.5065	-2.2221	-7.9330	-0.5766	3.5009	2.2558
neuron 9 hidden layer 1	-2.6099	-0.1212	1.4947	-3.4834	5.0173	-1.3496	6.3763	-0.7221	-2.0967	0.0250	-9.5647	7.2898
neuron 10 hidden layer 1	0.4987	-1.0760	0.6106	-1.2231	4.9695	1.7080	7.5174	1.3704	9.8188	1.8741	-6.7151	2.1246

(b) *Hidden layers weights*

	neuron 1 hidden layer 1	neuron 2 hidden layer 1	neuron 3 hidden layer 1	neuron 4 hidden layer 1	neuron 5 hidden layer 1	neuron 6 hidden layer 1	neuron 7 hidden layer 1	neuron 8 hidden layer 1	neuron 9 hidden layer 1	neuron 10 hidden layer 1
neuron 1 hidden layer 2	0.1402	0.3609	-0.1065	1.1769	-0.7456	0.5439	-0.5778	-0.6621	0.2307	-0.2836
neuron 2 hidden layer 2	0.7861	2.0953	-0.6018	-0.5144	0.1150	-1.0347	-0.4078	-0.7287	-0.6504	1.4093
neuron 3 hidden layer 2	-0.2895	-0.3107	0.4485	-1.8964	-1.5635	-0.2361	0.6552	-0.9146	1.7564	-0.3133
neuron 4 hidden layer 2	-2.8773	0.4747	-0.3552	-1.0091	2.2259	2.4939	1.0512	0.0534	1.9119	0.1972
neuron 5 hidden layer 2	-0.9330	1.6947	-0.0093	-1.1721	0.1412	0.2906	0.3043	0.1260	-0.3464	-0.2470
neuron 6 hidden layer 2	0.5303	-0.4058	0.3103	-0.1115	-1.4366	-1.0082	-0.7707	-1.0415	-1.2806	-0.9585
neuron 7 hidden layer 2	1.8436	0.3859	0.0415	-0.1013	1.5550	0.0632	1.0058	1.1387	1.2046	-1.9748
neuron 8 hidden layer 2	-0.9126	1.6149	-0.3813	-1.1216	-0.5687	2.0172	0.8956	-0.5594	0.7001	1.6369
neuron 9 hidden layer 2	-0.0035	-0.3262	1.5654	-1.0258	0.2401	0.8619	-1.1266	-0.6183	0.6367	-1.7574

(c) *Output weights*

	neuron 1 hidden layer 2	neuron 2 hidden layer 2	neuron 3 hidden layer 2	neuron 4 hidden layer 2	neuron 5 hidden layer 2	neuron 6 hidden layer 2	neuron 7 hidden layer 2	neuron 8 hidden layer 2	neuron 9 hidden layer 2
neuron 1 output layer	1.5783	-0.8410	0.8204	-2.5248	-1.9422	-0.0142	0.5269	1.4546	-0.0756

(d) *Biases*

neuron 1 input	-39.0451	neuron 1 hidden layer 2	-2.7353	neuron 1 output layer	-0.2779
neuron 2 input	-21.8202	neuron 2 hidden layer 2	-1.9091		
neuron 3 input	-47.8978	neuron 3 hidden layer 2	0.6422		
neuron 4 input	4.0387	neuron 4 hidden layer 2	1.7783		
neuron 5 input	6.4569	neuron 5 hidden layer 2	2.3754		
neuron 6 input	1.2515	neuron 6 hidden layer 2	2.5418		
neuron 7 input	21.3523	neuron 7 hidden layer 2	0.3631		
neuron 8 input	25.9176	neuron 8 hidden layer 2	-0.9615		
neuron 9 input	11.5018	neuron 9 hidden layer 2	0.0098		
neuron 10 input	6.4487				

Table A7: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of ED for all carriers*(a) Input weights*

	neuron 1 input	neuron 2 input	neuron 3 input	neuron 4 input	neuron 5 input	neuron 6 input	neuron 7 input	neuron 8 input	neuron 9 input	neuron 10 input	neuron 11 input	neuron 12 input	neuron 13 input	neuron 14 input	neuron 15 input	neuron 16 input
neuron 1 hidden layer 1	0.2947	1.2086	-0.2761	0.0766	2.7757	-0.4350	2.6366	3.6481	1.2961	3.9342	0.4292	-0.1123	-2.0150	2.8616	-4.2872	3.8273
neuron 2 hidden layer 1	1.8655	0.5594	-0.0193	-0.4117	1.0972	-2.1710	-1.2308	2.9361	0.6296	1.3593	-1.3288	-1.1991	-2.9528	1.3333	2.6693	-0.6031
neuron 3 hidden layer 1	-0.1639	-0.7003	2.0433	-1.8373	1.8062	-4.8505	-3.9623	3.4616	-0.0882	3.8183	0.3030	2.2993	0.0897	-1.4481	-3.2644	-3.4412
neuron 4 hidden layer 1	-0.2257	-1.9770	1.1311	-0.1471	1.4299	-1.3648	-3.2156	1.2863	2.0438	4.1783	0.3949	-0.3305	0.4728	-1.5560	3.7627	-3.1160
neuron 5 hidden layer 1	1.6682	-0.2325	1.5592	-1.0205	4.0007	2.3508	-2.3981	4.9590	-0.7039	-0.1672	-0.7776	-0.0376	-2.1834	-0.2432	2.0331	-0.0738
neuron 6 hidden layer 1	-0.0825	-0.7661	0.3710	-0.2504	2.8686	-5.3343	2.7871	-4.4118	-2.2445	-2.7643	-1.0088	0.1841	-2.0137	0.6936	0.6296	0.2240
neuron 7 hidden layer 1	-1.8454	1.9024	0.5627	1.8890	-0.8796	-3.9526	-3.2827	-5.5446	-2.8079	-3.7145	0.9790	-0.6017	0.2847	0.8602	1.5772	0.0998
neuron 8 hidden layer 1	0.9956	2.6876	0.7880	-1.7169	2.6267	4.4358	1.6338	-1.3290	1.3632	6.6710	0.5319	0.3406	1.8652	-2.3902	1.9994	1.5315
neuron 9 hidden layer 1	-0.4793	0.4806	-0.7542	0.1763	2.7566	1.3306	-1.4749	3.0893	-1.1681	-5.1515	1.8539	0.7879	1.8472	1.3492	2.2846	-2.7531
neuron 10 hidden layer 1	1.5371	-2.5642	-1.8947	1.5500	-2.1413	-3.9390	1.6694	-5.1198	0.3535	-3.0601	2.2629	-0.9210	3.3998	0.8770	-0.7890	-0.7247
neuron 11 hidden layer 1	-0.0448	1.8868	0.6795	-0.1735	2.5680	-3.2845	-2.4108	-1.1050	1.8371	2.4214	-0.7177	2.3310	0.5688	0.7278	-1.5504	0.1076
neuron 12 hidden layer 1	-0.0025	1.7893	0.2304	-1.4728	0.7373	0.6658	-2.6321	-1.0664	-2.3274	3.5563	0.6387	0.6924	1.5055	-0.0502	2.1956	3.3075
neuron 13 hidden layer 1	2.1378	2.9360	1.7538	-1.3471	0.9969	-2.5266	0.6717	-3.8441	-1.4469	-1.7405	0.9534	0.1196	-1.2978	0.5350	3.3708	0.1582

(b) Hidden layers weights

	neuron 1 hidden layer 1	neuron 2 hidden layer 1	neuron 3 hidden layer 1	neuron 4 hidden layer 1	neuron 5 hidden layer 1	neuron 6 hidden layer 1	neuron 7 hidden layer 1	neuron 8 hidden layer 1	neuron 9 hidden layer 1	neuron 10 hidden layer 1	neuron 11 hidden layer 1	neuron 12 hidden layer 1	neuron 13 hidden layer 1
neuron 1 hidden layer 2	-1.0157	-2.0904	0.2916	-1.5604	1.3571	0.6703	0.1472	-1.8240	-0.0528	1.0838	-1.4118	0.9319	-0.4128
neuron 2 hidden layer 2	0.0201	1.7212	1.5247	-1.4660	-0.4621	-0.0270	0.4298	-0.5790	0.0893	0.8031	0.0752	0.8253	-1.4759
neuron 3 hidden layer 2	0.9651	-0.8940	-1.0050	0.0385	-1.6841	1.8997	-0.5254	-0.4102	-0.3199	0.0607	-0.6200	0.1798	1.6975
neuron 4 hidden layer 2	1.8260	0.6595	-0.9506	1.8597	1.1396	-0.2439	-1.3342	-1.3975	-0.9032	-0.8346	-1.0574	-1.2660	-2.5429
neuron 5 hidden layer 2	0.5949	1.7231	-0.5852	0.0657	0.3123	-1.9626	-0.7525	-1.9482	-1.8696	-0.7689	0.1300	1.9383	2.1826
neuron 6 hidden layer 2	-0.3925	-1.1537	-0.9432	-1.6703	-0.3890	0.2057	-0.9653	-3.0086	0.6251	0.1586	2.8475	0.8918	0.4832
neuron 7 hidden layer 2	-0.5883	1.3248	-1.2861	-1.7827	0.0028	1.0025	-0.2644	0.6375	0.5214	-0.9543	-1.2244	-0.6217	-0.0302
neuron 8 hidden layer 2	-0.7554	0.5335	0.7869	-1.0239	0.5456	-0.0090	-1.8898	0.5210	0.4736	-1.0263	1.0238	-0.2633	-1.8151
neuron 9 hidden layer 2	1.2879	-0.4583	0.8576	-1.2115	-0.8274	1.6704	-1.1495	-1.6700	-0.4123	1.5048	1.3057	-1.6813	0.3444

(c) *Output weights*

	neuron 1 hidden layer 2	neuron 2 hidden layer 2	neuron 3 hidden layer 2	neuron 4 hidden layer 2	neuron 5 hidden layer 2	neuron 6 hidden layer 2	neuron 7 hidden layer 2	neuron 8 hidden layer 2	neuron 9 hidden layer 2
neuron 1 output layer	0.1008	-1.1415	-0.6691	1.9776	2.0020	1.6720	-0.4618	0.8470	0.8699

(d) *Biases*

neuron 1 input	52.1969	neuron 1 hidden layer 2	-2.7932	neuron 1 output layer	1.8808
neuron 2 input	-58.8266	neuron 2 hidden layer 2	1.1854		
neuron 3 input	-51.6920	neuron 3 hidden layer 2	-1.2530		
neuron 4 input	-64.0172	neuron 4 hidden layer 2	-0.7147		
neuron 5 input	-61.5817	neuron 5 hidden layer 2	-0.6157		
neuron 6 input	-12.3705	neuron 6 hidden layer 2	1.5002		
neuron 7 input	-7.3823	neuron 7 hidden layer 2	-1.3334		
neuron 8 input	-48.8509	neuron 8 hidden layer 2	-0.0725		
neuron 9 input	47.7173	neuron 9 hidden layer 2	1.1354		
neuron 10 input	-28.7819				
neuron 11 input	55.4569				
neuron 12 input	-67.4726				
neuron 13 input	11.5817				

Table A8: Input layer, hidden layers, and output layer weights and biases for ANN in prediction of FPF for all carriers

(a) Input weights

	neuron 1 input	neuron 2 input	neuron 3 input	neuron 4 input	neuron 5 input	neuron 6 input	neuron 7 input	neuron 8 input	neuron 9 input	neuron 10 input	neuron 11 input	neuron 12 input	neuron 13 input	neuron 14 input	neuron 15 input
neuron 1 hidden layer 1	- 0.1295	- 0.7201	- 1.5305	- 2.6029	- 2.9725	- 1.9574	- 1.6536	- 4.5417	- 2.9810	- 0.9273	- 7.7227	- 2.9252	- -1.2484	- 0.9985	- 5.0060
neuron 2 hidden layer 1	- 0.4985	- 1.0318	- 3.2740	- 4.3951	- 1.8833	- 2.5997	- 1.7526	- 5.2358	- 2.5891	- 3.1834	- -5.1163	- -1.0641	- -0.6535	- -2.2167	- -1.9036
neuron 3 hidden layer 1	- 0.1281	- 3.7257	- 0.5481	- 1.7591	- 2.2341	- 1.0051	- 2.7717	- 2.7255	- 5.5151	- -1.0242	- 4.6786	- 5.3812	- 1.5442	- 0.0089	- 5.7920
neuron 4 hidden layer 1	- 0.2800	- 0.1931	- 3.4117	- 1.7434	- 0.9077	- 2.2675	- 3.8087	- 2.9997	- 5.8002	- 2.3115	- -6.2255	- 0.7158	- -0.9385	- -3.3553	- -0.7117
neuron 5 hidden layer 1	- 2.1355	- 2.6679	- 0.8241	- 0.7016	- 1.3019	- 1.8885	- 0.5903	- 3.3692	- 4.9129	- 1.3439	- 10.0142	- -2.5248	- -3.2426	- -0.7256	- -4.6821
neuron 6 hidden layer 1	- 0.4333	- 0.6558	- 0.0599	- 1.8029	- 1.4644	- 0.1171	- 1.3893	- 0.3271	- 4.3225	- -2.2857	- -4.8033	- 0.0031	- 3.6744	- 0.2174	- 6.6022
neuron 7 hidden layer 1	- 1.3894	- 0.9145	- 2.5174	- 2.9587	- 0.2133	- 0.1589	- 1.2194	- 3.5186	- 0.5289	- 1.3303	- 3.6933	- -4.0326	- 2.4203	- 2.6550	- 5.0676
neuron 8 hidden layer 1	- 0.1498	- 4.2932	- 2.6537	- 1.4452	- 0.8807	- 0.5185	- 2.0813	- 4.5405	- 1.9479	- -1.4567	- 2.0749	- 4.0253	- -2.1283	- 2.3621	- -6.8100
neuron 9 hidden layer 1	- 0.0447	- 1.6986	- 0.9946	- 2.9019	- 0.1162	- 2.2719	- 1.1429	- 2.9693	- 2.1407	- -3.4126	- 0.9415	- 0.3937	- -2.5787	- 1.6583	- -8.4454
neuron 10 hidden layer 1	- 1.5749	- 2.9526	- 1.4826	- 1.5160	- 0.4546	- 3.4295	- 2.4559	- 5.8612	- 0.5824	- 1.0500	- -3.0445	- -3.9289	- -5.6034	- -1.5729	- 2.0514

(b) Hidden layers weights

	neuron 1 hidden layer 1	neuron 2 hidden layer 1	neuron 3 hidden layer 1	neuron 4 hidden layer 1	neuron 5 hidden layer 1	neuron 6 hidden layer 1	neuron 7 hidden layer 1	neuron 8 hidden layer 1	neuron 9 hidden layer 1	neuron 10 hidden layer 1
neuron 1 hidden layer 2	0.9780	-0.7371	0.7134	-0.5136	-0.2021	1.8554	-0.1403	1.0460	1.0651	-0.1270
neuron 2 hidden layer 2	-0.6906	0.4060	1.7023	0.2486	1.6635	-1.9048	-1.1992	-0.3783	0.7626	-0.6934
neuron 3 hidden layer 2	0.0512	-1.6679	-0.3319	-0.9632	-2.5817	-1.1195	-0.4689	-0.0161	1.3842	0.6875
neuron 4 hidden layer 2	0.1674	0.0629	-0.1014	-0.1448	0.3814	0.1905	-0.9752	-1.0591	0.6174	-1.6114
neuron 5 hidden layer 2	0.2609	0.3345	-1.1606	-1.0039	-1.1585	0.2991	-0.5645	-0.7975	1.6153	-2.2110
neuron 6 hidden layer 2	0.4169	0.0154	0.2520	0.5615	1.3896	0.2231	0.7338	-1.1247	2.0557	2.0650
neuron 7 hidden layer 2	-1.5440	0.2780	-1.6963	-0.6024	2.1633	1.4336	0.5684	-0.8109	-0.1855	1.9365
neuron 8 hidden layer 2	-0.4356	-0.2097	1.8676	-1.5757	-1.1267	0.3812	0.6608	-1.2760	0.5236	0.6057
neuron 9 hidden layer 2	2.0544	-0.1487	0.7704	0.7071	0.1484	-1.4467	1.1138	-0.3950	0.4319	-1.2883

(c) *Output weights*

	neuron 1 hidden layer 2	neuron 2 hidden layer 2	neuron 3 hidden layer 2	neuron 4 hidden layer 2	neuron 5 hidden layer 2	neuron 6 hidden layer 2	neuron 7 hidden layer 2	neuron 8 hidden layer 2	neuron 9 hidden layer 2
neuron 1 output layer	-1.1368	-0.7901	0.0159	-0.0611	0.4013	0.1662	0.2776	0.3418	-0.0212

(d) *Biases*

neuron 1 input	22.7425	neuron 1 hidden layer 2	1.6561	neuron 1 output layer	-0.6342
neuron 2 input	-4.3155	neuron 2 hidden layer 2	-0.5000		
neuron 3 input	-27.8391	neuron 3 hidden layer 2	0.7425		
neuron 4 input	-4.6380	neuron 4 hidden layer 2	-0.7305		
neuron 5 input	-7.9502	neuron 5 hidden layer 2	-1.4642		
neuron 6 input	-10.8011	neuron 6 hidden layer 2	1.0272		
neuron 7 input	-27.7253	neuron 7 hidden layer 2	-0.9362		
neuron 8 input	35.1591	neuron 8 hidden layer 2	-1.9644		
neuron 9 input	-25.8650	neuron 9 hidden layer 2	0.2082		
neuron 10 input	-7.2365				