



TAMPEREEN TEKNILLINEN YLIOPISTO  
TAMPERE UNIVERSITY OF TECHNOLOGY

HESAM JAFARIAN

**Process Modeling Optimization in Additive Manufacturing Using Artificial Neural Networks**

Master of Science Thesis

## ABSTRACT

**Hesam Jafarian:** Process Modeling Optimization in Additive Manufacturing Using Artificial Intelligence

Tampere University of technology  
Master of Science Thesis, 117 pages  
November 2018

Master's Degree Program in Automation Engineering  
**Major:** Factory Automation and Industrial Informatics  
**Examiner:** Professor Eric Coatanea

**Keywords:** Artificial Intelligence- Artificial Neural Network-Wire Arc Additive Manufacturing

The need for production has roots in human life and its history. This date back to primitive days of human life, where he or she had to apply surrounding materials in order to manufacture the tools necessary for survival and durability against any insecurity. This was legitimizing the use of any means in order to obtain the tools and reach the goals at any cost. However, with human development primarily within the knowledge and understanding domain and also with the desire of humanity for best, expectations have risen. This was the time not only the cost mattered but also the simplicity of design, massive production, and diversity, less waste, autonomy, and implementation within a shorter time gained a higher momentum.

On the other hand, the conventional manufacturing method was based on subtractive manufacturing with cutting and eliminating the unwanted sections or parts of an object. The disadvantage of such a method is that it requires a complicated production process design and is accompanied by waste. However, with the rise of additive manufacturing and three-dimensional printing equipment back in the 1980s, it became possible to build parts which could have almost any shape or geometry. Moreover, this also empowered the possibility of using digital and 3D models built by computer-aided design software.

Simultaneously, on the other side, the foundation and application of artificial intelligence were maturing. This was due to the demand for machines to assist human beings in the domain of knowledge reasoning, learning, and planning. These were the pillars for making machines autonomous and to benefit from such features.

Accordingly, this research work studies and overviews the applications and techniques of machine learning and artificial intelligence in the domain of additive manufacturing. It aims to determine the interaction of influential parameters on the process and to find the best solutions for improving the quality and mechanical features of manufactured parts.

Moreover, this research tends to enable the experts to grasp a better understanding of AM process during manufacturing and additionally intends to infuse the experts' knowledge in additive manufacturing field utilizing the artificial neural network and finally generate a model with the ability of prediction and selection for promising results.

## PREFACE

که گردد بدو مرد جوینده مه	به او گفت شاه، از هنرها چه به
که داننده بر مهتران بر، مه است	چنین پاسخ داد که دانش به است
تن مرده و جان نادان یکیست!	ز دانش به اندر جهان هیچ نیست!

-فردوسی کبیر

“Once questioned a king about the best artifice one can achieve, which with it the seeker can become extreme. A sagacious man responded to the king the Knowledge indeed it is. The owner of the knowledge to see or if to deem will become above all without to dream. Nothing would be better than knowledge in the world, where a corpse or alive without knowledge will be equal in accord.”  
–The great Ferdowsi

This thesis is conducted within the consent of Mechanical Engineering and Industrial Systems Faculty at the Tampere University of Technology in the years 2017-2018 and under the supervision of Professor Eric Coatanea.

I at this moment would like to express my utmost gratitude to all who supported me spiritually during the hard time I experienced with the loss of everything, merely the word ‘Father’ who was, is and will be an inspiration forever with his smiles and hopes for future.

I also like to express my highest appreciation again for Professor Coatanea who with all of his supports indeed played a turning point in my life. I want to appreciate the efforts of my family and dearest friends especially Romaric Prodhon, Marzieh Zare, Alireza Zare ,Saboktakin Hayati, Amir-homayoun Aliyazdi for their non-stopping supports.

Additionally, a significant amount of appreciation goes for our research group in the MEI-LAB.

Finally, I hope this thesis provides some but not all the means to pave the path for future development in the additive manufacturing field and envision the possible future works.

Tampere, 20.11. 2018  
Hesam Jafarian

## TABLE OF CONTENTS

ABSTRACT.....	I
PREFACE.....	II
TABLE OF CONTENTS.....	2
LIST OF FIGURES .....	5
NOMENCLATURE .....	8
INTRODUCTION .....	9
Objective and goal.....	9
Problem re-statement.....	9
Goals and Objectives.....	10
Study Limitations .....	10
THESIS STRUCTURE.....	11
DESIGN AND ANALYSIS ARCHITECTURE .....	12
ADDITIVE MANUFACTURING .....	13
History .....	13
Digital workflow .....	14
Design Manufacturing and Assembly .....	14
Design Requirements.....	15
Design and Opportunities .....	15
Constraints.....	15
Costs 15	
FUSED DEPOSITION ADDITIVE MANUFACTURING .....	16
WIRE-ARC ADDITIVE MANUFACTURING (WAAM).....	19
Welding Techniques.....	19
GMAW welding .....	19
MIG / MAG Welding.....	19
TIG welding .....	19
CMT Welding.....	19
CMT Pulse.....	20
CMT Advanced .....	20
CMT Advanced and Pulse.....	20
Shielding Gas Types.....	21
ACETYLENE (C <sub>2</sub> H <sub>2</sub> ) .....	21
ARGON (Ar).....	21
ARGON / CARBON DIOXIDE (Ar / CO <sub>2</sub> ) .....	21
CARBON DIOXIDE (CO <sub>2</sub> ) .....	21
NITROGEN (N <sub>2</sub> ) .....	21
OXYGEN (O <sub>2</sub> ).....	21
TRI-MIX (Ar / O <sub>2</sub> / CO <sub>2</sub> ) .....	21
Wire Feeds or Filler.....	22

WELDING DEFECTS.....	23
Cracks .....	24
Undercuts.....	24
Spatters .....	25
Porosity.....	25
Overlap defect .....	26
Crater & Slag.....	26
Fusion incomplection .....	27
Necklace Crack.....	27
Grooves and Penetration.....	27
ENVIRONMENT .....	28
ABB robot manipulator .....	29
IRBP A-750 Positioner.....	30
ABB teach-pendant .....	30
Cold Metal Transfer System.....	32
Torch.....	33
METHODOLGIES AND STATE OF THE ART .....	36
Dimensional Analysis Methodology .....	37
Computational Methodologies .....	38
Analysis of Variance (ANOVA) Methodology.....	38
Machine Learning Methodology .....	39
Supervised learning (learning by a teacher) .....	42
Non-supervised learning (learning without teacher) .....	42
Reinforcement learning .....	43
Learning rules.....	43
Backpropagation (BP) network .....	44
General issues in ANN development .....	44
Classical approximation using Surface Response Models .....	47
Knowledge-Based Artificial Intelligence System Design.....	48
Fuzzy Thinking, Neuro Fuzzy Rules and Inference .....	49
Fuzzification and Rules Generation .....	50
Hedge Categorization and Linguistic Variables.....	52
Ruled-Based Systems .....	54
Mamdani Inference Systems .....	55
Constraints.....	55
Adaptive Neural Network.....	56
RESULTS AND DISCUSSIONS.....	57
Case study 1: Training of a data-set using artificial neural networks .....	57
Case study 2: Sphere and Fluid a Knowledge Based Application .....	67
Case study 3: Fused Deposition Modeling.....	74
Case study 4: Bead segmentation using clustering approach.....	83

Case study 5: Radial-Based Neural Networks.....	90
Case study 6: NeuroFuzzy.....	94
CONCLUSIONS AND PROSPECTIVE WORKS.....	109
Conclusion.....	109
Prospective works.....	110
REFERENCES .....	111

## LIST OF FIGURES

FIGURE 1: ILLUSTRATION OF THE THESIS STRUCTURE. ....	11
FIGURE 2: ILLUSTRATION OF THE MAIN ARCHITECTURE DESIGN. ....	12
FIGURE 3: ILLUSTRATION OF TISSUE PRODUCTION BY MEANS OF AM PROCESS [7]. ....	13
FIGURE 4: ILLUSTRATION OF AN ARTIFICIAL HIP GENERATED USING AN AM PROCESS [9]. ....	13
<i>FIGURE 5: ILLUSTRATION OF A GAS TURBINE BLADE PRODUCED BY ADDITIVE MANUFACTURING.</i> ....	14
FIGURE 6: ILLUSTRATION OF A FUSED DEPOSITION MODELING PROCESS. ....	17
FIGURE 7: ILLUSTRATION OF THE PRINTING STEPS. ....	17
FIGURE 8 : ILLUSTRATION OF A G-CODE BLOCK WITH CORRESPONDING FUNCTIONS. ....	18
FIGURE 9: ILLUSTRATION OF PRE-PROCESSING WORKFLOW. ....	18
FIGURE 10: ILLUSTRATION OF CMT IN PULSE MODE [7]. ....	20
FIGURE 11: ILLUSTRATION OF CURRENT AND VOLTAGE DIAGRAM OF CMT IN DIFFERENT MODES [30]. ....	20
FIGURE 12: ILLUSTRATION OF LIST OF SHIELDING GASES. ....	21
FIGURE 13: ILLUSTRATION OF CATEGORIZATION OF THE WELDING DEFECTS. ....	23
FIGURE 14: ILLUSTRATION OF A PROPER WELDING. ....	23
FIGURE 15: ILLUSTRATION OF THE CRACK DEFECT. ....	24
FIGURE 16: ILLUSTRATION OF THE UNDERCUT DEFECT. ....	24
FIGURE 17: ILLUSTRATION OF WELDING SPATTERS AROUND A WELDING BEAD. ....	25
FIGURE 18: ILLUSTRATION OF POROSITY DEFECT IN WELDING. ....	25
FIGURE 19: ILLUSTRATION OF THE OVERLAPPING DEFECT IN WELDING PROCESS. ....	26
FIGURE 20: ILLUSTRATION OF THE SLAG DEFECT IN WELDING. ....	26
FIGURE 21: ILLUSTRATION OF AN INCOMPLETE FUSION DEFECT. ....	27
FIGURE 22 ILLUSTRATION OF THE GROOVED PENETRATION. ....	27
FIGURE 23: ILLUSTRATION OF THE EQUIPMENT OF ADDITIVE MANUFACTURING USING WAAM TECHNIQUE. ....	28
FIGURE 24: ILLUSTRATION OF THE WORKING RANGE OF THE ROBOT. ....	29
FIGURE 25: ILLUSTRATION OF DIMENSION OF ABB IRB 4600-40-2.55 MANIPULATOR. ....	30
FIGURE 26: ILLUSTRATION OF THE IRBP A-750 POSITIONER. ....	30
FIGURE 27: ILLUSTRATION OF OPERATION MODES OF THE ROBOT. ....	31
FIGURE 28: ILLUSTRATION OF AN ABB FLEX PENDANT AND ITS VIRTUAL FORMAT IN ROBOT STUDIO. ....	31
FIGURE 29: ILLUSTRATION OF THE WELDING SYSTEMS ARCHITECTURE. ....	32
FIGURE 30: ILLUSTRATION OF THE FRONIUS WELDING TORCH. ....	33
FIGURE 31: ILLUSTRATION OF DIFFERENT TORCH NECKS WITH CORRESPONDING ANGLES. ....	33
FIGURE 32: ILLUSTRATION OF THE TORCH NECK WEAR PARTS. ....	34
FIGURE 33: ILLUSTRATION OF ATTACHMENT OF ROBACTA DRIVE AND 5000 TORCH BOD. ....	34
FIGURE 34: ROBACTA DRIVE CMT. ....	34
FIGURE 35: ROBACTA 5000 WITH A TOUCH SENSOR. ....	34
FIGURE 36: ILLUSTRATION OF THE WELDING UNIT. ....	35
FIGURE 37: ELEMENTS OF REINFORCEMENT LEARNING. ....	43
FIGURE 38: ILLUSTRATION OF THE FUZZY AND BOOLEAN LOGIC. ....	49
FIGURE 39: ILLUSTRATION OF A UNIVERSAL SET X AND ITS SUBSETS AS CRISP AND FUZZY SUBSETS. ....	51
FIGURE 40: ILLUSTRATION OF THE HEDGES, CORRESPONDING MATHEMATICAL RELATIONS AND GRAPHICAL REPRESENTATION [100]. ....	53
FIGURE 41: ILLUSTRATION OF MONOTONIC RELATION OF THE HEIGHT AND WEIGHT. ....	54
FIGURE 42: ILLUSTRATION OF THE ARCHITECTURE OF A NEURAL NETWORK WITH TWO HIDDEN LAYERS. ....	58
FIGURE 43: ILLUSTRATION OF NETWORK ARCHITECTURE. ....	60
FIGURE 44: ILLUSTRATION OF A SIGMOIDAL AND A PURELIN TRANSFER FUNCTION. ....	60

FIGURE 45: ILLUSTRATION OF THE ANN PREDICTION OUTPUT AND THE EXPECTED TARGET VALUES FOR 120 DATA POINTS. .... 61

FIGURE 46: ILLUSTRATION OF THE PERFORMANCE CURVE FOR THE 120 DATA SAMPLES. .... 61

FIGURE 47: ILLUSTRATION OF THE PROGRESS PARAMETERS FOR THE 120 DATA SAMPLES..... 62

FIGURE 48. ILLUSTRATION OF TRAINING OF 2601DATA SAMPLES. .... 62

FIGURE 49: ILLUSTRATION OF PROGRESS DIAGRAM FOR 2601 SAMPLES. .... 63

FIGURE 50: ILLUSTRATION OF RESULTS FOR THE SECOND DATA SET ..... 63

FIGURE 51: ILLUSTRATION OF THE RESULTS FOR THE TWO DATA-SETS WITH PRESENCE OF A NOISE. .... 64

FIGURE 52: ILLUSTRATION OF A TRAINING WITH HIGH AMOUNT OF HIDDEN NODES AND LAYERS. .... 64

FIGURE 53: ILLUSTRATION OF THE NETWORK RESULTS WITH HIGH AMOUNT OF NEURONS AND HIDDEN LAYERS..... 65

*FIGURE 54: ILLUSTRATION OF THE SURFACE RESPONSE MODEL USING A LINEAR LOWESS WITH 1% SPAN AND NO ROBUST. .... 65*

FIGURE 55: ILLUSTRATION OF SURFACE RESPONSE MODELING ON THE DATA SETS. .... 66

FIGURE 56: ILLUSTRATION OF THE SPHERE IN THE FLOW FLUID WITH DEMONSTRATION OF THE DRAG FORCE. .... 67

FIGURE 57: ILLUSTRATION OF A KBANN NETWORK CONSTRUCTED USING DIMENSIONAL ANALYSIS..... 68

FIGURE 58: ILLUSTRATION OF PREVIOUS ATTEMPT TO PREDICT THE MODEL WITH MODELS SENSITIVITY TO PRIMARY CONDITIONS. .... 68

FIGURE 59: ILLUSTRATION OF THE RESULTS OF PREDICTION BY KB-ANN. .... 69

FIGURE 60: DIFFERENCE OF THE REAL DATA POINTS (TARGET VALUES) AND THE PREDICTED DATA BY THE NETWORK (OUTPUT) VALUES FOR THE KB-ANN..... 69

FIGURE 61: ILLUSTRATION OF PERFORMANCE OF TRAINING PROCESS FOR TRAINING, VALIDATION AND TEST DATA POINTS..... 69

FIGURE 62: ILLUSTRATION OF PERFORMANCE CURVES FOR THE KB-ANN NETWORK. .... 70

FIGURE 63: ILLUSTRATION OF CONVENTIONAL ANN STRUCTURE MODEL FOR THIS PROBLEM. .... 71

FIGURE 64: ILLUSTRATION OF THE RESIDUALS AND THE PREDICTION RESULTS OF THE CONVENTIONAL ANN. .... 71

FIGURE 65: ILLUSTRATION OF REGRESSION PLOT FOR THE CONVENTIONAL NETWORK. .... 72

FIGURE 66: ILLUSTRATION OF THE PERFORMANCE CURVES FOR THE CONVENTIONAL NEURAL NETWORKS..... 72

FIGURE 67: ILLUSTRATION OF THE RESULTS FOR BOTH NETWORKS AND THE COMPARISON OF THE CONVENTIONAL (Up) AND KB-ANN (Down) DATA PLOTS. .... 73

FIGURE 68: ILLUSTRATION OF THE FDM CASE USED FOR THE STUDY. .... 74

FIGURE 69: ILLUSTRATION OF THE CAUSAL RELATIONS IN THE FDM PROCESS FOR THE SELECTED CASE STUDY OBTAINED BY THE DACM. .... 75

FIGURE 70: ILLUSTRATION OF THE LIST OF IMPACTING PARAMETERS AFFECTING THE FDM PROCESS IN THIS CASE STUDY. .... 76

FIGURE 71: ILLUSTRATION OF CLASSICAL ANN MODEL USED FOR FDM PROCESS. .... 76

FIGURE 72: ILLUSTRATIONS OF THE PERFORMANCE CURVES FOR THE CLASSICAL TRAINING OF THICKNESS..... 77

FIGURE 73: ILLUSTRATION OF THE PERFORMANCE RESULT FOR THE CLASSICAL ANN FOR THE HEIGHT. .... 78

FIGURE 74: ILLUSTRATION OF THE BEST PERFORMANCE FOR MASS..... 78

FIGURE 75: ILLUSTRATION OF THE PERFORMANCE CURVES WITH THE BEST VALIDATION FOR THE ANN1 IN KB-ANN NETWORK. .... 79

FIGURE 76: ILLUSTRATION OF THE ANN2 PERFORMANCE DIAGRAM RESULTS IN THE KB-ANN NETWORK. .... 80

*FIGURE 77: ILLUSTRATION OF THE ANN3 PERFORMANCE DIAGRAM RESULTS IN THE KB-ANN NETWORK. .... 80*

FIGURE 78: ILLUSTRATION OF THE ANN4 PERFORMANCE CURVES IN KB-ANN NETWORK. .... 81

FIGURE 79: ILLUSTRATION OF GENERALIZATION TEST AND THE AMOUNT OF STANDARD ERROR FOR CLASSICAL AND KB-ANN NETWORKS. . 81

FIGURE 80: ILLUSTRATION OF A TURBINE BLADE OF A GAS COMPRESSOR. .... 83

FIGURE 81: ILLUSTRATION OF A CROSS SECTION OR SLICE SURFACE OF THE CAD MODEL. .... 83

FIGURE 82: ILLUSTRATION OF BEAD WITH DIFFERENT QUALITIES..... 84

*FIGURE 83: ILLUSTRATION OF THE PARAMETERS USED IN THE WELDING PROCESS. .... 84*

FIGURE 84: ILLUSTRATION OF THE BEADS IN POLYWORKS AFTER BEING SCANNED USING THE LASER SCANNER..... 86

FIGURE 85: ILLUSTRATION OF A CROSS SECTION CONDUCTED TO PROVIDE THE ANALYTICAL DATA..... 87

FIGURE 86: ILLUSTRATION OF A SCANNED BEAD WITH THREE CORRESPONDING ZONES. .... 87

FIGURE 87: ILLUSTRATION OF BEAD 1 AND 2 HEIGHT AND WIDTH DATA..... 88

FIGURE 88: ILLUSTRATION OF THE K-MEDOID CLUSTERING TECHNIQUE FOR ZONE SEGMENTATION FOR A WAAM PROCESS..... 88



FIGURE 89: ILLUSTRATION OF K-MEDOID CLUSTERING APPROACH FOR BEAD SEGMENTATION IN A WAAM PROCESS. ....	89
FIGURE 90: ILLUSTRATION OF THE TYPICAL RBF NETWORK TOPOLOGY WITH RADIAL FUNCTION IN THE HIDDEN LAYERS. ....	90
FIGURE 91: ILLUSTRATION OF THE RESULT OF RBF TRAINING WITH LEARNING RATE OF 0.0003 AND 5 CLUSTER CENTERS.....	91
FIGURE 92: ILLUSTRATION OF THE RBF NETWORK RESULTS WITH LEARNING RATE OF THE 0.00003 AND 4 CLUSTERS.....	92
FIGURE 93: ILLUSTRATION OF RESULTS OF THE RBF NETWORK FOR THE LEARNING RATE OF 0.00008 AND 4 NUMBER OF CENTERS.....	92
FIGURE 94: ILLUSTRATION OF AN APPROXIMATE USING THE RBF NETWORKS FOR A KNOWN $\sin(x)$ FUNCTION PLUS A RANDOM WHITE NOISE. THE BEST FIT IS OBTAINED FOR A SIX CLUSTER CENTERS AND WITH A LEARNING RATE OF 0.03 .....	93
FIGURE 95: ILLUSTRATION OF A POTENTIAL CAUSAL RELATION IN THE WAAM PROCESS.....	94
FIGURE 96: DEFINITION OF THE CAUSAL GRAPH.....	95
FIGURE 97: ILLUSTRATION OF THE INTERACTION AMONG THE PARAMETERS ON THE MEAN OF AVERAGE HEIGHT.....	96
FIGURE 98: ILLUSTRATION OF THE VARIABLES AND THEIR CORRESPONDING HEIGHT VALUES.....	98
FIGURE 99: ILLUSTRATION OF THE MAIN EFFECT PLOT FOR THE OUTPUT HEIGHT AND THE OTHER MAIN INPUT VARIABLES.....	98
FIGURE 100: INTERACTION OF THE INPUT VARIABLES IN THE WAAM PROCESS AND THE HEIGHT VARIABLE AS THE OUTPUT.....	99
FIGURE 101: ILLUSTRATION OF THE LIST OF FACTORS INFORMATION AFTER NORMALIZATION FOR THE HEIGHT OUTPUT VALUE.....	100
FIGURE 102: ILLUSTRATION OF THE RESULTS OF A TWO WAY ANOVA ON THE WAAM FOR ESTIMATION OF THE HEIGHT.....	100
FIGURE 103: ILLUSTRATION OF THE CORRELATION BETWEEN INPUT VARIABLES AND THE OUTPUT HEIGHT.....	101
FIGURE 104: BEAD SEGMENTATION TO SEPARATE THE BEAD ZONES.....	101
FIGURE 105: ILLUSTRATION OF THE FUZZY INFERENCE ENGINE WITH 3 INPUTS AND 1 OUTPUT VARIABLE.....	101
FIGURE 106: TS VS HEIGHT VALUES (TOP) AND MEMBERSHIP FUNCTION FOR THE TS (DOWN).....	102
FIGURE 107: WFR (M/MIN) VS HEIGHT AND CORRESPONDING WFR'S RANGE (TOP) AND ILLUSTRATION OF THE MEMBERSHIP FUNCTIONS FOR THE WFR.....	103
FIGURE 108: A ( $^{\circ}$ ) VS HEIGHT EFFECT PLOT AND THE RANGE OF VALUES.....	104
FIGURE 109: ILLUSTRATION OF THE MEMBERSHIP FUNCTIONS FOR THE TORCH ANGLE.....	104
FIGURE 110: ILLUSTRATION OF THE RANGE OF HEIGHT'S VALUES FOR 55 SAMPLES.....	105
FIGURE 111: ILLUSTRATION OF THE MEMBERSHIP FUNCTIONS FOR THE HEIGHT PARAMETER.....	105
FIGURE 112: ILLUSTRATION OF THE RULES DEFINITION IN FUZZY TOOLBOX IN MATLAB.....	106
FIGURE 113: ILLUSTRATION OF THE FUZZY INFERENCE WITH THE CORRESPONDING RULES DEFINE FOR ESTIMATION OF THE HEIGHT.....	107
FIGURE 114: ILLUSTRATION OF COMPARISON OF THE PREDICTED VALUES AND ACTUAL VALUES FOR THE TEST DATA ON FOR THE WAAM.....	107

## LIST OF TABLES

TABLE 1: AXIS INFORMATION OF THE ROBOT AND CORRESPONDING SPEEDS.....	29
TABLE 2: TALLNESS MEMBERSHIP COMPARISON OF THE BOOLEAN AND FUZZY REASONING BASED ON THE 1.85 CM CRITERIA.....	50
TABLE 3: THE COMPARISON OF THE TWO NETWORKS IN PRESENCE AND ABSENCE OF THE NOISE.....	64
TABLE 4: LIST OF PARAMETERS INVOLVED IN WAAM PROCESS.....	95
TABLE 5: TABLE OF PREDICTION RESULTS.....	108

## NOMENCLATURE

AM	Additive Manufacturing
ANN	Artificial Neural Network
ANOVA	Analysis of Variance
ASTM	American Society for Test and Material
BL	Backpropagation Learning
BL	Boltzmann Learning
BNN	Bayesian Neural Network
BP	Back Propagation
CAD	Computer Aiding Designs
CL	Competitive Learning
CMT	Cold Metal Transfer
DA	Dimensional Analysis
ECL	Error Correction Learning
ECL	Error-correction learning
FDM	Fused Deposition Modeling
FNN	Fuzzy-Neural-Networks
GA	Genetic Algorithm
G-Code	Geometric code
HL	Hebbian Learning
KB-ANN	Knowledge Based Artificial Intelligence
LBM	Laser Beam Melting
LM	Levenberg-Marquardt
LSTM	Long Short-Term Memory
ML	Machine Learning
MLP	Multi-Layer Perceptron
MSE	Mean Squared Error
NHN	Number of Hidden Nodes
RBF	Radial Basis Function
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
RSM	Response Surface Methodology
SSE	Sum-of-Squared Error
WAAM	Wire Arc Additive Manufacturing

## INTRODUCTION

Although additive manufacturing (shortly as AM) dates back to 1980s, it has received much attention recently. One reason is that it is considered as a quick prototyping solution for many scientist and researcher. Besides, many companies find it as a mean for a better rapid prototyping and production planning in their research and development procedures. Consequently, prominence of the most influential parameters in an AM processes have become more evident and turned to a trend of demand. Therefore, depending on the type of AM process and based on the method applied, concepts such as diffusion, melting, porosity, orientation of fillers, types of materials, mechanical features, heat transfer and fluids behavior are among the most discussed topics and a lot of research and studies are conducted accordingly at the moment to address these topics [1].

### Objective and goal

With correspondence to the first illustration portrayed in the introduction section and based on the demand mentioned at the moment by the companies, this research intends to provide a study of AM processes with the objectives demanded in this field. Therefore, following this section, the problem in the field of AM is re-stated based on recent studies and then an illustration of the objectives and dealing with limitations are shown consequently.

### Problem re-statement

Based on the Testing and Materials American Society (Shortly ASTM) and within the designation of the F42, the AM process is categorized into the following categories and standards.

- Material Extrusion
- Power Bed Fusion
- Material Jetting
- Binder Jetting
- Direct Energy Deposition
- Vat Photo Polymerization
- Sheet Lamination

Nevertheless, among these AM processes, Fused Deposition Modelling hereafter called FDM and Wire Arc Additive Manufacturing shortly as (WAAM) would be the focus of this study [2] [3]. Meanwhile, in order to conduct a novel study and to obtain a physical model, it is required to reach an understanding of the controlling parameters and properties, which play a chief role for instance in the amount of surface roughness, fatigue properties, porosity and so forth.

Also, based on the type of the technology which is applied in the AM process, it is required to know the limitations and challenges within that concept. The main areas where special attention is required are [4]:

- AM filament and materials
- Type of AM technology and process
- AM equipment items
- Qualification of the productions
- Process controls possibilities such as real-time and closed-loop process controls
- Simulation and modeling
- Sensors such as their application in melt-pool measurements and temperatures
- Consistency
- Quality
- Reproduction

## Goals and Objectives

In order to reach a novel design concerning quality and strength it is required as an objective to illustrate and simplify involving parameters into the most useful parameters. This is realized by applying the concept presented as the dimensional analysis hereafter called as DA [9]. This method is used to identify the involving parameter and reduce the redundant relations. Additionally, it provides a graphical representation of existing relations in the AM process which can be applied for further and better analysis. The objective of this thesis is to study the appropriate machine learning approaches by utilizing DA in order to achieve a novel design regarding quality and structure. This is due to the reason that the additive manufacturing is subject to non-linearity and requires parameter adjustment. One another issue is that it is subject to uncertainty and this makes obtaining a process model quite tricky through the application of traditional approaches.

## Study Limitations

Although the study aims to present a novel model, however, it is important to notice that there are complexity and roughness which exists within additive manufacturing processes. This is perhaps due to the constraints imposed by the manufacturer devices and within the context of their design. Accordingly and based on a much dimension of parameters; resources must be adopted and allocated intelligently. One another essential issue mentioned is the amount of uncertainty which exists in the additive manufacturing processes. For instance, the uncertainty of the heating transfer patterns and precise interaction of the parameters. Also, the data collection within this context where there are limited sensors as the sources of data and also measurement techniques set a vast amount of constraints on the studies.

## THESIS STRUCTURE

As illustrated in the following figure and based on the structure a brief review of the background, history, constraints, and definition of the additive manufacturing alongside state of the art is represented. Moreover, appropriate machine learning techniques are reviewed in order to illustrate the possibilities, advantageous and disadvantageous of this practice in the context of the AM processes. Afterward, some of the application and the cases would be demonstrated accordingly and a final conclusion is presented at the end.

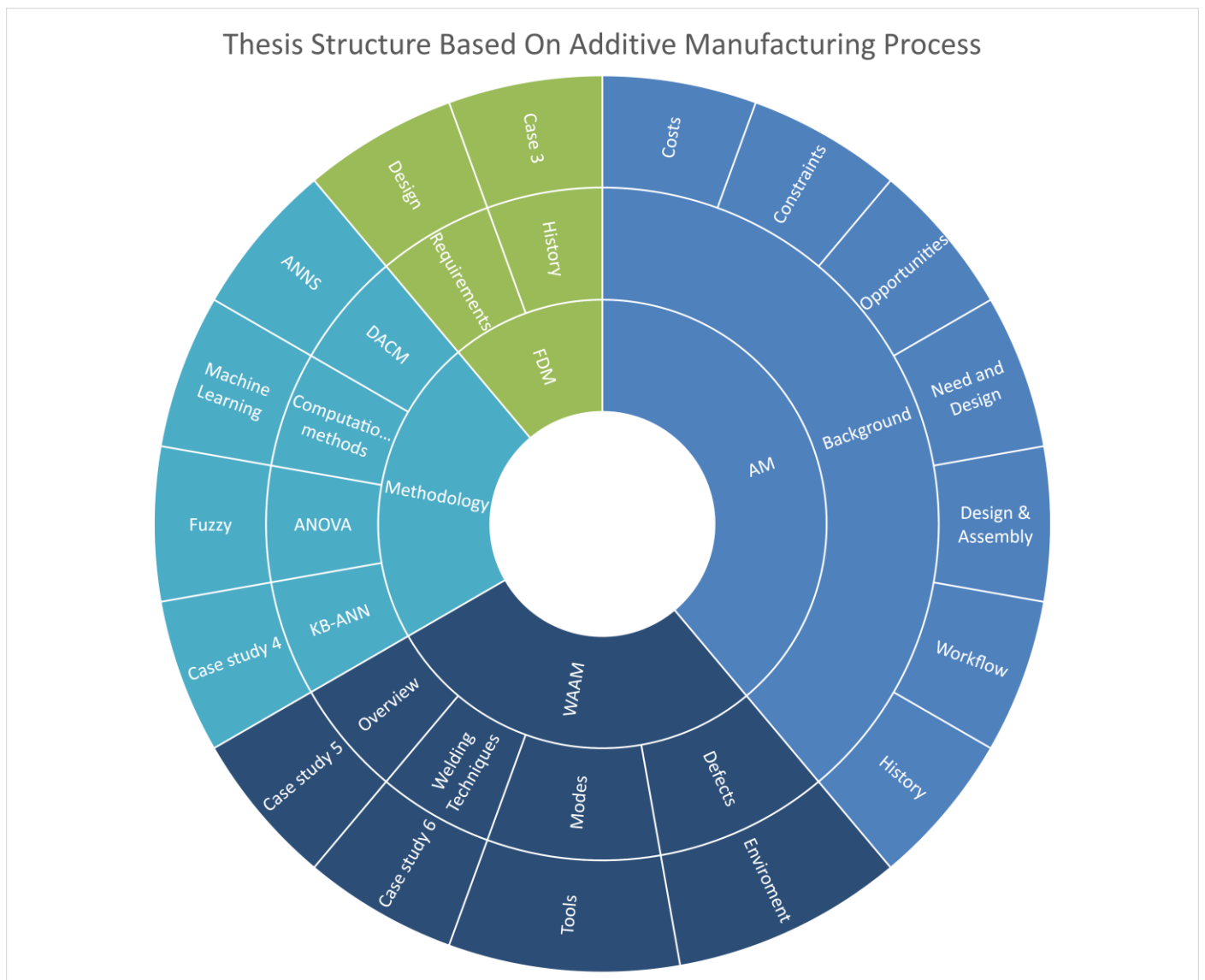


Figure 1: Illustration of the thesis structure.

## DESIGN AND ANALYSIS ARCHITECTURE

The design architecture is constructed based on the main elements of design. This includes the type of the process, the practice to obtain the most influential parameters and defining the dimensionality of the process (dimension analysis and reduction), design of the experiments in order to collect data, implementation of the experiments, measurements and data processing and finally applying the machine learning algorithm with the goal of estimation and predictions of parameters. All of these form the essential means to implement a novel design.

FDM and the WAAM processes are selected in this case for the design process. Next, it is required to use analytical methods to detect the most influential parameters for analysis. With the parameters at hand, a design of the experiment is nominated to find the interactions among parameters and especially on the desired outputs. Subsequently, with the experiments implementation, outputs are measured using for instance scanners and other measurements devices. Finally, the generated output data will be used for training artificial neural networks to analyze the data and the impact of the parameters.

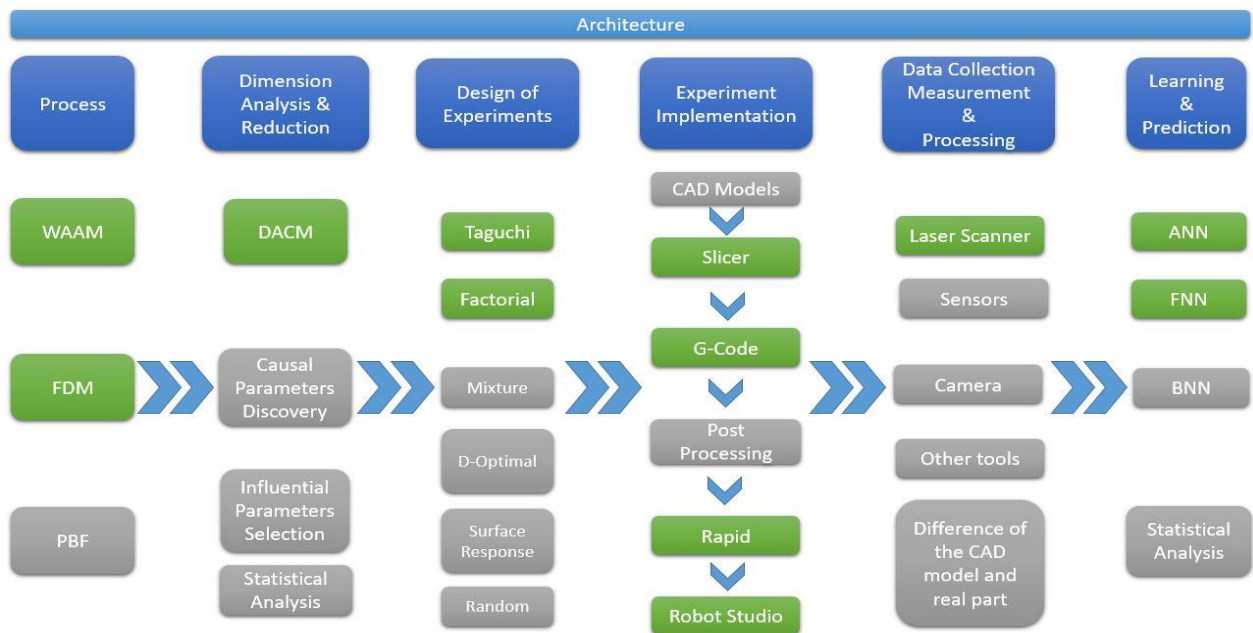


Figure 2: Illustration of the main architecture design.

## ADDITIVE MANUFACTURING

Additive manufacturing is a process where the desired part is built by deposition or an incremental layer-by-layer production procedure. Most of AM technologies use either wire or powder as the source. Then the source is deposited or melted by a focused heating process, and then it will form a part with solidification [5]. This part or object can be made of plastic, a metal material such as steel, aluminum, and titanium or any other material (body tissue)[6][7][8][9]. The advantage that the AM holds is the unique design and short lead-time in comparison to subtractive manufacturing methods [10]. As it is stated, the AM inclines to maneuver from rapid proto-typing to applications with rapid manufacturing. However, there is a necessity for knowledge acquisition in such processes. Especially in micro-structures that are made as a consequence of optimized process parameters and manufactured parts [11].



Figure 4: Illustration of an artificial hip generated using an AM process [9].

Figure 3: Illustration of tissue production by means of AM process [7].

### History

The AM history dates back to 150 years ago. This was done by building topographical maps and sculptures from 2D layers. Later that in 1960 and 1970 some proof of concept has been presented as a use of photopolymerization, powder fusion or sheet lamination in 1979 [12].

In 1989, MIT presented a patent of a 3D printing and later that in 1990 the laser beam melting (LBM) was introduced successfully [13][14]. On the other hand, the FDM along-side of the stereolithography, solid ground curing and laminated object manufacturing were introduced in 1991[15]. Meanwhile, it is worth to refer to this fact that the AM advancement was very dependent upon the development of the programmable logic controllers which later on with availability of the computers (such as CAD and CAM) systems, substantial means paved the way to help the improvement of this process.

<sup>1</sup> Courtesy of Instrumentaria [9]

<sup>2</sup> Courtesy EOS GmbH [7]

In the 2000s, newly commercialized technologies fueled the growth of the AM as with the introduction of the electron beam melting (EBM) [16]. This development was affected by file formats AM processes, such as STL, LEAF, and LMI where they stand for Stereo Lithography, Layer Manufacturing Interface and Layer Exchange ASCII Format [17]. All of these tools have assisted in quality improvement, and with the emergence of the internet and with the support of open-source hardware and software developments, AM processes regarding development has grown immensely.

3



*Figure 5: Illustration of a gas turbine blade produced by additive manufacturing.*

## Digital workflow

AM process holds a data flow, which is made for the machines in order to generate the required instructions and to generate the final parts. This is also accompanied by a physical workflow, which matches the data preparation and digital workflow. On the other side, the solid work-flow is initiated using the AM technologies; these include the binder jetting, deposition, extrusion, material jetting, powder fusion, lamination, and polymerization. With this direct production; models and prototyping would be possible for the AM processes. Meanwhile, Meanwhile, in [18] the need for data structure, data generation and formation with a novel approach of data decomposition is illustrated.

## Design Manufacturing and Assembly

AM design is considered an optimization and as a goal to reduce the development time and the cost. Hence, this helps in quality and performance boosting. It also can lead to profitability concerning consideration of constraints. Meanwhile, it can be stated that the design process would have three levels where for the first level concrete tools and the required techniques are addressed. On the other hand for the second level, it will tend to provide a better understanding of the quality of process design. Moreover, at the highest level is that it investigates the impacts of relations between the design and the manufacturing on the designer, process and the practice[19][20].

---

<sup>3</sup> Courtesy of Siemens [www.siemens.com/](http://www.siemens.com/)



## Design Requirements

In practice, the design is different from the conventional designs perhaps all regarding knowledge, rules or even tools. Therefore, it is required to obtain the design rules and tools which are compatible with different types of constraints and features imposed by manufacturing processes. That is where the need for the process tools and similar rules emerge while it is considered as among the most disputed challenges an AM process can face [21].

## Design and Opportunities

The designs will be accompanied by opportunities and benefits. These benefits are mainly associated with the additive manufacturing process with freedoms. Mostly, the design is evaluated regarding part level or micro-macro complexity or multi-scale complexities. In process material choice, color and topologies would be considered as other production design factors[22][23].

## Constraints

The AM process with all the capabilities and potentials is also facing limitations, which must be considered by the designers. These limitations are manifested in forms of constraints, which are categorized in groups. This includes the discretization and directionality and the requirements for a novel design. Another critical constraint is the machine with their features and capabilities. Manufacturing parts are specifically impacted by this. On the other hand, materials that are being used in the process are influencing an AM process. In this term, raw materials can be used though they must be adapted accordingly. Metrology and control quality techniques are also imposing significant difficulties in an AM process. They are mainly related to material verification and geometry. Embedded design and materials from a different point of view will apply maintenance, repair and recycling constraints. This is while the design concerning regulations is also subject to constraints this is applied in such as aerospace and medical application where an approval grant is necessary before final use. In conjunction with this, some of the limitations are connected to Computer Aided Design Models shortly as CADs. Therefore, the design of complex and final products would be challenging due to the parametric nature of natures. That would because of the appropriateness of such models for old-style processes and not gradual shapes[23] [24].

## Costs

For the industrial applications, the AM is adopted as a production method where it is subject to cost. This is where the cost emerges as a barrier for AM process in industry. However, it is helpful in terms of economic to understand it and benefit where it is a motivation for manufacturing companies. One of the significant cost which is studied, is the production and parts cost. This can include the labor cost machine and material. Others can include the failures; transportation etc. different cost models are represented in order to optimized this concept and increase the feasibility. Build time and energy consumptions are the other factor. The goal would be to discover the strategies in all life-cycle which can result in process and product cost optimization [23][25].

## FUSED DEPOSITION ADDITIVE MANUFACTURING

Fused Depositing Modeling shortly referred as FDM is developed by Stratasys and is considered to be as one of the essential extrusion-based AM processes. This method uses a heating section denoted as a heat chamber to melt a polymer-based material by applying a filament feeder system. The chief merits of the FDM process are categorized and presented as follows:

- Material type and ranges
- Mechanical property
- The strength of the material

Meanwhile, this technique cannot be efficiently welcomed due to the following shortcomings:

- Build time is usually high
- Speed and acceleration is slow
- It requires a point-wise plot and deposition [26]

However, the polymer extrusion technologies and especially the FDMs are the AM processes where they are applied habitually and are the subject of research even in other fields such biology.

FDM usually encompass melting, extrusion, deposition, and solidification whereby application of thermal energy conduction the material will become molten and distributed for the cooling while it is infusing the layers by thermal bonding. One important thing to note is that in the process different methods can be used to deliver the filament concerning deposition. These includes

1. Self-extrusion
2. High-pressure plungers
3. Rotary pumps
4. Thermoplastic liquefier delivery [25].

In this thesis, a liquefier delivery technique with an extruding filament is studied. In FDM the layers are deposited in the same direction of the flow of polymer melted by a nozzle and resulting in shapes that would have a form of a flat ellipse. In FDM, observation of the complicated phenomenon in order to control the process has been the subject of studies for a long time. These consists of tackling with the problems such as thermal transfer and tension, layers deposition, bonding process and cooling time.

Meanwhile, nozzle behavior understanding and similar design of it besides the control of thermal behavior must become the main focus of the studies since they have shown a quality increase in cases [27] [28].

Moreover, two questions seem to be addressed as part requirements are achievable or not and what are the optimal manufacturing parameters to reach the desired part specification. In this case, the prerequisite for an approach to use as a closed loop control method would be required such as metamodeling techniques in order to realize a novel controlling unit [29].

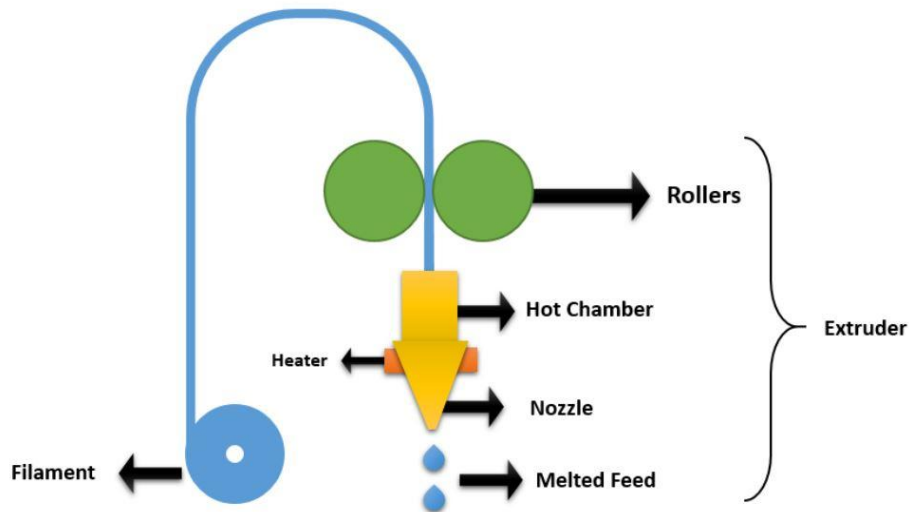


Figure 6: Illustration of a Fused Deposition Modeling Process.

Furthermore, a significant factor in the design of 3D products especially the FDM is the consideration of the processes in three necessary forms. This is very vital for designers to keep track of the product and its status in these steps.



Figure 7: Illustration of the printing steps.

For instance, in the pre-processing section, a model of a product is created using the 3D CAD models. After that, this model fundamentally is sliced into 2D layers of equal heights where the 2D layers are made of contours. These contours are in next step translated into trajectories using path planning algorithms where they will be translated into actuators commands. In some cases, these trajectories will become present as geometric codes or simply as g-codes.

G-codes are prepared codes where they determine the type of motions in contours. The g-codes are following a program based files where they include a sequence of words in a block. Each block and each word has a corresponding meaning where it establishes a command.

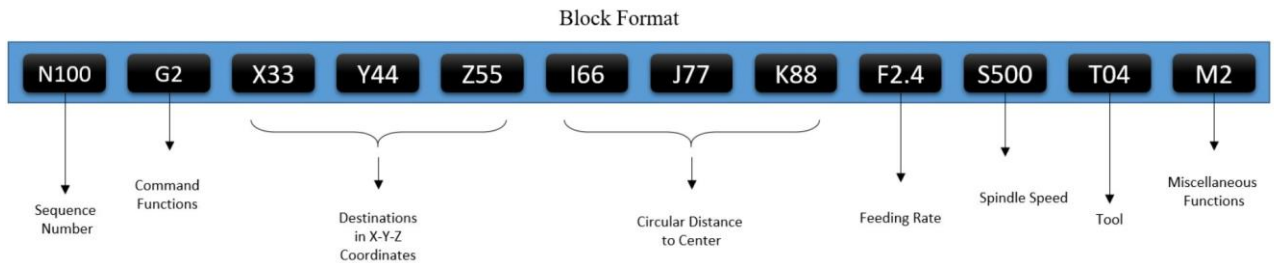


Figure 8 : Illustration of a G-Code block with corresponding functions.

In general, the flow of the work in the pre-processing section of printing can follow the subsequent model:



Figure 9: Illustration of pre-processing workflow.

## WIRE-ARC ADDITIVE MANUFACTURING (WAAM)

One of the section which is the part of WAAM is the welding unit. This unit is constructed from the following compartments and establishes the welding unit[30].

### Welding Techniques

There are numerous welding techniques with different characteristics can applications. Among these welding techniques, some hold greater importance concerning application in additive manufacturing which is studied subsequently.

### GMAW welding

It is one of welding used in the additive manufacturing and is mostly reached by its subcategories namely MIG and MAG welding which are introduced as follows.

### MIG / MAG Welding

It stands for the metal inert gas welding or the metal active gas welding. Within these processes, an arc is made and with heat generated by the arc work piece and metal are in-fused for joining. Additionally shielding gas is provided to prohibit the working space from the contaminations. There are two methods one is a constant voltage, and the current one possibility is the alternating current. Historically they are used for aluminum welding and non-ferrous materials[31].

### TIG welding

It is a gas tungsten arc welding or shortly GTAW and is referred to as tungsten inert gas welding. It is applied for arc welding using tungsten electrodes to produce the weld[32].

### CMT Welding

The CMT stands for the cold metal transfer and is a technique which provides a stable arc and precise process control. What advantage this process has over the conventional method (MIG and MAG welding) is that it provides accurate welding by an alternation of the hot and cold process. This leads to spatter-free welding and seams with brazing. Lower heat input, less distortion, extremely stable arc control precise droplet detachment infusion of thin-even coated which this would also lead to welds with joins of steel and aluminum. During this process, an integrated wire motion is provided which by retracting the wires by 130 times a second. As illustrated in the below the metal filler is moved toward the welding surface, and when the filler touches the surface or get close enough to the surface an arc is initiated, and the current correspondingly melts the material. Then the filler is retracted and helps the droplet detach with short circuits this would reduce the excessive heat. In the end, the wire is reversed, and the process begins again[33]. Due to difficulty and special requirements, three process is introduced for the CMT as follows:

## CMT Pulse

It is a combination of the two processor cycles where a pulsed cycle and CMT cycle is in-fused and produces inputs with more heat. As a consequence, there would be an increase in performance and flexibility.

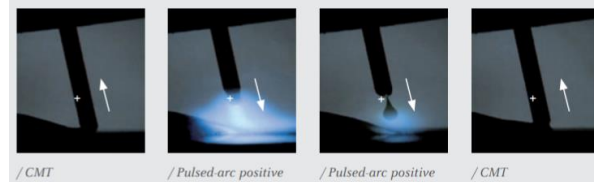


Figure 10: Illustration of CMT in Pulse mode [7].



Figure 11: Illustration of current and voltage diagram of CMT in different modes [30].

As illustrated in the above diagram the CMT modes and other welding processes are drawn based on the current and voltage which in general indicates the amount of power and heat. The CMT modes as illustrated are producing less heat and lower distortion. Arc also is controlled precisely.

## CMT Advanced

This mode benefits from welding current polarity. The main feature of this process method is the reversal polarity. It short circuits the polarity of the reversal and provides a stable CMT. It also provides better deposition at a higher rate.

## CMT Advanced and Pulse

This mode is a infusing the negative pole CMTs and poles pulsing cycles and provides a high precision arc[34].

## Shielding Gas Types

Numerous types of shielding gases exist in this process where they would be as follows:



Figure 12: Illustration of list of shielding gases.

### ACETYLENE (C<sub>2</sub>H<sub>2</sub>)

Acetylene, when used with oxygen, is suitable for *welding and cutting steel* and other metals.

### ARGON (Ar)



Argon is often used for *MIG welding aluminum*. It's also used for *TIG welding*.

### ARGON / CARBON DIOXIDE (Ar / CO<sub>2</sub>)



Argon/carbon dioxide blends are frequently used as shielding gas when *MIG welding mild steel*.

### CARBON DIOXIDE (CO<sub>2</sub>)



Carbon dioxide can be used as shielding gas when MIG welding carbon steel. KMS in Coquitlam fills 5 and 10 lb. cylinders, ideal sizes for beer carbonation and freshwater aquarium systems.

### NITROGEN (N<sub>2</sub>)

Nitrogen is commonly used to clean refrigeration systems. It's also an option for tire inflation. Nitrogen is less susceptible to temperature variations than compressed air. Plus it reduces heat and therefore reduces rolling resistance, which can translate into fuel economy. For plasma cutters, nitrogen can replace compressed air.

### OXYGEN (O<sub>2</sub>)



Oxygen, when used with acetylene and other fuel gases, is ideal for welding and cutting steel and other metals.

### TRI-MIX (Ar / O<sub>2</sub> / CO<sub>2</sub>)



Tri-mix applications include short arc, spray and pulsed spray arc welding of stainless steel and MIG welding mild steel.

## Wire Feeds or Filler

Different types of filler are allowed to be given as the input to the wire feeder section. These include the 26 types of materials. Each of these fillers is offered in different kinds of diameters, and depending on the type of the diameter there is a corresponding shielding gas for each. The main fillers include:

- Nickel
- Stainless steel
- Stellite



## WELDING DEFECTS

A defect in the welding process is referring to the realization of any non-desired formation in welding which is considered as improper and based on the criteria categorization an incorrect welding pattern. These defects usually manifest in the form of:

- Size
- Quality
- Or Shape

Moreover, based on the location of that they occur, they can be categorized as:

- Internal
- Or external

Namely, the welding defects would be as the list of following:

1. Cracks
2. Porosity
3. Undercut
4. Slag
5. Spatter
6. Fusion
7. Penetration
8. Overlap
9. Crater

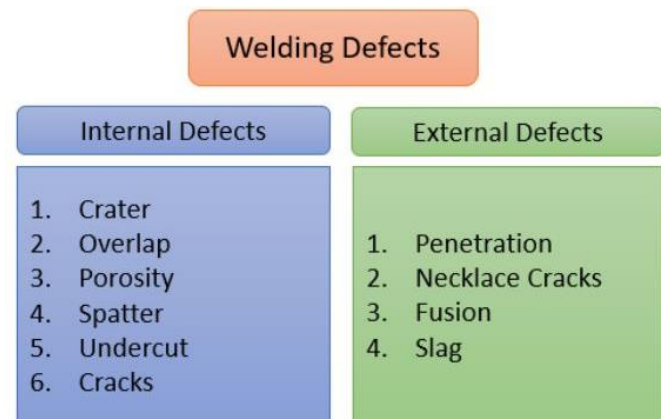


Figure 13: Illustration of categorization of the welding defects.

In Figure 14 a proper welding is portrayed. Subsequently, some of the defects are considered more in-depth as in following sections [35][36][37][38][39].



Figure 14: Illustration of a proper welding.

## Cracks

This type of defect is more present on the surface and also in depth of the welds around the heat-zones **Figure 15**. Temperature is one of the affecting factors where it divides the crack into two sub-categories.



*Figure 15: Illustration of the crack defect.*

- Hot  
This sort of the crack is happening while crystallization is occurring.
- Cold  
This sort also happens usually at the end of the welding process and where the temperature is low. Visibility of this crack regularly must be checked after sometimes and even days later after welding.

Cracks usually happen due to the ductility of the base material. Carbon and Sulphur existence on the material can also cause such issues. Application of the hydrogen as a shielding gas through the process besides the rigidity of the materials for expansion or contracts. Cracks can usually be the cause of stress in the process too.

To reduce the risk of having cracks in the welding process; adopting a novel material is required. Meanwhile, another method is to pre-heat the process and reduction of the cooling time. The gap of materials for the base and the filler of the welding and during the welding process must be addressed to avoid the cracks.

## Undercuts

This phenomenon happens when the weld and base melt away from each other. At this stage, a groove of notch shaped is formed which is referred to as the undercut **Figure 16**. The general impact of this process will manifest in a reduction of fatigue strength of the weld and joints.

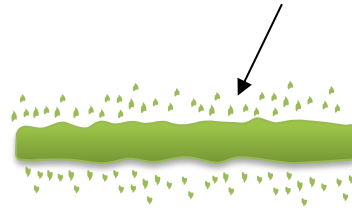
This usually happens when the arc voltage is very high and also if a wrong filler is used. One another important reason for such defect is the welding speed where with a higher value it would become more susceptible for undercut defect. To reduce the risk of an undercut defect, it is required to decrease the arc length or arc voltage. Proper setting of the filler diameter as becoming smaller can help in reducing the undercuts. Also, adequate adjustment of the angle of the welding with consideration of travel speed can help reduce the undercuts.



*Figure 16: Illustration of the undercut defect.*

## Spatters

Spatters or spattering of metal drops which is observable and well-known in welding processes refers to a phenomenon where droplets of metal during the welding will become expelled and will be left on the surfaces **Figure 17**. One of the main reasons for this effect is the amount of high welding current. The long arc with an incorrect polarity will also lead to a higher amount of spatters. As a consequence for reducing the amount of spatters it is required to Arc Length reduction, and current of the weld besides checking the right polarity of welding can help in spatters reduction. Adjustment of welding angle and use of a proper shielding gas will assist in reducing the spatters in a welding process.



*Figure 17: Illustration of welding spatters around a welding bead.*

## Porosity

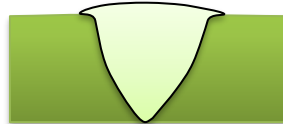
The porosity defect happens when gases in the form of bubbles will become trapped in the welding **Figure 18**. Filler property and larger arcs with a high current can lead into porosity problem. Increased welding currents. The hazardous material on the welding base such as oil or rust can lead in spatters. Therefore, to reduce the main effecting causes the reduction of the current, adopting smaller arcs and removal of hazardous material on the surface of the welding properly would assist in the removal of porosity.



*Figure 18: Illustration of porosity defect in welding.*

## Overlap defect

This defect happens when the weld is going to be extended beyond the weld toe and occurs where the weld forms an angle usually smaller than 90 degrees **Figure 19**. This often occurs with the improper welding and large wires or electrodes and also with using high current. Consequently, for reducing the risk of having overlap defect, it is suggested to have proper techniques for welding and using smaller wire or electrodes and reduce the welding current.



*Figure 19: Illustration of the overlapping defect in welding process.*

## Crater & Slag

This type of defect emerges in the welding and when the craters have the trouble of being poured before an Arc is initiated. This would lead to outside edges to become cool earlier and will lead to stress and cracks. The main reason is the torch angle and large electrodes or fillers. For not confronted with such an issue it is suggested to adjust the angle and use a proper welding material.

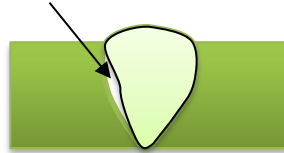
While, Slag defect usually affects toughness or structure quality **Figure 20**. Usually, it either emerges on the surface, or the welding turns. The primary cause of the slag is the current and corresponding density when it is small. This causes the reduction of heat on the surface of the melting metal. In addition, if the speed of welding is high, this can lead to slag occurrence. Welding surface's cleanness is another crucial factor for the slag to happen. The angle of the welding and the travel rate would have an impact on the slag. To assist the process with slag, increase of current, adjustment of the speed (feed rate), modification of Travel speed and clean surface with a proper welding angle can help the reduction of the welding slag risk.



*Figure 20: Illustration of the slag defect in welding.*

## Fusion incomplection

The fusion incomplection is usually due to the improper welding. This will lead in solidification of the metal and constructing gap in melted metal **Figure 21**. Low heat input, large welding pool, wrong torch angle, and bead position are considered as the leading causes of an incomplete fusion. The increase of welding current, reduction of the travel speed and deposition rate, adjustment of the torch angle properly and adjustment of the bead position can help immensely in reducing the risk of the fusion incomplection.



*Figure 21: Illustration of an incomplete fusion defect.*

## Necklace Crack

This usually happens in the electron beam welding. When welding has not enough penetration molten metal will not have proper flow and especially an in cavities and lead to such defect. The primary cause is related to the welding technique and when working with the material such as stainless steel, nickel base alloys, carbon steels, and tin alloys. Besides, another leading cause of this would be due to a high-speed beam welding. Hence, to diminish the risk of this defect, it is necessary to adjust the welding speed at a constant value properly and similarly to use a proper material and welding techniques.

## Grooves and Penetration

This welding defect is happening when the groove of the metal is not filled **Figure 22**. This defect is also denoted as an imperfect diffusion or penetration. Typically, the grooves effect occurs when there is not enough deposition of welding and electrodes or welding filler is not as a right size. This is while welding technique is also playing a significant role. To tackle with this defect, it is essential to use adequate welding deposition and use an appropriate electrode or filler.



*Figure 22 Illustration of the Grooved penetration.*

## ENVIRONMENT

The environment which is applied for the manufacturing using the WAAM consists of the following groups of equipment. Each section would be considered as an influential section on the AM process and correspondingly will affect the process and quality of the part which is printed. Therefore, analysis and addressing the impact which each section would have on the process would be required to lead an enhanced additive manufacturing process.

4

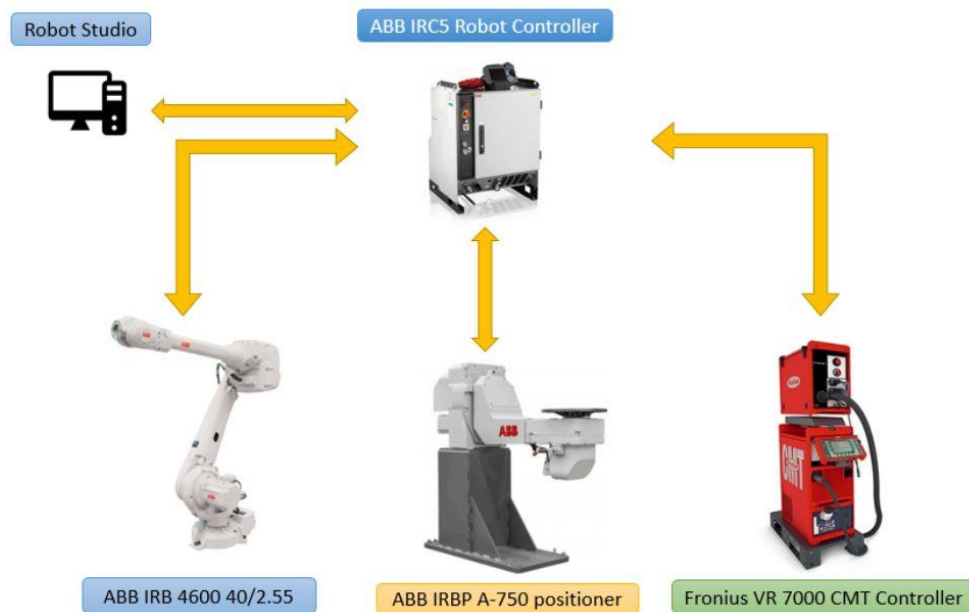


Figure 23: Illustration of the equipment of additive manufacturing using WAAM technique.

<sup>4</sup> Courtesy of ABB adopted by <https://new.abb.com/products/robotics/industrial-robots/irb-4600/>

## ABB robot manipulator

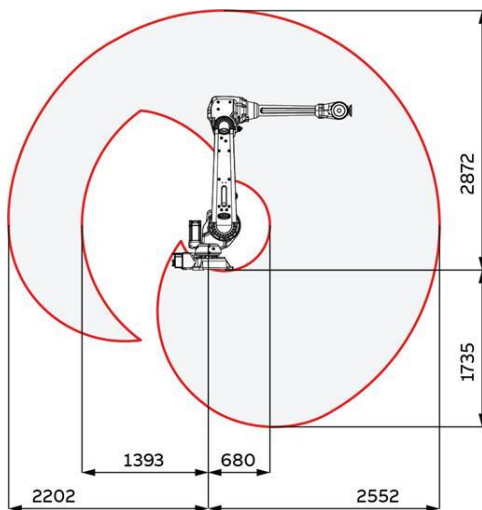
As it is illustrated in **Figure 23**, an IRB 4600 40/2.55 robot is used for the manipulation of the environment. This robot has 6 degrees of freedom referred to as the axis. Each axis is equipped with the following features where the working range and maximum speed of the axis are illustrated in the following table. This information must be taken into account as it will have a corresponding impact on the path and trajectory of the manipulator on any process.

The main application of this robot is manifested in the form of the:

1. Arc Welding
2. Packing
3. Laser Cutting
4. Assembly
5. Laser Welding
6. Machine tending and removal

The robot has the weight of the 435 Kg, and its respective dimension can be found in **Figure 24**. The dimension of the robot also is applicable in modeling and trajectory planning which have significant impacts on the qualitative features of the AM process's products especially for applications related to wire arc additive manufacturing [40] [41].

5



IRB 4600-40/2.55

Axis movement	Working range	Axis max speed
Axis 1	+180° to -180°	175°/s
Axis 2	+150° to -90°	175°/s
Axis 3	+75° to -180°	175°/s
Axis 4	+400° to -400°	250°
Axis 5	+120° to -125°	250°
Axis 6	+400° to -400°	360°

Table 1: Axis information of the robot and corresponding speeds.

Figure 24: Illustration of the working range of the robot.

<sup>5</sup> [https://search-ext.abb.com/library/Download.aspx?DocumentID=ROB0109EN\\_G&LanguageCode=en&DocumentPartId=&Action=Launch](https://search-ext.abb.com/library/Download.aspx?DocumentID=ROB0109EN_G&LanguageCode=en&DocumentPartId=&Action=Launch)

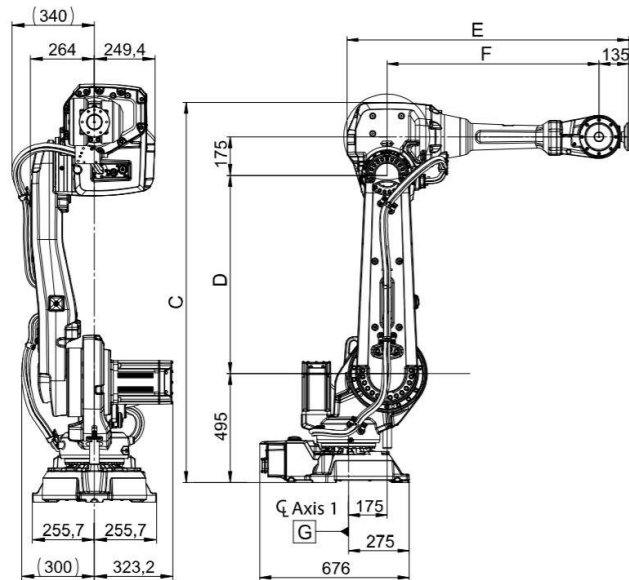


Figure 25: Illustration of dimension of ABB IRB 4600-40-2.55 manipulator.

### IRBP A-750 Positioner

This positioner is an ABB product among other two similar variants (IRBP A-250, A-500) and is applied in order to rotate the work pieces on two axes. This ensures the reachability of the positioner in the process as they can handle the work pieces up to 740 kg load. Because this product is design as modular, it is considered as friendly service device [42].

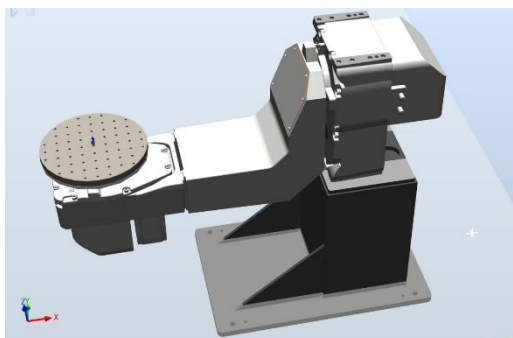


Figure 26: Illustration of the IRBP A-750 Positioner.



## ABB teach-pendant

The Teach Pendant is joystick similar to a PlayStation or Xbox controller where it provides the tool for control, monitor and works with any manipulator designed by the ABB company. This controller is connected using a cable to the IRC5 controlling box. On IRC5 control box it is possible to adjust the robot operating mode whether to be in:

1. Automatic Mode
2. Manual Mode
3. Full speed 100% Manual Mode

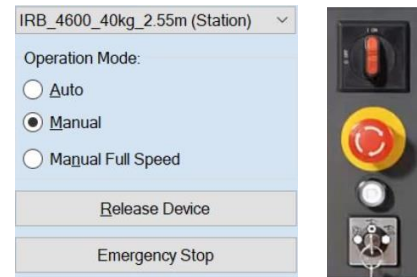


Figure 27: Illustration of operation modes of the robot.

With a nominated mode for operation, it is promising to launch, edit (program, positions), monitor and jog the robot in different procedures using the Flex Pendant. This, also, delivers a mean to connect different other external devices and to monitor or edit a networking interface with online services such as cloud systems in order to collect data, model or modify the robotic arms beside the external devices such as in this case the Fronius welding devices. For instance, the information about the welding process and the position of the robots could be available using cloud services and IoT technologies. Further information regarding the Flex Pendant cloud is found on the ABB robot website and through the catalogs.

6

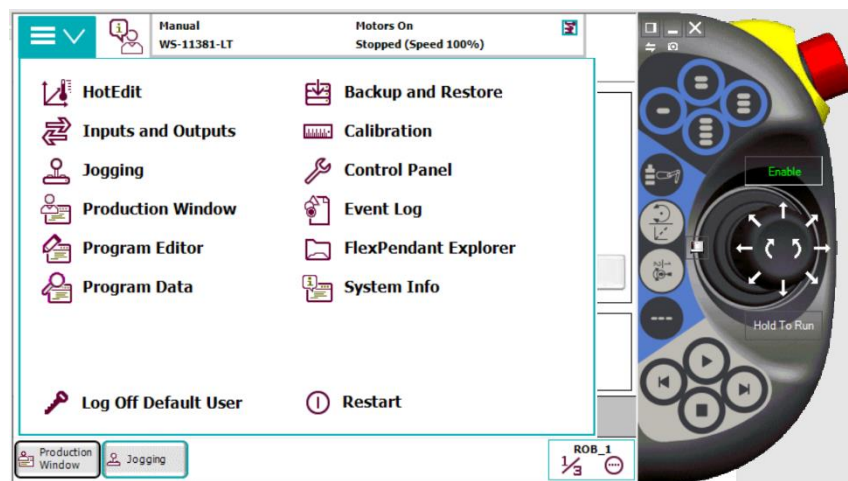


Figure 28: Illustration of an ABB Flex Pendant and its virtual format in Robot Studio.

<sup>6</sup> [new.abb.com/products/3HAC028357-001/teach-pendant](http://new.abb.com/products/3HAC028357-001/teach-pendant)

## Cold Metal Transfer System

The Cold Metal Transfer System is constructed from multiple subunits with a corresponding role in the welding process. These as illustrated in the following **Figure 29** consists of a wire feeder unit where a roll of material, for instance, steel, aluminum, etc. is inserted in where 4-roller drive feed the material to the wire-spool until it reaches the working area.

The next section is known as the power source where the electrical unit is located and defines the amount of power delivered to the working pieces. It is benefiting from a microprocessor where it inverts the power necessary for the welding process and ensures the welding process embrace all the defined properties.

The other section is the cooling unit which provides a cooling medium for the welding system and guarantees the water cooling process for the welding torch. The RCU 5000i is the remote controller unit where the parameter, modes of welding (TIG-MIG, CTM modes, 2-step, 4-Step modes), external or internal control commands are defined as jobs in this controller and allowed authorities (with the login option) can modify the jobs to perform the welding process.

The next unit is the Wire Buffer. This Unit is a coupler which connects the front and the rear drives. This paves the path for the wire to travel smoothly. The Robacta Drive CMT unit is a gear-less unit where forward the wire material and almost 90 times back up in second to ensure the reliability of the constant wire contact pressure. This would bring about a better quality of the welding process. The contec contacting system has the role of ensuring the contact surfaces and the force of the contacting system and welding wire is kept at a specific range. This unit is useful to maintain and keep the contact tip uniformed where to control the uneven wear and behavior of the process.

7

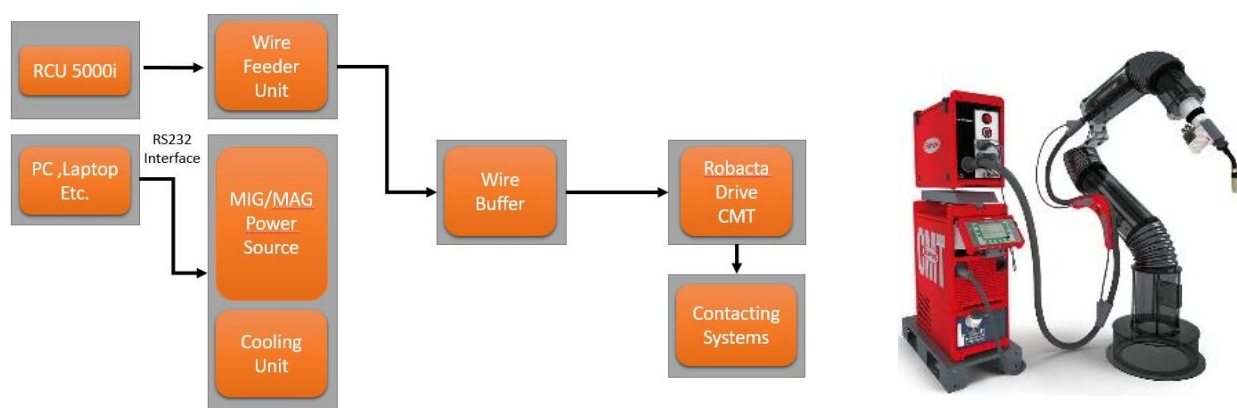
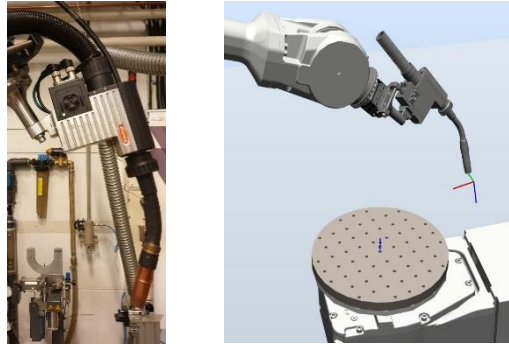


Figure 29: Illustration of the Welding Systems Architecture.

## Torch

The ABB IRB4600 40-2.55 robot is equipped with a tooltip of a Fronius welding torch as illustrated in the following **Figure 30**. This welding torch consists of two main parts where they are as follow:

- Robacta Drive CMT
- Robacta 5000 torch neck

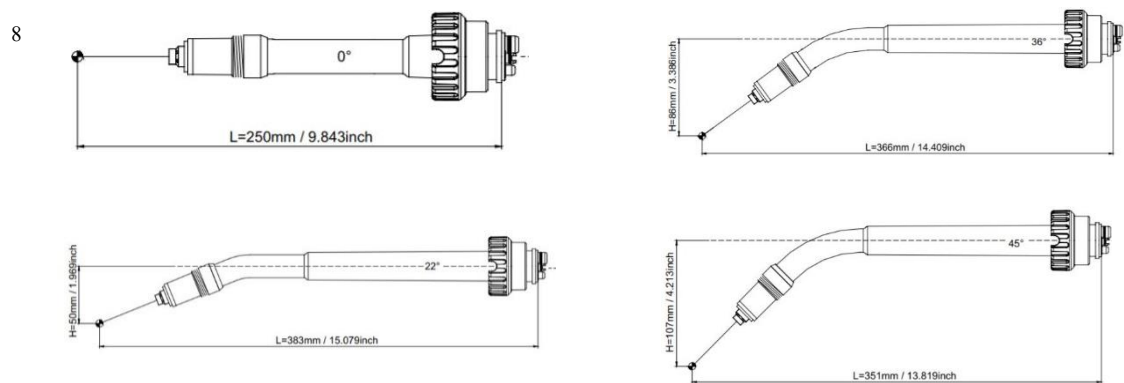


*Figure 30: Illustration of the Fronius Welding Torch.*

The Robacta Drive CMT is constructed by

1. A Torch
2. Wire Buffer
3. Drive Unit
4. Hose pack

Depending on the type of the application the torch is chosen as either of Robacta 280, 300, 400, 500, 700 or 5000. These torch necks would have a corresponding angle as  $22^\circ$  or  $36^\circ$ . However, these are also presented in  $0^\circ$  and  $45^\circ$  format too **Figure 31**. The torch angle and specifically the angle made to the working object is playing a key role in the quality of the welding and must be considered respectively.

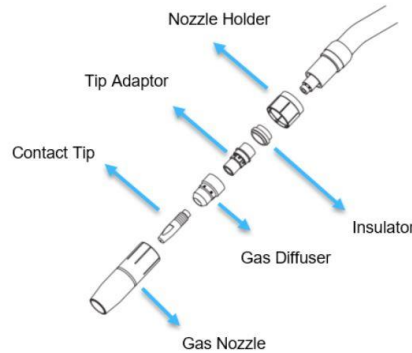


*Figure 31: Illustration of different torch necks with corresponding angles.*

<sup>8</sup> [www.fronius.com](http://www.fronius.com)-Robacta 5000 manual

The torch necks are also constructed from sub-components **Figure 32** referred to as wearing parts which they shape the characteristics of the welding torch. One of the weightiest components is the contact tip which defines the size of the wire feed. Moreover, as it will be illustrated they would be considered as significant influential parameters in welding. Meanwhile, another sub-component is the gas nozzle. The gas nozzle defines the behavior of the flow of shielding gases. This affects the quality of the welding.

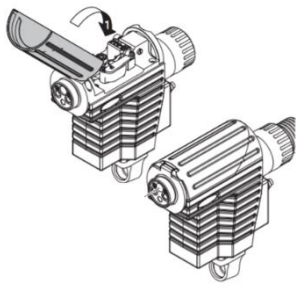
9



*Figure 32: Illustration of the torch neck wear parts.*

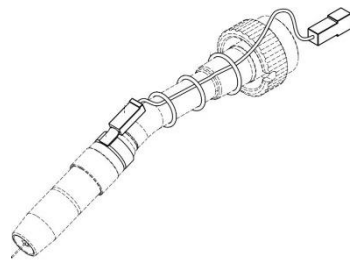
When the attachment of the wear parts and other related components is finalized; the torch neck is attached to the drive section forming the welding torch **Figure 33** and **Figure 34**.

10

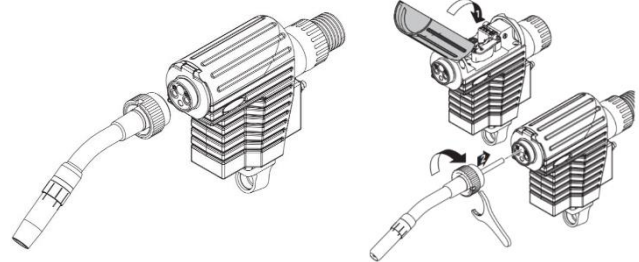


*Figure 34: Robacta Drive CMT.*

11



*Figure 35: Robacta 5000 with a touch sensor.*



*Figure 33: Illustration of attachment of Robacta Drive and 5000 torch bod.*

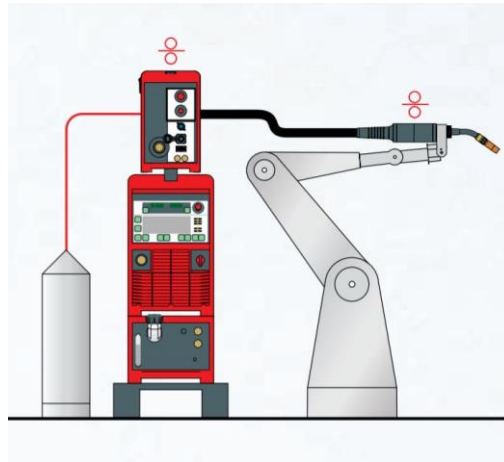
<sup>9</sup> [www.fronius.com/downloads-](http://www.fronius.com/downloads-) Robacta 5000 torch body manual page 12

<sup>10</sup> [www.fronius.com/downloads-](http://www.fronius.com/downloads-) Robacta CMT Drive manual

<sup>11</sup> [www.fronius.com/downloads-](http://www.fronius.com/downloads-) Robacta 5000 torch body manual page 38

The welding torch also as mentioned before will be connected to the shielding gas capsule and wire feed system where they form all together the architecture of the welding compartment illustrated in the following **Figure 36**:

12



*Figure 36: Illustration of the welding unit.*

The red signs indicate the existence of the welding material inside the wire feed unit. This passes through a hose beside other signals and power cables, reaching the welding gun which makes the wire feed material available in the torch for the welding process. This hose also supplies the shielding gas to the welding gun. For instance, MISON 8 gas which is a combination of the CO<sub>2</sub>, Ar and NO is supplied for the construction steel welding and manufacturing applications<sup>13</sup>.

<sup>12</sup> Robacta, Robacta Drive MIG/Mag Robot welding torches manual by Fronius

<sup>13</sup> [http://www.aga.se/en/products\\_ren/cutting\\_welding\\_gases/shielding\\_gases/mison\\_shielding\\_gases/index.html](http://www.aga.se/en/products_ren/cutting_welding_gases/shielding_gases/mison_shielding_gases/index.html)

## METHODOLOGIES AND STATE OF THE ART

Experiments implementation and especially experiments that one parameter value is changing over the time is considered to have the price of time and cost in order to evaluate. Therefore, it is required to optimize the collection of the data by applying methods which are time and cost optimized. The design of experiments shortly referred as DoEs are the methods that will maximize evaluation optimization by utilizing sufficient experiments. The following list is referring to some of DoEs which are used for design analysis and finally for inferences.

- Factorial design
- Combinational designs with different levels
- Surface response design
- Screening design
- Mixture design
- Tagouchi design

These methods are among the designs which hold different selecting criteria for analysis of the most influential parameters with their corresponding values [43]. Within the context of the AM processes, the extent of parameters that are involved in the quality of the manufactured parts is great. This also demands to discover the experiments' design space and parameters combinations simultaneously. In such cases usually, sampling methods are more convenient to explore the design space in an efficient manner [44][45]. Additionally, one important concept that must be considered is the parts quality with ensuring repeatability. One reason for this maybe is the presence of unwanted phenomenon and it is necessary to investigate it during the experiments implementation. Besides, in order to analyze the designing methods which are mentioned earlier, it is shown that fractional space consideration in DoEs can lead to promising results [43][46][47].

Alternatively, the poor repeatability in AM processes itself is controllable with consideration of the main parameters and by applying a closed loop control system on it. The novelty of this approach and the subsequent differences are more observable for variant parameters and especially in processes such as FDM. However, it is worth to mention too that usually, the repeatability problem is originating from the latency of control processes. This is resolvable by ensuring the controlling factors at a promising level [48]. Therefore, such issues especially the latency of controlling factors must be well-addressed during the process and this requires great computation need. Meta-modeling techniques with the help of artificial neural networks in such cases can help to address the controlling of parameters efficiently. This illustrates better the importance of the modeling and can assist to reduce costs and to increase the quality. Likewise, experimental data and knowledge will be necessary to model the AM processes appropriately. In order to construct models for a given data, statistical regression model and other techniques could be applied.

In [49] mathematical formulation of the bead width models using genetic programming approach is studied. Similarly, in [50] by using Tangent Overlapping Model, a bead analysis is conducted for WAAM processes. Moreover, in order to increase the forecasting capabilities, the genetic programming modeling and hybridization with computational methods such as artificial neural networks, fuzzy logics, regression trees, and support vector machines are conducted [51][52].

Furthermore, in [53] a 3D printed FDM part is processed by evolutionary calculation algorithms while the study of the AM processes using computational techniques are found to perform better and even artificial neural networks has shown better performance in comparison to genetic programming and regression algorithms[49].

Nevertheless, these computational models are used correctly; however, they are time-consuming and there is a necessity for using strategies that can be applied to avoid intensive computations. Metamodeling approaches can provide such strategy by using artificial neural networks and in order to control the involving parameters. This type of modeling can be beneficial concerning prediction and it also can support closed-loop control systems. In these conditions, the cost of the learning approaches in artificial neural networks would be high. Then it is required to apply a technique which reduces the amount of experimental learning data but with better or close to approximations that are used with a high amount of training data.

One approach is the minimization of the data by using an expert or a physical knowledge and applying management procedures and integrating this knowledge into artificial neural networks. The process of knowledge extraction will provide the possibility of using the knowledge that is exciting especially in, and it can be used for design and planning [54][55].

## **Dimensional Analysis Methodology**

One of the frameworks which are used to translate the existing knowledge is the Dimensional Analysis Conceptual modeling [49] and its practices and encodes the existing knowledge linked to a system. This representation provides a causal graph which is capable of constructing a ground for an adaptive artificial neural networks training. This modeling technique benefits from system boundary and objective definition which provides a better understanding of a system. This method similarly provides a function representation that contributes to the functions and their order in a system and demonstrates its various performances. The DACM approach has the role of transforming these functions into generic functions and applies causal rules inherited from Bond graph theory for the description of a system.

Subsequently, behavioral equations are constructed by using this method for the bond graph, and with the similar nature, a color is allocated to each node for representation of parameters. The model of the system can in future be used for simulations purposes such as quantitative or qualitative. Finally, the modeling process in this method ends at the point whether details of a system of interest is provided or not [56][57][58].

## Computational Methodologies

DACM besides another type of component analysis modeling approaches will provide the mathematical relationships that manifest the linked between variables. This is where in [59] principal component analysis is used as a feature extraction tool for obtaining sensory data and similar associations. These sensory data are applied for analysis by statistical approaches or artificial neural networks such as Bayesian Neural networks to predict the desired output. Knowledge and features are also extracted by employing another type of machine learning and information processing such as Mutual information method. With such tools, there is a possibility to define the amount of information mutuality among the parameters [60]. One another advantage is the ability to check a linear or a non-linear reliance with these computing tools.

This is while, application of other machine learning techniques such as Bayesian neural networks in manufacturing processes are still the subject of research such as the application of the quality analysis in manufacturing processes [61].

Artificial neural networks as computational tools are also used in variant processes and have the strength of prediction of output parameters of an interesting process. These tools can model different developments by using computational power and optimization techniques and to provide precise solutions for the estimation of desired variables. The chief merit of using artificial neural networks is the capability of computing unclear data and describing the behavior of complex systems. In the meantime, it is worth to mention that the artificial neural networks are the models which not only require a large number of data samples but also they would require a high amount of computational power too [62][63].

## Analysis of Variance (ANOVA) Methodology

This statistical method is an analytical technique that interprets the difference between sets of parameters or values. This method assists the researchers in the recognition and assessment of the parameters and discovers the possible interactions between variables. ANOVA usually address the questions that indexed variables are affected by each other or not? Moreover, this method is used to compare the different values of the variables. It is usually conducted for different type of independent and dependant variables and also for different number of groups this is conducted as One-way ANOVA, Two-way ANOVA or Multivariate ANOVA. In Addition, this technique is accompanied with follow up tests to evaluate variant groups of variables and [64][65].



## Machine Learning Methodology

At this stage, it is worth to refer to the artificial neural networks and overview different types of these computational tools. Therefore, in this case, the computational intelligence or machine intelligence would be a branch of computational science that intelligent agents are designed for analysis and estimation purposes. Also, this would assist in perceiving an environment and to reach an optimized evaluation and perception of the desired environment. Artificial neural networks are bio-inspired networking systems that can communicate, learn, reason and adapt to different conditions. Meanwhile, they can perform statistical modeling and being an alternative solution for logistic statistical regression methods with the power of being easily transferred to different platforms [66].

Recently, Artificial Neural Networks (ANNs) are used as computational modeling tools to model and develop complex real-world problems. The main characteristics of ANNs are known as nonlinearity, high parallelism, robustness, fault and failure tolerance, learning ability to handle imprecise and fuzzy information, and their capability to generalize [67]. Nonlinearity makes better fit to the data and high parallelism indicates fast processing and hardware failure-tolerance. The main goal of ANN-based approaches (neuro-computing) is providing and developing mathematical algorithms that facilitate ANNs to learn by mimicking information processing and knowledge acquisition in the human brain.

In 1958, the procedure of the single artificial neuron and the perceptron have been introduced to find a solution for problems of character recognition [68]. The inputs of an artificial processing neuron are received from the environment, they act as stimuli. Then, the neuron combines them in a special method to form a net input, applies a linear threshold gate, and transfers the output to another neuron or the environment. A system with such structure is called the Perceptron. The Perceptron generates a mapping between the inputs, stimuli, and the output.

ANNs can provide a lot of advantageous and require less need for training and can detecting complex relations among parameters with non-linearity. Correspondingly, ANNs can discover the interaction between variables independently and intrinsically.

Some of the application is that in general, ANNs are more robust and often perform better than other computational tools in dealing with a variety of problems from different categories such as follows:

- I. **Pattern classification:** In pattern classification tasks, supervised learning is used to assign an unknown input pattern to one of predefined classes based on one or more features that describe a given class.
- II. **Clustering:** is based on unsupervised learning. In clustering problems, clusters (classes) are designed based on similarities and dissimilarities between the input patterns according their inter-correlations. In the clustering network, samples in similar patterns are clustered in a class.
- III. **Function approximation and modeling:** the goal is to approximate the principal rules associate with the inputs to the outputs. Function approximation is used to solve tasks where theoretical-based models are not available, data is gained from observations or experiments or to replace theoretical models with complex computation by using acquired data from such models.
- IV. **Forecasting:** given a time series samples from an observation with a certain phenomenon, an ANN is trained and then is applied on another observation to forecast (predict) its behavior in future.
- V. **Optimization:** the target is to maximize or minimize a function subject to a set of constraints.
- VI. **Association:** an ANN is trained by using an ideal data without noises and then is applied on a noisy data to classify the data.

To refer to some of the disadvantages, it is worth to refer to the fact that there is less amount of understanding about the hidden layers of neural networks and they are susceptible to overfitting and consequently generalization problem. Also, they put a lot of burden by immense computational power on users [63][66][69].

By the same token, it is common to face several forms of complications and the interest is to find the paramount and optimal solution with finding algorithms or set of instructions that transform the input of a system to the target outputs efficiently (especially concerning computational speed and memory requirement). For another kind of applications the knowledge of transformation (inputs to outputs) is at hand, but in most of the cases it is required to obtain the knowledge.

For the reason of being the ANNs are categorized as the following classifications meaning that the ANNs are classified based on the significant features. Based on these features the classification can be done based on:

- The purpose of using the ANN, for instance, clustering or pattern recognition.
- The type of learning method used to measure the performance of the system regarding the network's output and update the parameters of the network.
- The degree of the connection between layers and neurons, full or partial connection.
- The direction of the information flow in the network, recurrent or non-recurrent. In recurrent networks, feedback loops, such as Backpropagation, are used through time to loop information back into the network.
- The level of the learning supervision required to train the ANN. In supervised learning, the network is trained by using the right answer.

In such cases, the collection of data is acquired to obtain the knowledge and the transformation rules. If machines or computers do this task without programming, it is referred to as Machine Learning shortly ML. ML would be the method that uses sample data and learns from past experiments in order to obtain and optimize the transformation of inputs to outputs [70][71].

ML-based on learning technique falls into the following main categories:

1. Supervised learning methods
2. Unsupervised learning methods
3. Reinforcement Learning Methods

## Supervised learning (learning by a teacher)

Supervised learning refers to the practice of ML where an example of data with desired values are given, and the goal of the machine is to learn from these sample data and obtain the existing rules or the transformations and accordingly map the inputs to outputs. These data are usually categorized as training, test and validation data. On the other hand, based on the application type the supervised learning falls into the following categories:

- Classification applications
  - This type of learning practices the uses where the desired inputs will be checked to fall into a finite group. The output could be represented either as 0 not being a member of a group, or 1 as being a member or even a number where it corresponds to a group number which an input belongs to.
- Regression applications
  - If a problem or a task obeys this form that outputs consists of one or more continuous variables the problem is referred to as regression. An instance of this learning method could be in the chemical processes. This is where the yields prediction as output is represented as an input of the reactants, pressure and temperature [71][72][73][74].

Some of the application of supervised learning is mentioned as:

1. Pattern Recognition
2. Function Approximation

## Non-supervised learning (learning without teacher)

Within the context of this practice, the learning is conducted without assistance or any feature or label for training. Therefore, the data would be labeled based on the resemblances or correspondence and also based on hidden information. In case of having a set of observation the goal would be as inferring the properties inside of the inputs without any supervision. Besides, in contrast with the supervised learning, there would be no correct answer or a degree of error for the inputs. In such applications, the dimension is usually higher, and the properties of interest are broader [70][71] [74][75].

The application of the non-supervised learning falls into the following categories:

- Clustering
- Auto-encoders
- Hebbian learning
- Self-organized map
- Principal component analysis

## Reinforcement learning

This learning is constructed based on an intelligent agent and an environment. This is where the environment is providing a feedback to the agent and the agent also provides an in-tended action namely reward as a good feedback. This type of learning in comparison to the other learning techniques is not using any optimal output and is ordered to perform a trial and error procedure to discover the best solutions [76].

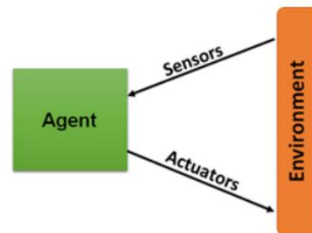


Figure 37: Elements of reinforcement learning.

## Learning rules

A learning rule is defined to modify the network weights of the connections between layers and neurons[77]. Based on this definition, there are four main types of rules[78][79].

- Error-correction learning (ECL) rule. ECL is used in supervised learning to update the connection weights to minimize the overall network error. In this technique, the difference between the ANN result, at any stage of training, and the corresponding right answer is computed.
- The Boltzmann learning (BL) rule. BL is a defined as a stochastic rule which works such as ECL, with the difference that the output of each neuron is generated based on a Boltzmann statistical distribution [67].
- Hebbian learning referred as HL is one of classical learning methods. It assumes and checks if the neurons of both side in a synapses are able to become active synchronously and the weight of the synapse is validated to increase. In contract to BL and ECL a local learning is conducted and weights are adjusted accordingly based on the activities of neurons[80].
- The competitive learning which is referred as CL is a type of learning method that forces all the neurons to compete among themselves and one neuron will be active in an iteration and with all the weights allocated [67].This type of learning method is existing in the biological application and systems [78].

Numerous type of networks exists and new ones come as sometimes some are just a modification of present existing ones meaning that they are continually being developed[81]. Some networks are noticeable in solving the perceptual problems while this is suitable only for modeling data and approximation purposes. Among these networks ANNs with Backpropagation network topologies and Radial basis function (RBF) networks are among the most popular one.

## Backpropagation (BP) network

A backpropagation (BP) network is a type of MLP with an input layer, containing nodes or input variables, an output layer, with nodes or dependent variables, and one or more hidden layers. The nonlinearity in the data can be obtained with the help of nodes of hidden layers. In these networks, supervised learning, with the ECL rule, can enable the networks to learn how to map from one data space to another by using samples. The goal of using backpropagation is to update each of the weights in the network to gain the actual output to be closer the target output, minimizing the error for each output neuron and the network. These networks are so flexible and can be applied for data modeling, classification, control, and pattern recognition [79].

In this thesis, the backpropagation algorithm is implemented, as follows:

1. Run the network forward with the input data to get the network output
2. For each output node we compute delta,  $\delta_k = O_k(1 - O_k)(O_k - t_k)$ , where  $O_k$  is output of node  $k$  and  $t_k$  is the target value of node  $k$ .
3. For each hidden node we calculate delta,  $\delta_j = O_j(1 - O_j) \sum_{k \in K} \delta_k W_{kj}^l$ , where  $W_{ij}^l$  is the weight from layer  $l - 1$  node  $i$  to layer  $l$  node  $j$ .
4. Update the weights,  $\Delta W = -\eta \delta_l O_{l-1}$  and then apply the new weights,  $W + \Delta W \rightarrow W$ , where  $\eta$  is the learning rate.

## General issues in ANN development

There are a variety of matters that should be considered before training any network. Some of the following issues are only relevant to BP ANNs while others are applicable to the design of all ANN types.

- Database size and partitioning
  - One of the significant information about the ANNs is that they are constructed and obtained with a variable size of databases. Therefore the generalization of the models is subject to study when data outside of the database is used. The development of the ANNs model need partitioning or segmentation for being uses as training and so forth. Therefore the subset data for training is used in the training phase and updates the weights of the network according to their data. The test data is the subset that checks the learning response of the network. Moreover, the last segment of the data is allocated to validation as to check the quality of the training and confirms its accuracy during the learning process or sometimes it is referred to as the final test.

In conclusion, illustrates the need for a good database and a proper way of segmentation of the data for the training, validation and test phases. Data preprocessing

- Data normalization
  - The neural networks also sometimes require a uniformed range of data for instance from (e.g., 0-1). Normalization is vital since it prevents the larger data samples not to override the smaller ones. It also assists in preventing the networks premature the saturation of hidden nodes that impedes the learning process.
- Data preprocessing
  - Neural networks also required several types of preprocessing of the data in order to provide a better fit and converge to solutions faster. One method is noise removal. Another method is input reduction and dimensionality. Data transformation and treatment of non-normality. Distributing the data and inspecting it and deleting any outlier data point[82][83][84].
- Input/output representation
  - Proper data representation also plays a role in the design of a successful ANN[84]. The data inputs and outputs can be continuous, discrete, or a mixture of both. Binary inputs and outputs are very useful in extracting rules from a trained network [85]. For this purpose, a continuous variable may be replaced by binary numbers by partitioning its range into a number of intervals, each assigned to a unique class.
- Network weight initialization
  - Initialization of a network involves assigning initial values for the weights (and thresholds) of all connections links. Some researchers [86] indicated that weights initialization can have an effect on network convergence.
- Convergence criteria
  - Three different criteria may be used to stop training: (i) training error ( $\rho \leq \epsilon$ ), (ii) gradient of error ( $\nabla_{\rho} \leq \delta$ ), and (iii) cross-validation, where  $\rho$  is the arbitrary error function, and  $\epsilon$  and  $\delta$  are small real numbers. The third criterion is more reliable, however it is computationally more demanding and often requires abundant data. Convergence is usually based on the error function,  $\rho$ , exhibiting deviation of the predictions from the corresponding target output values such as the sum of squares of deviations. Training proceeds until  $\rho$  reduces to a desired minimum. The most commonly used stopping criterion in neural network training is the sum-of-squared-errors (SSE).
- Hidden layer size
  - In most function approximation problems, one hidden layer is sufficient to approximate continuous functions [87][68]. Generally, two hidden layers may be necessary for learning functions with discontinuities. The determination of the appropriate number of hidden layers and number of hidden nodes (NHN) in each layer is one of the most critical tasks in ANN design. Unlike the input and output layers, one starts with no prior knowledge as to the number and size of hidden layers.

## Radial Basis Function (RBF) Networks

Basically, radial functions are a type of functions that can be used any variant of model, linear or nonlinear, and any type of network, single-layer or multi-layer[88][66]. In Broomhead and Lowe's paper[89], radial basis function networks (RBF networks) are presented with radial functions in a single layer network [88][66].

The most common formula for any RBF is as blow:

$$h(x) = \phi((x - c)^T R^{-1}(x - c)),$$

where  $\phi$  is the applied function, e.g., multi-quadric or Gaussian [78],  $c$  is the center, and  $R$  is the used metric. Therefore, from the term  $(x - c)^T R^{-1}(x - c)$  the distance between the input sample  $x$  and the center  $c$ , in the metric  $R$ , is measured.

The RBF networks are known as a special case of a multilayer feedforward error-backpropagation network containing three layers [88][90][91]. The purpose of using hidden layer in these networks is clustering the inputs of the network, the nodes in hidden layer are named cluster centers. Accordingly, the selection between RBF neural networks and the BPANNs is based on the application and the raised tasks [92][88]. In comparison with BP networks, RBF train faster, however they are not as adaptable and are relatively slower for use [93][66].



## Classical approximation using Surface Response Models

Traditional curve fitting models address the dimensional data. In those cases, the fitting model concerns to a dependent parameter or variables with only one input.

$$Y=f(X)$$

However, such a fitting model is subject to challenges as standard fitting models. These challenges include:

1. Functionality of fitting
2. Plotting and analysis
3. The requirement for significant programming expertise

Meanwhile, in the Surface Response Modelling or Response Surface Methodology (RSM) what is considered is a statistical fitting technique to develop models from experimental or simulation data. It is mainly used to address the sensitivity of the parameter to input variables and provided an optimized response. The surface response provides a better representation of input/output parameters and their relation. The model is mostly applied to three-dimensional data (3D) and is for dependent parameters with two inputs.

$$Z=f(X, Y)$$

Therefore, it is possible to plot, make a comparison of different fits, do linear or nonlinear regression and perform a local smoothing regression and interpolation. It also enables the possibility to view the goodness of the statistical fits, illustrate the confidence interval and view residuals and remove outliers' meanwhile reaching a good fit using validation data [94] [95].

## Knowledge-Based Artificial Intelligence System Design

Knowledge is defined as the ability to perceive a subject or a domain, whether theoretical or practical. However, some believe it is knowhow which is acquired or provided by someone else called merely “an expert.” Experts are considered the most influential persons in a field, and they must be capable of defining knowledge in the form of rules applicable in problem tackling procedures. However, experts are considered as limited and are highly expensive. The goal of the AI is not only to provide the foundation that uses the expert’s knowledge or rules but also to extract these rules and knowledge by employing the conventional machine learning techniques, especially in an autonomous fashion.

In order to illustrate more desirably, we can consider the case of crossing a street which a robot is missioned to pass. In order to teach the robot to cross the street (a non-expert who lacks the knowledge and the rules of crossing a street), we are entitled to define and teach the rudimentary rules and knowhow in a manner where the robot subsequently be able to do this task autonomously.

These rules can be represented in form antecedents and consequents rules. These antecedents and consequents can also be manifested in the form of IF and THEN rules where what is coming after the IF condition is called antecedent and what is coming after the THEN condition is called consequent.

In this example we establish the rules as follow: IF the light was green, THEN the robot can cross the street and IF the light was red, THEN the robot must wait. In this case, the rules are provided by an expert like us humans. However, in general, the goal of designing an intelligent system is to extract these rules autonomously through empirical testing and by means of observations.

Subsequently, later future works would be to infuse other expert’s knowledge with the one extracted autonomously earlier. A unique example of such knowledge extraction and infusion exists in the baby human example, where at first the kids or the baby extracts these rules and know-how through its surrounding in the first phase of his or her life and by hearing and observing. Following this, he or she will obtain other rules and know-how through learning from other experts and infuses those rules with her or his own knowledge for a novel inference or reasoning.

To summarize this, it is entitled to discover the governing rules in a problem and apply those rules for the sake of intelligence and autonomy.

## Fuzzy Thinking, Neuro Fuzzy Rules and Inference

Real problems always require real knowledge and rules which are defined commonly through common sense. Also when tackling with such problems, there is no sharp distinction for categorization of the members within a subset. For instance, there is no sharp line to distinct a hill from a mountain, or there is no exact line to define a person as tall or short. This is while we apply phrases such as very, quite, seldom, a little and so forth to describe the quantity and greatness of our measurements.

As a consequence Fuzzy logic in some sense is a tool which reflects the way of human thinking and can model the words and decision making sense. Perhaps, this would lead to the design of an intelligent system. Fuzzy logic is also titled as multi-value logic. This concept deals with the mathematical representation of the terms where conventionally the logic is limited to true or false or 0 and 1. With this concept, the conventional logic is segregated into multiple sets and extend intervals into 0 and 1. Therefore, the possibility is introduced as a value to show the amount of truthfulness and falsehood of a statement of a parameter. Perhaps this value leads to a reasoning technique which is referred to as the theory of possibility. This theory, later on, addresses the vagueness and fuzziness in problems as a form of probability and assigns a corresponding value. In the following **Figure 38** the concept of expanding the inference and reasoning interval is illustrated briefly [96].



*Figure 38: Illustration of the Fuzzy and Boolean logic.*

Fuzzy logic is defined to represent knowledge (rules) based on the degree of membership rather than classical binary membership or crisp membership. Based on the definition, a Fuzzy set accepts the multi-values and different degree of membership. This is done by providing the possibility of having a spectrum. The interval generated is not a crisp 0 and 1. It is extended to totally false, entirely true, partly false and partly true. One must consider this fact that in crisp sets the membership of an element, for instance,  $x$  is either permitted as being a member or not being a member. This school of thought imposes a sharp border for the membership however the Fuzzy theory membership of a set is a certain degree [97].

Besides, in **Table 2** a comparison of membership is illustrated between a Fuzzy and crisp (Boolean) reasoning approach.

Tallness Inference			
Person	Height (m)	Boolean Reasoning	Fuzzy Reasoning
A	2.10	1	1.00
B	2.08	1	1.00
C	1.95	1	0.95
D	1.85	1	0.84
E	1.83	0	0.78
F	1.78	0	0.56
G	1.75	0	0.45
H	1.68	0	0.30
I	1.61	0	0.10
J	1.52	0	0.00

Table 2: Tallness membership comparison of the Boolean and Fuzzy reasoning based on the 1.85 cm criteria.

## Fuzzification and Rules Generation

Formation of Fuzzy sets or Fuzzification, in short, is a concept which is constructed by the membership function definition. Membership function shows the degree which an element belongs to a Fuzzy set, and it is possible to apply it for the transformation of crisp sets into Fuzzy sets [98].

By definition membership function of an element  $x$  in a Fuzzy set is defined by  $\mu(x)$ . If  $\mu(x) = 0$ , it states that  $x$  does not belong to a Fuzzy set at all. On the other hand  $\mu(x) = 1$  states that  $x$  belongs to the Fuzzy set entirely. Also, other values of  $\mu$  between 0 and 1 ( $0 < \mu(x) < 1$ ) illustrate that  $x$  is partially a member of the Fuzzy sets. In such cases the  $\mu(x)$  indicate the amount or the degree of membership.

The next step is to use the membership function definition and create the Fuzzy sets. However, it is obligatory to define the membership of elements in a Fuzzy set by employing knowledge extraction and rules generation beforehand. Perhaps, this is done by applying the following techniques:

1. Knowledge and Rules extraction through one expert: One expert defines whether an element is a member of a Fuzzy set or not.
2. Knowledge and Rules extraction through multiple experts: Multiple experts give their opinion about the membership of the elements.
3. Knowledge and Rules extraction through Neural Networks and Machine Learning: With this technique, different methods of machine learning are practiced to define the membership of elements in a Fuzzy set. The novelty of this approach is that it is done autonomously.

Besides, it is valuable to illustrate the mathematical relations working behind the fuzzification process. The membership function value  $\mu$  has a crucial role in this process. Set A can be a Fuzzy subset of set X only if we have the below relation [99]:

$$A = \{(x, \mu_A(x)) \mid x \in X, \mu_A(x) \rightarrow [0,1]\}$$

This illustrates that set A is a Fuzzy subset if the element x's membership function value for set A belongs to an interval of [0, 1]. Therefore, if the interval turns into a set of crisp value {0, 1} consequently we would have a crisp subset in such case. This emphasizes that a Fuzzy set can cover the definition of a crisp set by nature.

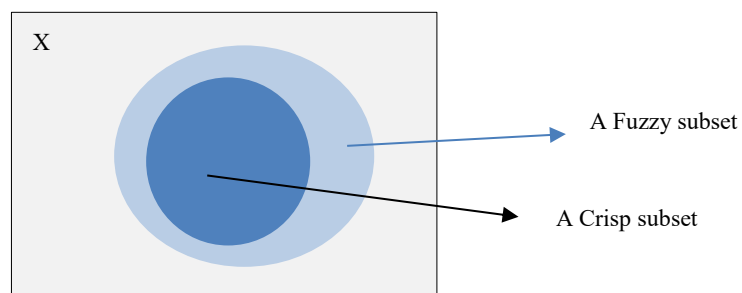


Figure 39: Illustration of a universal set X and its subsets as crisp and Fuzzy subsets

## Hedge Categorization and Linguistic Variables

One of the supremacy of the Fuzzy theory is the capability of using the linguistic parameters, which they lay the foundation of intelligent system designs. However, linguistic parameters are Fuzzy, and they can receive a range of value possibilities in their context of universal set definition. In this case, the Fuzzy theory can assist to address the vagueness of such parameters. Also. Fuzzy rules can be generated and applied for reasoning and inference applications.

To illustrate this more desirably, consider a parameter like the speed of a car. This parameter can get a range of values between 0 to 310 km (the universal interval or set). The linguistic words modifiers such as very slow, slow, average, fast and very fast are usually applied to indicate a measurement for this parameter although being Fuzzy. These modifiers and words accompanying with other modifiers such as quite, somewhat, more, slightly, less and so on together are called Hedges. Fuzzy sets are also constructed with hedge terms. Application of hedges relies upon the following categories:

- General purposes (hedges like: extremely, quite and very)
- Probabilistic purposes (hedges like: unlikely, likely)
- Fact-checking purposes (hedges like: true, false)

Human thinking can be represented with hedges. Hedges can act as operations too. They can do the following operations namely as the concentration of dilation. Modifiers such as very and extremely can act as concentration. Whereas more and less can act as dilators. This means with these modifiers the membership interval can cover a larger or a shorter range accordingly.

Meanwhile, there exists a mathematical relation for each hedge. Since they influence the membership function. Therefore as an example, we present the following linguistics as follow[100]:

$$\text{Very} = \mu^2 \quad \text{Extremely} = \mu^3$$

And for

$$\text{More or Less} = \sqrt{\mu}$$



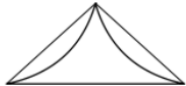





A little	$[\mu_A(x)]^{1.3}$	
Slightly	$[\mu_A(x)]^{1.7}$	
Very	$[\mu_A(x)]^2$	
Extremely	$[\mu_A(x)]^3$	
Very very	$[\mu_A(x)]^4$	
More or less	$\sqrt{\mu_A(x)}$	
Somewhat	$\sqrt{\mu_A(x)}$	
Indeed	$2[\mu_A(x)]^2$ if $0 \leq \mu_A \leq 0.5$ $1 - 2[1 - \mu_A(x)]^2$ if $0.5 < \mu_A \leq 1$	

Figure 40: Illustration of the hedges, corresponding mathematical relations and graphical representation [100].

## Ruled-Based Systems

The rules as mentioned before usually manifest in different forms and have perhaps corresponding inference. When tackling with complex systems, the fuzzy rules can be suggested as a means of analysis. Fuzzy rules again would follow the same format of conditional statements. Where they can be presented as IF-THEN rules. Perhaps the antecedent and consequent of such rules can be presented in the form of linguistic variables. One of the significant things to mention is the difference the traditional conditional rules with Fuzzy rules. This is manifested especially within the trigger rules. Where in classical rules binary logic and the Fuzzy rules trigger the condition are triggered with Fuzzy logic meaning that the condition can get a range or degree for the true conditions. Rules will be evaluated and get a degree of true condition, and they can be fired accordingly as a degree or to put it short partially. This process of inference is called merely a monotonic selection. As it is illustrated below, based on how much true the condition of being tall is a corresponding weight is assigned and indicates the amount which this true. For instance, this states that IF we have a tall man, THEN the weight would be heavy and a corresponding value will be assigned [100].

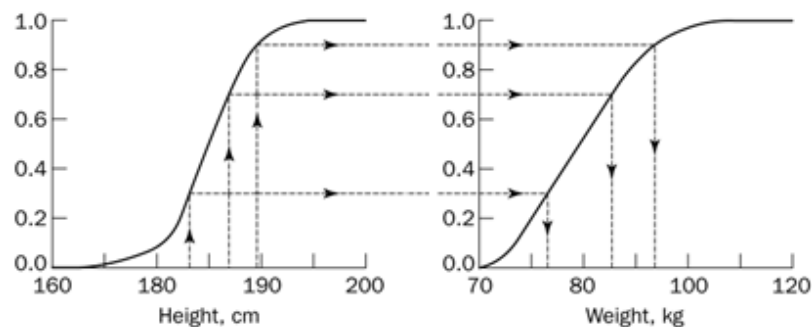


Figure 41: Illustration of Monotonic relation of the Height and Weight.

Rules can have multiple antecedent and multiple consequents. They can be mixed with Conjunctions as AND and Disjunctions as OR.

IF	Study duration is long
AND	number of students are large
AND	funding resource is inadequate
THEN	there is a high risk.
IF	Shower water is Hot
THEN	Hot water valve degree decrease
	Cold water valve degree increases



After the combination of these rules in order to reach a variable with crisp value Fuzzy system requires to sum up all Fuzzy outputs and obtain a single Fuzzy answer. Next, by a Defuzzification process, a numeric value would be assigned to the Fuzzy output.

The fuzzy inference systems apply three mostly used inference techniques as follows:

1. Mamdani
2. Sugeno
3. Zadeh

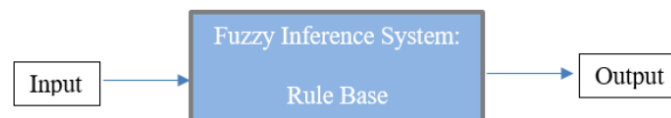
## Mamdani Inference Systems

This inference technique is one of the techniques which is used very often. This technique is presented by Professor Ebrahim Mamdani and was used as a tool to control a steam engine and a boiler in 1975 [101].

## Constraints

Construction of a fuzzy system include:

1. Determination of the inputs and outputs
2. Creating Membership Functions
3. Create rules
4. Result Simulation



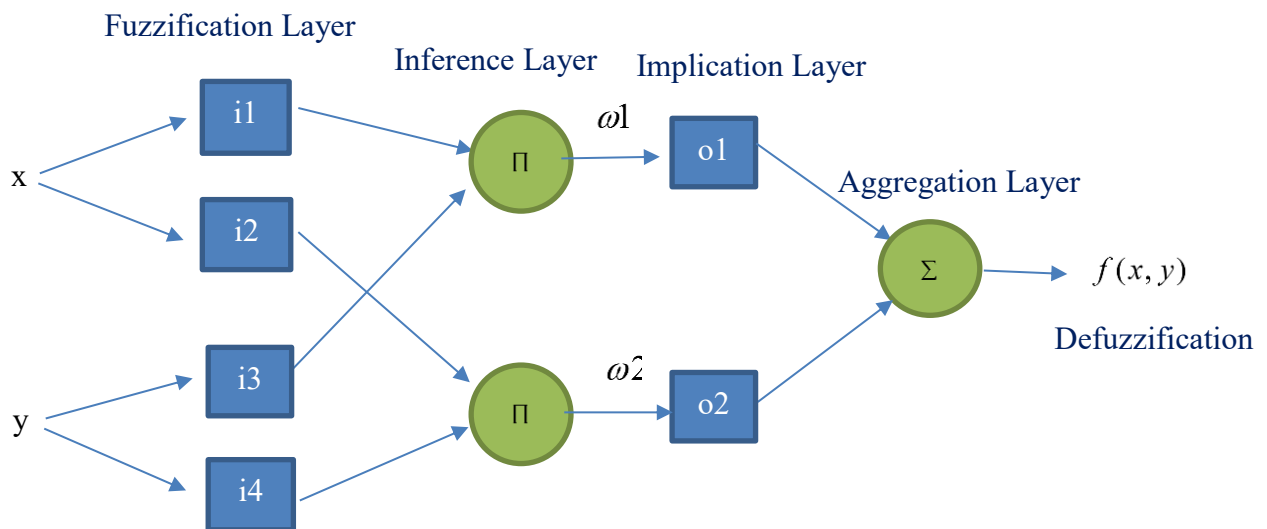
## Adaptive Neural Network

Fuzzy Neural Networks are widely used in many applications. Accordingly, a branch of neural networks classes referred to as adaptive neural networks are applied, where they are regarding inference equal to Mamdani inference systems. These networks construct a system which is referred merely as Adaptive Neuro-Fuzzy Inference Systems hereafter called ANFIS. However, the ANFIS can be merged by Mamdani inference engines and construct Mamdani Adaptive Neuro-Fuzzy Inference Systems hereafter as M-ANFIS. Among the inference engines, the Mamdani and Tagaki-Sugeno are between well [100].

Each M-ANFIS is going to be presented within five layers.

1. Fuzzification
2. Inference
3. Implication
4. Aggregation
5. Defuzzification

If X and Y be the input of the networks, an M-ANFIS is going to be constructed accordingly.



Rules: IF  $x=i1$  and  $y=i3$  THEN  $o1$

And: IF  $x=i2$  and  $y=i4$  THEN  $o2$

## RESULTS AND DISCUSSIONS

### Case study 1: Training of a data-set using artificial neural networks

The primary objective of this case study is the illustration and prediction supremacy of the artificial neural network with consideration and illustration of the impact of many data samples on the prediction accuracy in them. In this case, a multilayer perceptron network is trained with two different learning algorithms and also with consideration of the different number of hidden layers and nodes. Meanwhile, the training, validation and testing curves of proper training are illustrated and compared in different cases of training algorithm. Here, the Levenberg Marquardt algorithm hereafter referred to as LM, and also Bayesian Regularization referred as BR are studied within, and the results of training are compared accordingly.

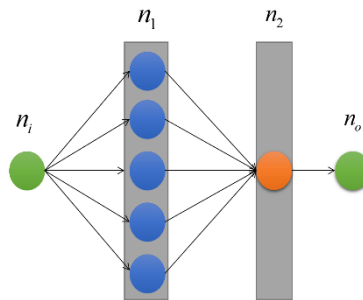
Moreover, other training algorithms are mentioned briefly, and their pros and cons are compared consequently. Next, the networks' performances are studied in terms of generalization with the presence of noise. The data-samples for training the networks are divided into multiple groups, and the results of the training are discussed after.

As mentioned earlier regression analysis and function approximations are the practice of supervised learning. In this case, a set of inputs referred to as  $x$  and  $y$  is provided as training data samples, and a corresponding result as output  $z$  is estimated using artificial neural networks. In this case training and learning criteria such as the number of hidden nodes, the number of hidden layers within the category of supervised learning method are compared. This assists in a genuine understanding of the best approximations and also help to understand the meaning of performance curves for the future case.

In this case, two data sample is created using the following code and with a known mathematical equation. Next, this data is provided to a neural network with the goal of checking the performance of the ANNs and their ability to predict the trend inside the data. Also, the number of data samples will indicate the amount of error in prediction while how this can impact on the training process. In this case, the dataset is divided into three groups. Training data set which contains 70 percent of randomized selected data. The validation data set is allocated 15 percent, and the same ration is kept for the testing datasets. This means that if a data set has 100 data samples 70 samples are used for training purpose and 15 samples for validation and 15 samples for testing. The performance of the networks is compared with their residuals and by using Mean Squared Error note. Shortly referred as MSE later. The root of the MSE as Standard Error can also be used for comparing the quality of the networks.

As illustrated in **Figure 42**, the topology of the neural network for applying function approximation on this data set is illustrated. The network benefits from two hidden layers where each layer encompass at least 6 hidden nodes **Figure 43**.

When applying the Levenberg Marquad or Bayesian Regulation learning algorithms the lower band and higher band of the number of neurons in a layer can be set using the following rule of thumb relations and this is illustrated approximately in context and also is usually obtained by trial and error procedure to find the optimal hidden nodes and layers where the networks will provide a perfect performance and fitting while not having any generalization or overfitting issues [66]. Though the lower band is obtain through a test and trial approach but there is an approximate rule of thumb for defining number of minimum number of neurons in a network. If we consider a Network as follow:



*Figure 42: Illustration of the architecture of a neural network with two hidden layers.*

Then we would have:

$n_i$	$n_1$	$n_2$	$n_o$
Input layer neuron number	Layer 1 neuron number	Layer 2 neuron number	Output layer neuron number

The number of nodes in layer 2 or must be equal to, the number of neurons in the output layer then we must have:

$$n_2 = n_o$$

Meanwhile, with  $k$  samples of data at hand, the number of known information or equations would be as follows where it would depend on the input neurons and output neurons and the number of samples.

$$K \times (n_i + n_o)$$

Each node would generate an unknown mathematical equation. For instance, in the input and first layer, the number of unknown equations would be:

$$n_1 \times (n_i + 1)$$

Moreover, for the next layer and output layer the number of unknowns must be as with consideration of layer two as one node:

$$n_o \times (n_1 + 1)$$

Hence the number of unknowns for one hidden layer network with n neurons would be:

$$n_1(n_i + 1) + n_o(n_1 + 1)$$

Regardless of activation function or transfer function and based on the number of known and unknown equations built by the number of neurons in order to solve this network the following equation must be satisfied.

Number Unknown equations < Number of known equations

$$n_1(n_i + 1) + n_o(n_1 + 1) \leq K \times (n_i + n_o)$$

Thus, from the above equation, it would be inferred that the number of the neurons in the first layer should be at least as:

$$n_1 \leq \frac{k(n_i + n_o) - n_o}{n_i + n_o + 1}$$

This would define the upper bound number of neurons for each layer meaning that if the training includes the more significant amount of nodes larger than this threshold the network would face saturation either bad training or overfitting to the data itself.

Meanwhile based on empirical methods for the lower band in order not to have an under-fitted model the number of neurons in the first layer must be set as below:

$$n_1 \geq 2(n_i + n_o)$$

The number of nodes for the training of this case study is obeying this rule, and it is possible to illustrate the under fitting and overfitting problem by adjusting the number of nodes larger or less than this value.

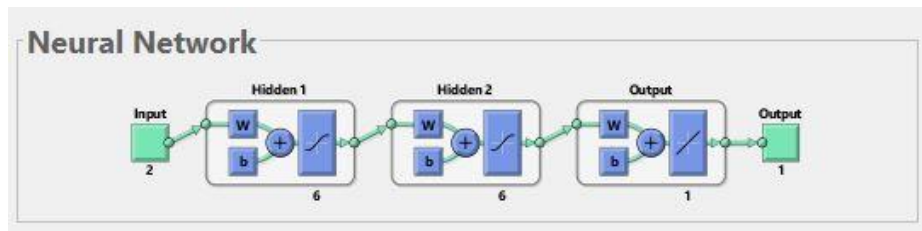


Figure 43: Illustration of network architecture.

Besides, the activation function which is used for training, in this case, is a tangent symmetrical sigmoidal function. This activation function is usually used for catching the non-linearity in the systems and squashes the elements of an input matrix into the interval of -1 and 1 with an s-shape format. The output layer also will use a transfer function usually a poslin or purelin function to map the outputs of layers linearly to the output, but this depends if negative numbers are also included in the data sets or not.

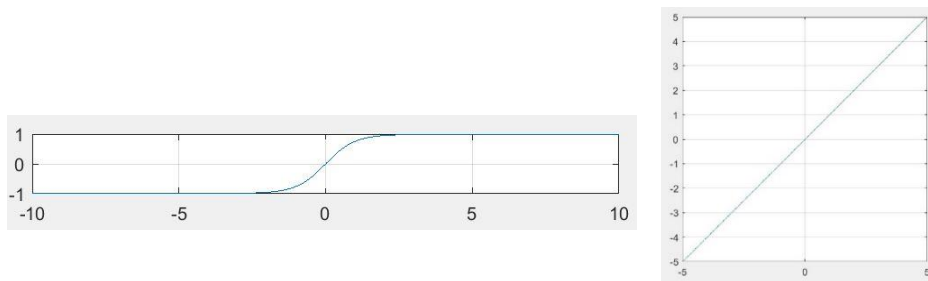


Figure 44: Illustration of a sigmoidal and a purelin transfer function.

The other type of transfer function is list as below which based on different topologies and application they are used:

1. Elliot sigmoidal
2. Hard limit
3. Competitive
4. Logarithmic sigmoidal
5. Radial base
6. Normalized radial basis
7. Saturation linear
8. Soft max
9. Triangular basis

As stated before in the first data set 120 points are used predictions. In **Figure 45** the ANN prediction out (blue) curve and the real target values that the ANN must have predicted is shown. This indicates for the starting and ending and even some areas the prediction of the ANN is not as very satisfying. This is more evident by getting the residual error of the inputs and outputs for each point. The second left down the curve in **Figure 45** shows the amount of difference the output of the ANN and the real data have. Finally, the histogram of the errors is shown in next right down diagram in **Figure 45**. This shows the integrity of the errors and their value around 0. If all the error residual have an integration around the zero points, the model is said to have a perfect fit and perhaps would have an MSE. The MSE of this network performance is  $4.832128931550165e-04$  which would have a standard error of 0.022.

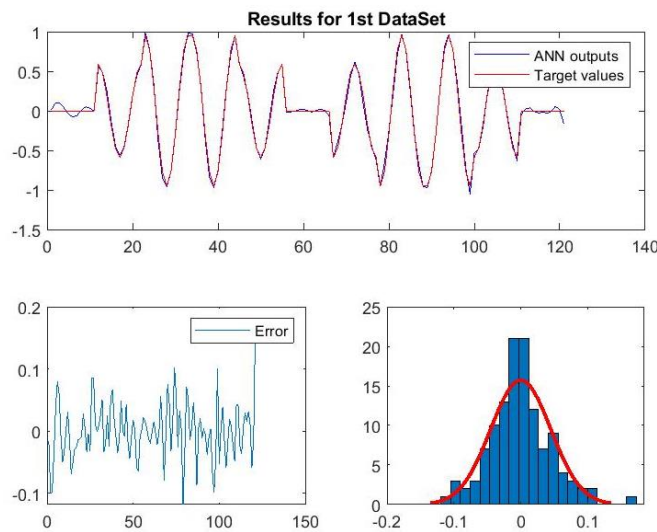


Figure 45: Illustration of the ANN prediction output and the expected target values for 120 data points.

Meanwhile, in order to check the performance of the ANN, three curves of MSE (the vertical axis) for the training data sample (blue curve), validation (green curve) and test samples (red curves) within each iteration noted as EPOCHs (the horizontal axis) is illustrated. These curves can carry the information for under-fitting and lousy training and show the performance of learning in the training process.

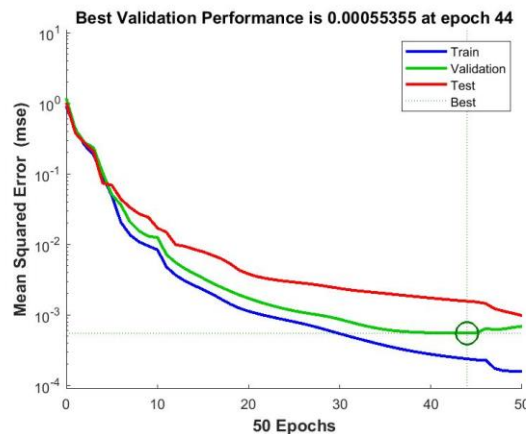


Figure 46: Illustration of the performance curve for the 120 data samples.

As it is shown, the training process is stopped with the validation checks. The validation check is the process which evaluate the amount of MSE in each iteration and if the networks fail after six times the training process stops. This means that the tanning of the network is not showing any MSE reduction and six consecutive evaluation for MSE is higher than the previous one. The training can stop based on time, gradient, performance, Mu parameter (mainly used in Bayesian networks) or validation checks. If we consider the training similar to the case of teaching to some students the validation checks can be similar to mid-term exams while the test is similar to the final exam. If a student fails consecutively in mid-term exams, it is very likely to fail on the final exam too. This indicates if the number of validation checks being set as a value. The network is allowed to fail as that much. This can lead to the under-fitting phenomenon in neural networks. In **Figure 47** some of these evaluation criteria are listed.

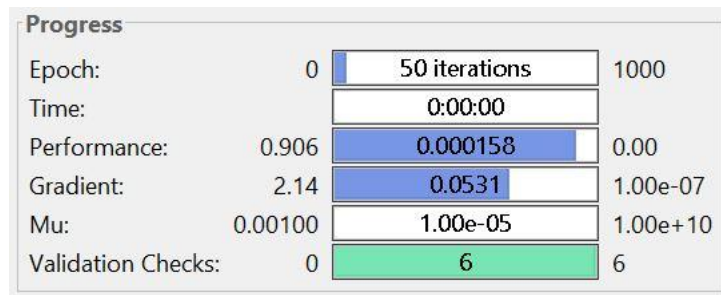


Figure 47: Illustration of the progress parameters for the 120 data samples.

In the next step, the amount of data samples is augmented to 2601 samples, but the training is following a similar architecture as it was used for the 120 samples. For this case, the amount of iterations increases to until it reaches the 1000th EPOCHS. At this stage the training stops and as it is shown in **Figure 48** the training performance shows that the amount MSE is decreasing very well as the train, validation, and test MSE curves follow each other. This indicates a good training process, but still, final validation test with other data out of the data set is required to check the generalization of the neural networks.

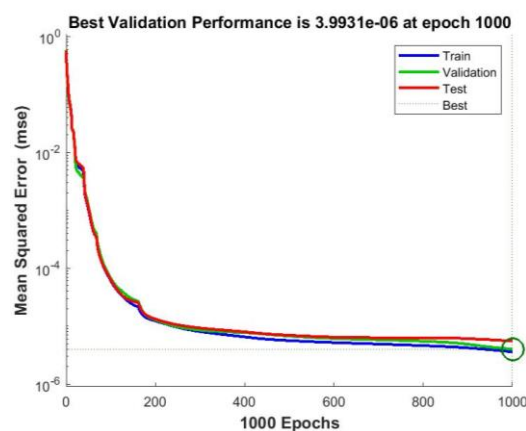


Figure 48. Illustration of training of 2601 data samples.



As illustrated in the following **Figure 49** the progress diagram of the ANN2 for the second data-set indicates that the network is run for 1000 EPOCHS and the network is stopped when it reaches the desired value. The value of the MSE for the network is equal to  $4.399819401461010e-07$ , and the standard error is  $6.633113448043091e-04$ .

Progress			
Epoch:	0	1000 iterations	1000
Time:		0:00:05	
Performance:	0.558	$3.58e-06$	0.00
Gradient:	1.04	0.000239	$1.00e-07$
Mu:	0.00100	$1.00e-06$	$1.00e+10$
Validation Checks:	0	0	6

Figure 49: Illustration of progress diagram for 2601 samples.

As it is shown in top section of **Figure 50** the network output (blue curve) is quite overlapping the target value which the network is expected to predict. Meanwhile, in the left corner side of the figure, the difference of the output is shown for each data sample, and the amount of error residual is calculated. The histogram diagram of the error on right corner side of **Figure 50** shows the accumulation of the error as approximately two bins of 1000 and 800 normalized around the point zero. This indicates the network is capable of predicting the values without knowing the corresponding mathematical equations.

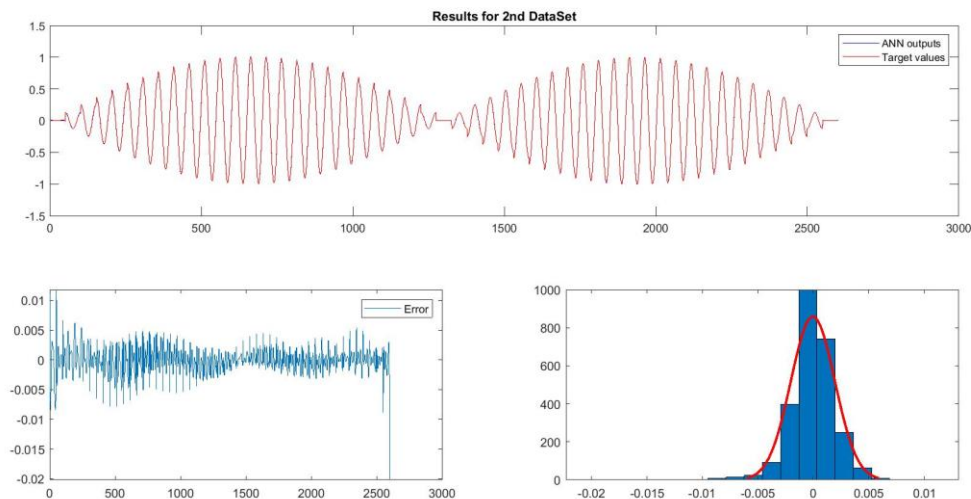


Figure 50: Illustration of results for the second data set

In the next step, the performance of the networks is evaluated by adding white noise to the data. This would help to assess the performance of the networks to see if they catch the noise or not. This is corresponding to the problem of over-fitting of a network to noise, and in such cases, the network cannot generalize to new situations or data other than the data-set which is used for the training.

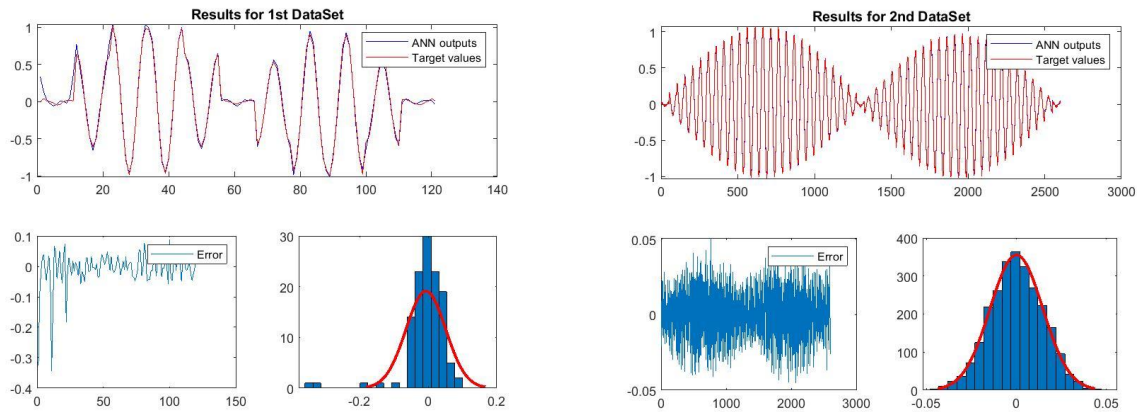


Figure 51: Illustration of the results for the two data-sets with presence of a noise.

The results of the analysis for these networks in terms of Mean Square Error and Standard Error same as root mean square error (RMSE) shows that the networks with higher sample rate can predict the and perform better while they are also able to generalize efficiently to other data spaces.

MSE	Data-set1	Data-set2	Standard Error	Data-set1	Data-set2
Without noise	4.83E-04	4.40E-07	Without noise	0.022	6.63E-04
With noise	0.003398	2.13E-04	With noise	0.058294	0.014599

Table 3: The comparison of the two networks in presence and absence of the noise.

In the case of increasing the number of hidden layers abruptly the network fails to train appropriately and error does not decrease by a sufficient gradient value. In such cases, the risk of overfitting and generalization will exist, and also the residual bin in the histogram of error will reduce. The **Figure 52** shows the failure of the training after ten iterations and indicates the gradient of learning is decrease immensely to a minimal value.

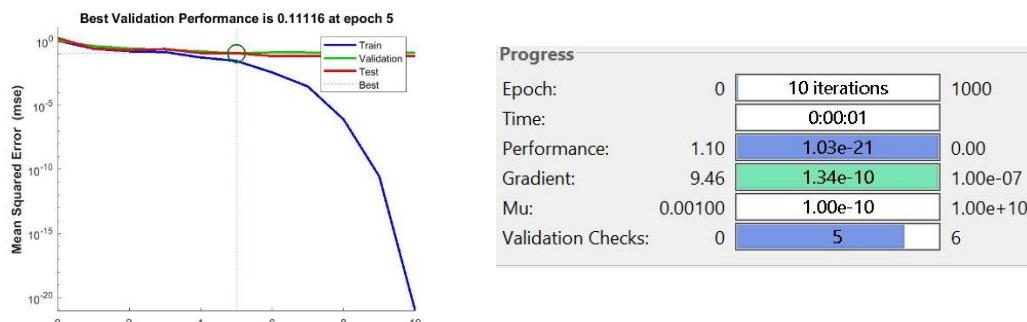


Figure 52: Illustration of a training with high amount of hidden nodes and layers.

The results of the training are shown in **Figure 53** diagram. The standard error for the 1st dataset increase to the value of 0.233 and the 2nd dataset the standard error will become 0.0143. With comparison to their optimal value, they indicate a significant variance.

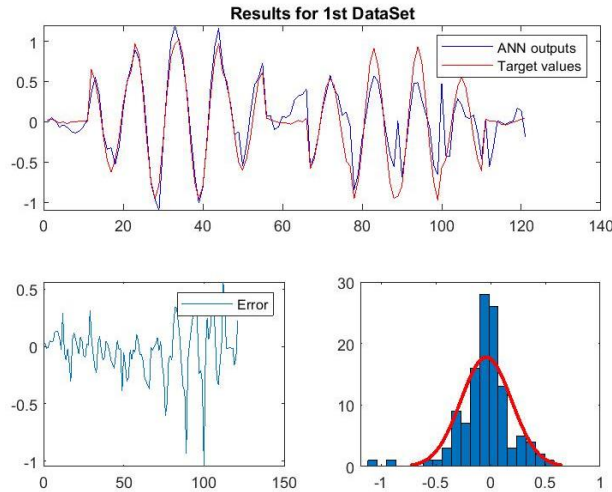


Figure 53: Illustration of the network results with high amount of neurons and hidden layers.

On the other hand, one of the classical approaches for analysis of such dimension of parameters is the surface response modeling technique and by using polynomial regression or locally weighted scatterplot smooth model shortly referred as LOWESS which by using the data samples, applies a locally weighted linear regression to smooth the data and provide a fitting model. This technique is referred to as weighted since the machine sets the weight function for the regression.

Besides it is possible to use a robust solution that can make the process of smoothing strong toward any outliers (data with high error deviation). With this method, it is also possible to compare the results of the ANNs and classical approaches. The **Figure 54** diagrams show the surface response model for the 1st dataset. The model which is obtained is having an RMSE= 0.02524 by using a Linear LOWESS technique. This value is obtained by a 1% span, and if it is changed to 10%, the RMSE rises to 0.09589 value. For the second data-set (right diagram) the RMSE = 0.008885 where with 10 % span it changes to RMSE= 0.07737.

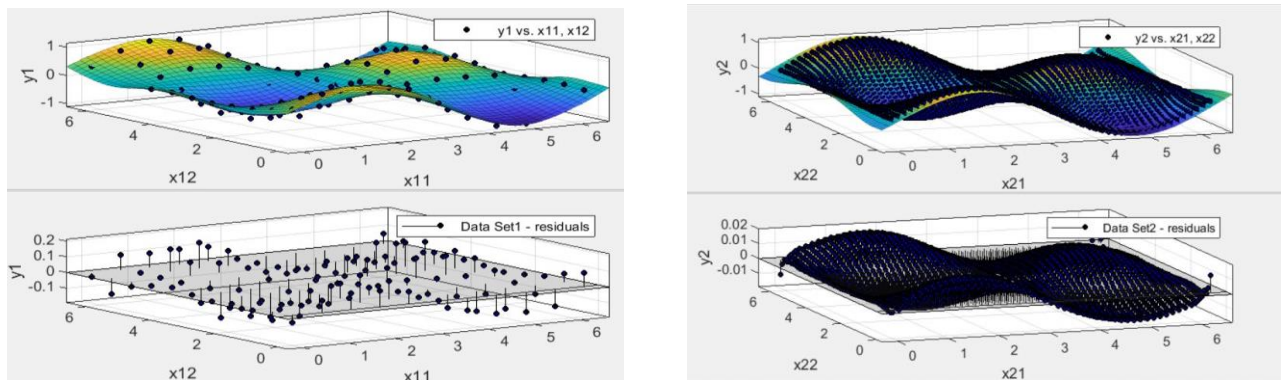
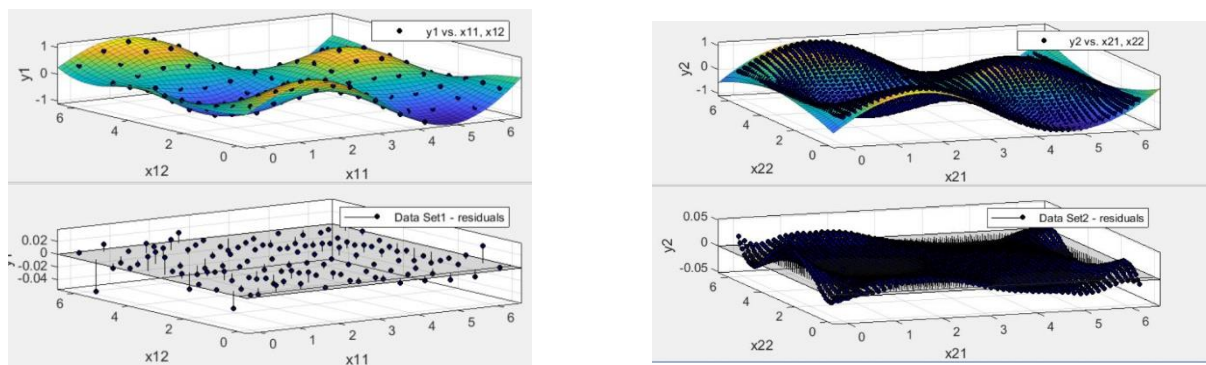


Figure 54: Illustration of the surface response model using a linear LOWESS with 1% span and no robust.

However, the quadratic analysis shows the  $RSME = 0.01237$  but without any robust. The quadratic or linear is the type of the polynomial model which is applied for calculating the regression models. Also, the span defines how much percentage of the data samples in a set is used by neighboring data points for defining the smooth values and fitting. By reducing the percentage, the data points which are used will be closer to the data this increase the risk of overfitting. Also, as mentioned the robust is an option of using robust weights for function with the purpose of estimation and resisting to outliers.

As it is illustrated in **Figure 55**, the surface response for both data-sets is shown. The left diagram indicates the response model besides the error residuals for the 120 sample data and the right diagram demonstrate the 2601 data set. As mentioned the 1st data set has the  $RMSE = 0.01237$  And the 2nd data set has  $RMSE = 0.009153$ .<sup>14</sup>



*Figure 55: Illustration of surface response modeling on the data sets.*

One of the critical things to address is the training functions. In this case, the training algorithm which was used is the LM algorithm. However, in most of the cases, it is not simple to address the fact that what type of training algorithm provides the most optimized solutions. This is depending on the number of points in a training set plus the complexity of the problem, weights, biases, the error top goal, and even the type of the application which the network is used for. But in general, the networks with feed-forward topology uses almost six options. The LM function is assumed to have a faster convergence, and in many cases, it can provide a lower MSE. However, the LM algorithm and its performance is susceptible to reduction when the amount of weights decreases, and it is not very decent for other applications such as pattern applications. Meanwhile, another algorithm which is used and constructed in the heart of the LM algorithm is the BR algorithm. BR algorithm will adjust the weights and biases in such a way that to provide a superior generalization for the network outputs [102][103].

<sup>14</sup> White noise is added to this data sets.

## Case study 2: Sphere and Fluid a Knowledge Based Application

In this case study, which the results have been published in the IDETC 2018 conference in Quebec [104], the application of the ANNs as a knowledge-based neural network (KB-ANN) is studied on a data collected for a flow around a sphere, Reynolds number ( $Re$ ), Drag coefficient ( $C_d$ ) and invariant or the Pi numbers. The primary objective of this case is to test the methodology of knowledge-based-artificial neural networks. This study. Subsequently, after that, the results are discussed in terms of training and generalization by using the MSE notation. The goal is to provide a topology which is benefiting from the prior knowledge and infuses this knowledge into the neural networks topologies. This case is benefiting from the knowledge of the DACM approach presented earlier and by defining the system behavior in terms of causal relations. One of the main features of the KB-ANN is the way of implementation and infusion of the knowledge in the network which indicates better working results with few data points.

As stated in [104], a sphere is positioned in a flow of fluid which is exposed to a drag force  $w$  and is subject to friction of fluid with the density of  $\rho$ , a viscosity of  $\nu$  and with the speed velocity of  $V$ . This sphere has the diameter of  $l$  which constructs a non-linear physical problem for solving.

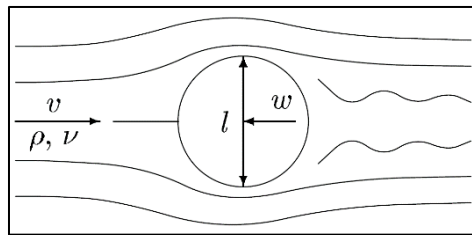


Figure 56: Illustration of the Sphere in the flow fluid with demonstration of the drag force.

In this classic example, the goal is the estimation and modeling of the drag force by utilizing of velocity. Therefore, an attempt is made to present the functional architecture of the neural network capable of modeling this phenomenon. In [104] the role of invariants on the construction of the network topology is described which the network is infused with a prior knowledge. Since in this problem there is little known about the relation of the  $Re$  and  $C_d$  the invariants of  $\pi_1$  and  $\pi_2$  the goal is to construct an ANN which with there is a possibility to provide a fitting model. Therefore by using the experiment's data obtain from the relations between variables it is possible to construct the rules for achieving the target (21). With the consequent knowledge of the model for the transformation at that point, it would be possible to predict the desired outcomes and predict the drag force. For this two networks are proposed as a fully connected network and a KBANN where the topology is using the prior knowledge.

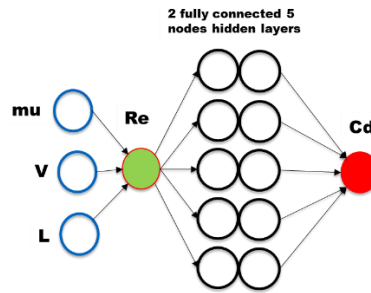


Figure 57: Illustration of a KBANN network constructed using dimensional analysis

For this case in order to obtain a novel model, it is required to normalize the data points to keep the proportionality of the value to each other. The normalization is conducted to map the data-point to the interval of 0 and 1. This while other normalization techniques could be used to the task. For instance application and using of standard deviation or minimum and maximum of the data-points. As illustrated in the **Figure 58** different layers and weight (models) were used in previous attempts to predict the behavior between the input data point and outputs. The regression models are claimed to have different MSE due to the initial randomized weights. This is due to the reason that since neural networks use initial randomized weights and use randomized data for training, different results would appear with each training. Therefore, one good remedy for this is to run the network for multiple time and check the validity and find the minimum error cost that the networks can provide. Unquestionably with an excellent generalization to other data points.

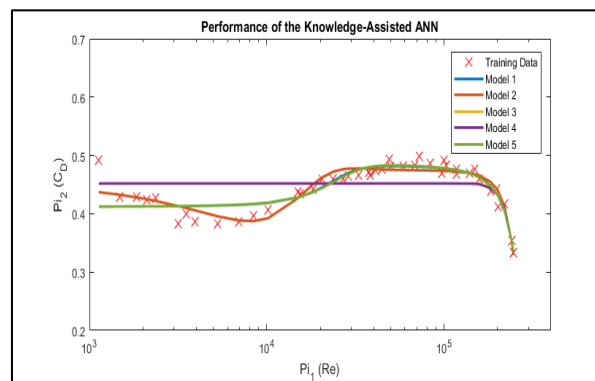


Figure 58: Illustration of precious attempt to predict the model with models sensitivity to primary conditions.

As illustrated in **Figure 59**, the prediction of the Cd based on Re is shown. A supervised learning approach is used by providing training, validation and test data points randomly. The blue circles are the ANN prediction points versus the Red Cross which is the target points and as what the ANN must have predicted. The training is conducted 25 times with different random initial starting points.

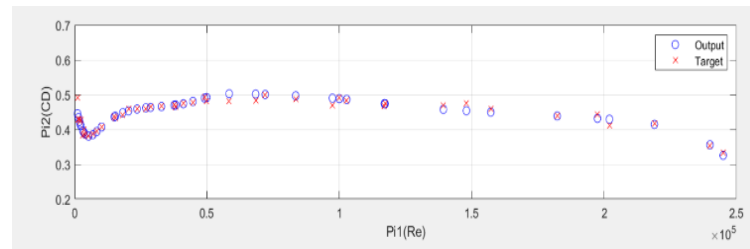


Figure 59: Illustration of the results of prediction by KB-ANN.

The amount of error for predicting each point is illustrated in **Figure 60** and the amount of test error the is equal to  $MSE=0.000128$ .

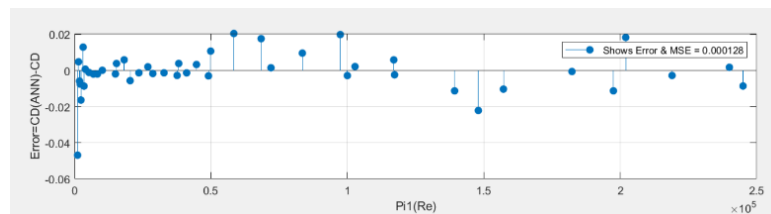


Figure 60: Difference of the real data points (target values) and the predicted data by the network (output) values for the KB-ANN.

In order to evaluate the performance of this network, it is necessary to evaluate the regression analysis of the network. The **Figure 61** illustrates the training quality besides validation and test. As it is shown the target values and output values for each batch of data indicate a good match. Especially for the test data points, the relation of the outputs and target values are close to each other by the R indicator equal to  $R=0.94952$ . The more this value is closer to 1 it indicates a better fit.

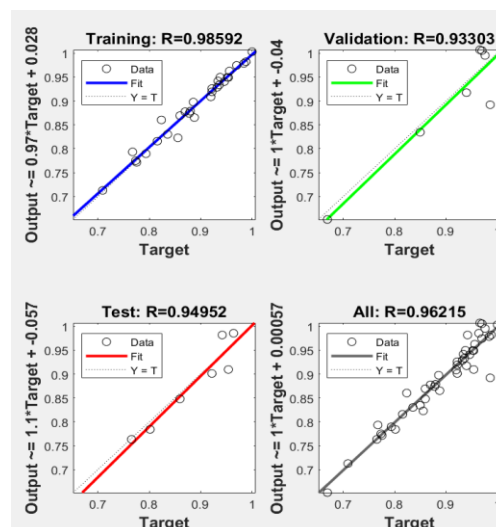


Figure 61: Illustration of performance of training process for training, validation and test data points.

However, one must be aware of the overfitting issue where the networks can catch the noise of the system, and this is not desired. The dashed black line present in the diagrams is the desired line for each of the training batch where the linear relation of the target and output values must fall on. Also, for this training, the dataset is portioned randomly to the batches of 70% for training points and 15% each for validation and testing data points.

Moreover, in order to evaluate the network's outputs quality, it is necessary to observe how the performance curves for each element of the network is behaving. As it is shown in **Figure 62**, the best performance validation for the KB-ANN is equal to  $1.8789 \times 10^{-3}$  at the 55th iteration. Though the network continues training until 65 iteration but from 55th iteration, the networks start to behave as overfitting. At this stage, the network stops the training with the best validation obtain thus far.

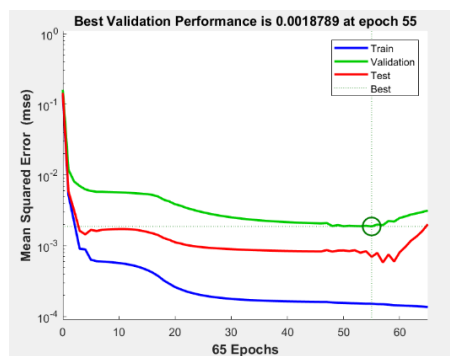


Figure 62: Illustration of performance curves for the KB-ANN network.

On the other hand, a conventional fully connected ANN is constructed for modeling the problem and for the reason of comparison to the knowledge assisted network. For this topology, the inputs include the physical properties of the problem, and the output is the drag coefficient. The main difference of this network with the knowledge assisted network is the application of the prior knowledge in this topology (Pi numbers calculated as in [104]).



The conventional network topology includes two hidden layers but requires at least six nodes. The activation function which is used for this network is a sigmoid function. The output of the network is the drag force  $W$  which with the known values for other parameters, it is possible to calculate the drag coefficients. **Figure 63** demonstrates the topology of the conventional ANN to model this problem.

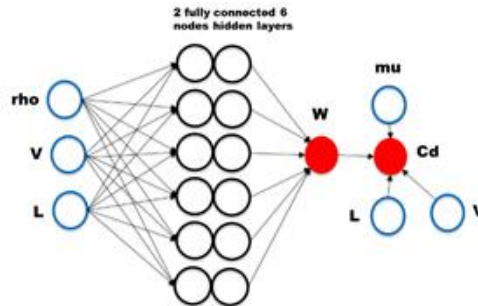


Figure 63: Illustration of conventional ANN structure model for this problem.

Normalization of data points in this process will assist the transfer functions not being trapped in the saturation zones. Unquestionably, after when the training is ended the values must return to the original values by de-normalization function. Because the training points are minimal in quantity, the network's output results are sensitive to initialization. This could lead to diverse results. However, the best results are intended to be used for later applications. Hence, this model is trained with 25 different initial points, and the solution with the lowest cost is nominated.

This model is also presenting a decent performance but in comparison to the KB-ANN network the initial points (Reynold's numbers) are showing higher error and will lead to higher residual for the network. The MSE of this network is equal to  $MSE=0.021067$ . In **Figure 64** the fitting results of the network are displayed. In the left diagram the outputs, blue rings are the ANN's output, and the red crosses stand for the target or actual values that the network must have anticipated.

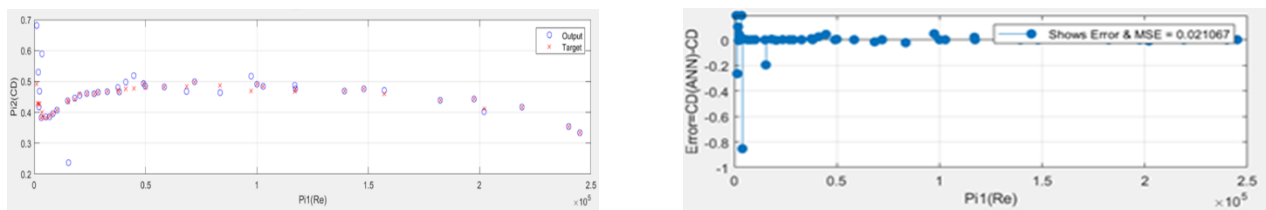


Figure 64: Illustration of the residuals and the prediction results of the conventional ANN.

The regression plot **Figure 65** of the networks indicates the evaluation of the conventional ANN networks and the quality of the training on the provided data. As illustrated the training data points will match the output points with the indicator of  $R=1$ , and for the validation data points the output and the targets match each other at a considerable rate. This applies to the test data points. Though these results are more promising for the conventional networks with a high amount of nodes, however, this is required to be checked for the overfitting and generalization phenomenon.

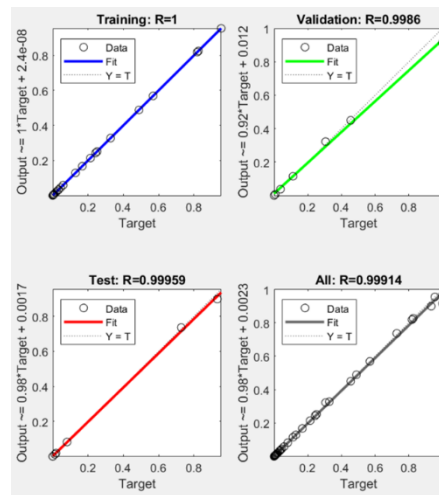


Figure 65: Illustration of regression plot for the conventional network.

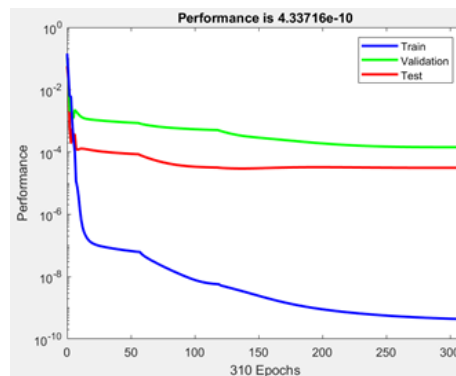


Figure 66: Illustration of the performance curves for the conventional neural networks.

As shown in **Figure 66** the best validation for the conventional networks is  $4.33716e-10$ . The best result of iteration is made at EPOCHs 310. This is while the best validation value for the KB-ANN (green curve) is obtained at the 55th iterations. This means the KB-ANN has been trained faster with a lower computational cost. Though the regression plot **Figure 65** for the conventional ANN is better with the analysis of the performance curves, it is evident that the conventional network is more susceptible to overfitting and generalization issue with the same amount of training data points. This is evident also in the plot of predicted points for both networks and the target values that they must have estimated.

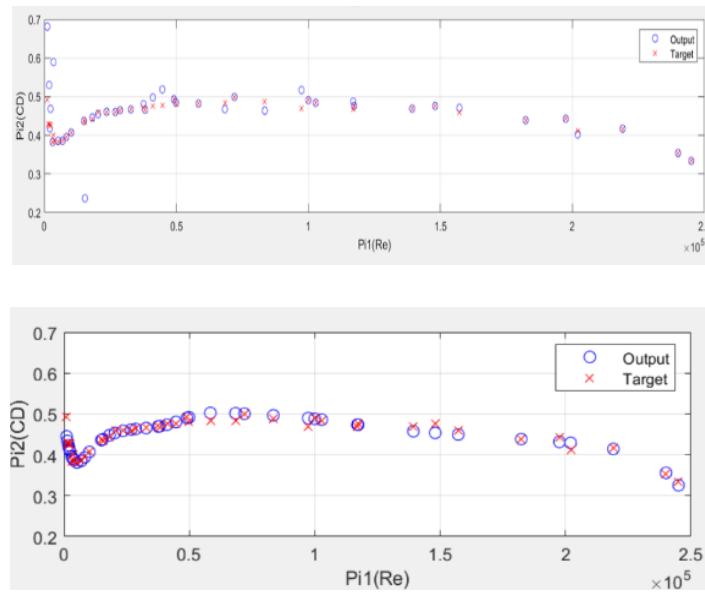


Figure 67: Illustration of the results for both networks and the comparison of the Conventional (Up) and KB-ANN (Down) data plots.

By consideration of **Figure 67** it is evident that in terms of prediction the KB-ANN is indicating a better fitting to the experimental data. In addition, the KB-ANN has shown better training despite a lower amount of iterations. This is while the number of nodes is less than the conventional network topology. Another significant result is that the knowledge provided by the Pi numbers is assisting the networks in better training while there is a fewer amount of data-points. The result of the KB-ANN is providing faster convergence due to fewer weights and this assists in terms of computation power. Meanwhile, from the point view of robustness, the KB-ANN is performing superior in terms of generalization whereas for the conventional topology due to having less amount of training data points it is more prone to error when exposed to new data-sets out the studied space. While there are 44 data points, the MSE for the KB-ANN network is less than the conventional one.

The limitation of this case study is manifested more in terms of the reduced data-set. This makes the topology of the neural networks more susceptible to the overfitting and under-fitting issues. As a consequence, the number of nodes and hidden layers must be selected accordingly and wisely to avoid such potential issues. In conclusion, the superiority of KB-ANN networks has been shown through this study and the details of this is accessible through the paper published and referenced as [104].

### Case study 3: Fused Deposition Modeling

This case study is the extended application of the KB-ANNs and is applied to a fused deposition modeling case as a branch of additive manufacturing. As mentioned earlier this case is also benefiting from the knowledge encoding techniques which is developed by Coatanea et al. [105]. The encoded knowledge of the fused deposition process is encoded and represented in the form of causality among the critical parameters. After that, it is simplified as a KB-ANN model where to be able to implement with ANNs. Functional representation is used to define the system behavioral. These functions are later used by DACM frameworks to define the causal rules existing in the system. Analytical approaches are used to define the dimensionality of the system and result in a colored graph and with the list of equations. Further analytical methods can be applied later to define the quality and quantity of the output of the system according to the obtained relations.

As mentioned earlier the FDM process deals with the thermal and energy processes and by the distribution of them ensures the bonding of materials and with the goal of a perfect deposition. In FDM several techniques and tools are used in order to deposit material in a novel fashion, and the subject of this case is to study the best approaches for delivering the material in place adequately.

In this case, the part which is printed is having a wall thickness of  $e=0.5 +0.05$  or  $-0.05$  mm, and it has the height of  $H_t=12 +0.05$  or  $-0.05$  mm. These ratios are fixed for the printing process.

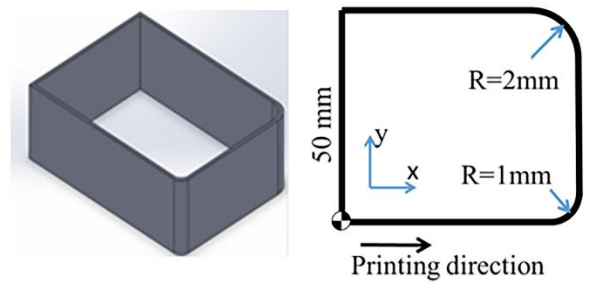


Figure 68: Illustration of the FDM case used for the study.

This case is promoting from the implementation of the experiments and testing the influential factors and after that defining the controlling procedures to address the latency in this system. Moreover, with the help of tools such as DACM framework the design of the KB-ANN network architecture is conducted and used for training the ANNs. Finally, the performance of the pieces of training is compared with the ANN classical networks to ensure the superiority of the infusing knowledge in modeling and optimizing this type of additive manufacturing process.

Initially, the printings are implemented by some selected process parameters and by using the slicing software such as Repetier. With an analysis of this additive manufacturing process. The main most influential parameters are recognized namely as the Height, Thickness and Mass. The notations for these parameters correspondingly are:

1. Height  $H_t$
2. Thickness  $e$
3. Mass  $M_t$

With the help of orthogonal arrays analysis, the performance evaluation of the experiments is achieved, and the level of latent parameters are kept accordingly low to minimize their effects on the targets. Correspondingly, the values which have been stated earlier for the height and thickness are nominated for deployment in the modeling. The remaining parameters such as

1. Layer Height  $h_i$
2. Nozzle Travel Speed  $TS$
3. Extruder Temperature  $T_{set}$

are also used in modeling this additive manufacturing process. The causal relations among these parameters are defined by applying the DACM frameworks. The causal relation between the most crucial variable is obtained while the green or independent variables are the input variables and there is full control of their adjustment. The blue or dependent variables are the ones which there is no possibility of a direct adjustment over them. However, there is a possibility to obtain them. The black or exogenous variables are the ones which there is less control on them and are imposed by the environment, and finally, the red or performance variables are the parameters of the system output which provides information about the performance of the system. The causal relation of the parameters wherever it was promising to obtain is constructed, and some relations are obtained using mathematical relations and also adjusted based on the optimal results extracted from the literature.

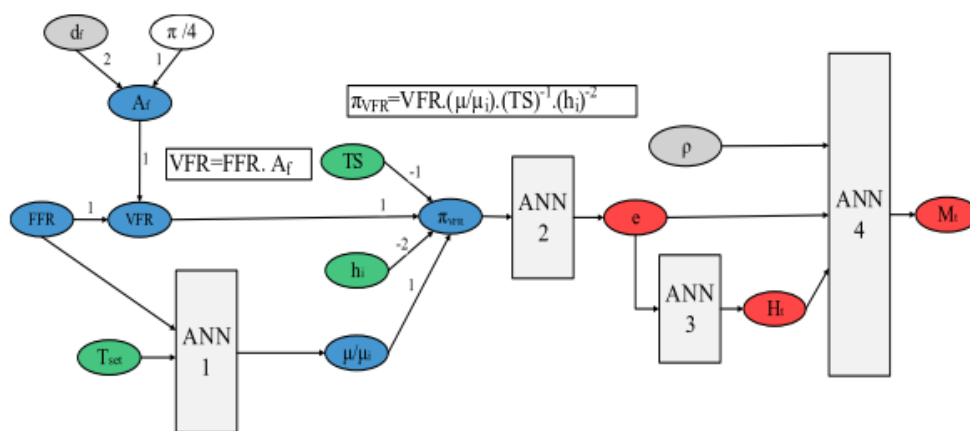


Figure 69: Illustration of the causal relations in the FDM process for the selected case study obtained by the DACM.

The hybrid topology of the KB-ANN is constructed with this fundamental assumption that to replace an ANN where the relation or casualty of the parameters is unknown. In such conditions, it is expected that the ANN by using available training data and where there is not any real data by using simulation procedure obtain existing relation among the selected parameters and pave the path for estimation and optimization of the additive manufacturing process, here as an FDM case.

Sign	Definition
$\lambda$	Relaxation time index
Af	Cross-sectional area of filament (mm <sup>2</sup> )
a	Dimensionless parameter describing the transition between the first Newtonian plateau and the power law zone
cp	Heat capacity (J/kg.K)
df	Diameter of filament (mm)
di	Diameter at ith section of liquefier nozzle (mm)
dx/dt	Nozzle velocity in x direction for dx (mm/s)
dy/dt	Nozzle velocity in y direction for dy (mm/s)
e	Intended wall thickness (mm)
FFR	Filament flow rate (mm/s)
hi	Layer height (mm)
Ht	Part height (mm)
k	Coefficient of conduction (W/m.K)
Li	Length at ith section of liquefier geometry (mm)
Mt	Part mass (g)
n	Power index
T0	Output temperature (°C)
Ti	Initial filament temperature (°C)
Tref	Reference temperature (°C)
TS	Nozzle travel speed (mm/s)
Tw	Wall temperature (°C)
$\Delta v$	change in nozzle travel velocity (mm/s)
$\theta$	Dimensionless temperature
$\mu$	Viscosity of polymer filament (Pa s)
$\mu_i$	Kinematic viscosity (m <sup>2</sup> /s) at reference temperature
$\mu_{inf}$	Viscosity at the infinite shear rate (Pa s)
$\mu_0$	Viscosity of fluid at zero shear rate (Pa s)
$\rho$	Filament density (kg/m <sup>3</sup> )
$\gamma$	Shear rate (s <sup>-1</sup> )
Fst	Force resulting from surface tension (N)

Figure 70: Illustration of the list of impacting parameters affecting the FDM process in this case study.

As illustrated in the Figure 70 above the list of all the impacting parameters and constant values are shown. With the help of the DACM framework and infusing the available knowledge such as mathematical relations, equations and empirical analysis Figure 69 is obtained. The constructing topology of KB-ANN network is going to follow this diagram. Whereas the classical ANN model will follow a fully connected topology as illustrated in Figure 71.

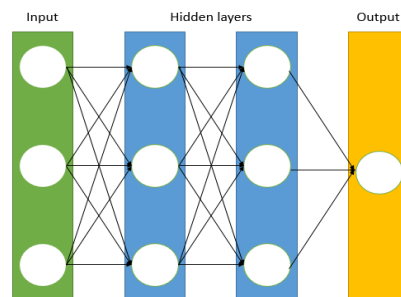


Figure 71: Illustration of classical ANN model used for FDM process.

The classical method of the ANN leads to generating three networks where they estimate the wall thickness, mass and the height of the part generated. The inputs of the ANNs are the nozzle travel speed, extruder temperature and the layer height. Due to the small amount of data the number of neurons is selected as three nodes this is due to the fact proved in earlier cases when the amount of neurons increases drastically the networks encounter the chance of an over-fitting problem. Meanwhile, since there is a small amount of data the risk of having an under-fitting model is very high.

This obliges the design to have a reasonable amount of nodes. The evaluating parameter for this model is mean square error or MSE. The training algorithm for these networks is selected as the Levenberg-Marquardt algorithm. The same distribution of training as it was used for the previous cases is used here, and 70 percent of the data points randomly are allocated to the training of these networks. Likewise, for validation and testing, 15 percent is allocated each.

The thickness ANN model, by using the classical approach has the following performance after 15 iterations. The best validation for this training is obtained as  $5.43 \times 10^{-4}$  with eight iterations.

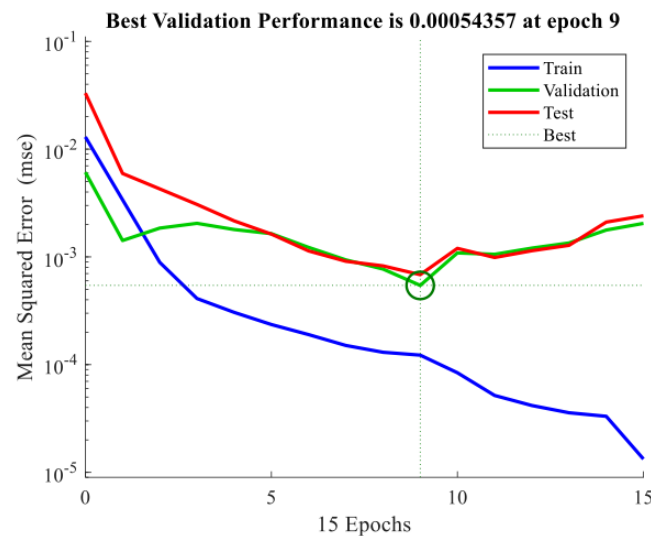


Figure 72: Illustrations of the performance curves for the classical training of thickness.

For the height, the best performance on the data set is obtained as the following diagram with the best  $1.15 \times 10^{-4}$  validation at the 10th iteration.

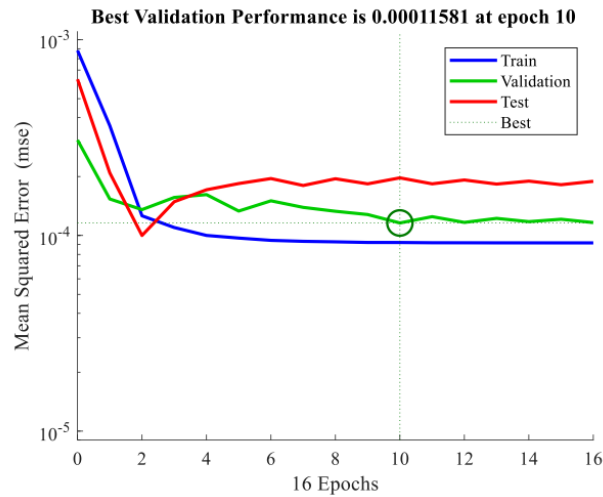


Figure 73: Illustration of the performance result for the classical ANN for the height.

The results for the estimation of the mass respectively is obtained, and the best performances are shown in the following **Figure 74** as the best validation was equal  $2.01 \times 10^{-3}$  on the 23rd iteration.

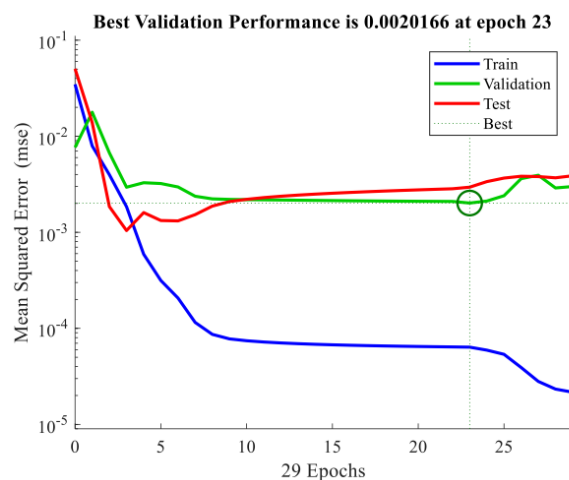


Figure 74: Illustration of the best performance for mass.

With the best performance curves, the best the quality of the training can be evaluated and proper information can be obtained for generalization concerns, overfitting concerns, error quality, etc. The thickness and the mass curves as following a descending trend though it is subject to a low generalization for inputs and the curves ascending with MSE while the training curve is descending indicate the risk of overfitting. The less number of iterations indicates the ANNs are less likely to find a good fit in comparison to previous scenarios that the high number of iterations was accompanied with MSE reduction.



The knowledge-based neural networks as illustrated in the causal diagram is designed with four interconnected modular ANNs. These benefits from the knowledge discovery that is infused in the model. The first ANN is constructed for the viscosity ( $\mu$ ) of molten polymer material at a specific nozzle temperature and a reference viscosity at the temperature of 175 °C. A ratio of such two is implemented with a filament feed rate and extruder temperature (FFR and Tset) as inputs. The corresponding dataset for the ratio as being a blue variable is not possible to obtain but through simulation and computation techniques presented by the DACM framework [106][107].

The KB-ANN is inhibiting the Vaschy-Buckingham  $\pi$  numbers with the independent dimensionless primitives which are referred to as Pi numbers. With the assistant of this concept, there is a possibility of characterization of the physical phenomenon and its dimensionless primitives, and it is possible to reduce the complexity of the system. For the second ANN, this concept has assisted in obtaining the data points for the training of the ANN. The ANN3 in the KBANN network topology is also constructed to model the height (Ht) with the inputs of the layer height (hi) and the number of the layers (n). This is while the ANN4 is used for modeling the mass (Mt) with the input of the thickness (e) and height and material density. In order to perform a better comparison, the same distribution of data for training as 70 percent for training purpose and 15 percent each for validation and testing is chosen. The topology of this type is constructed using three nodes and one hidden layer. The ANN1 performance curves are illustrated as the **Figure 75** where the best performance is obtained at 7.7186e-05 and at the 53rd iteration. The network is trained for almost 60 iterations, and the training curves indicate a better training on the data samples with a downward MSE reduction trend.

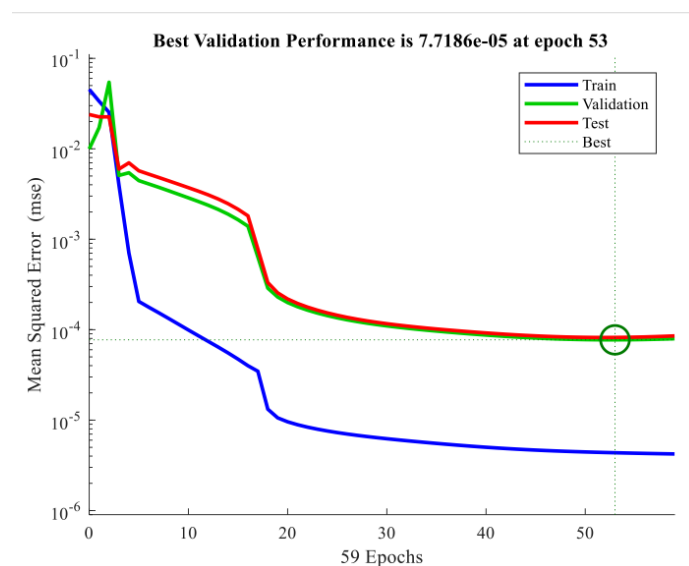


Figure 75: Illustration of the performance curves with the best validation for the ANN1 in KB-ANN network.

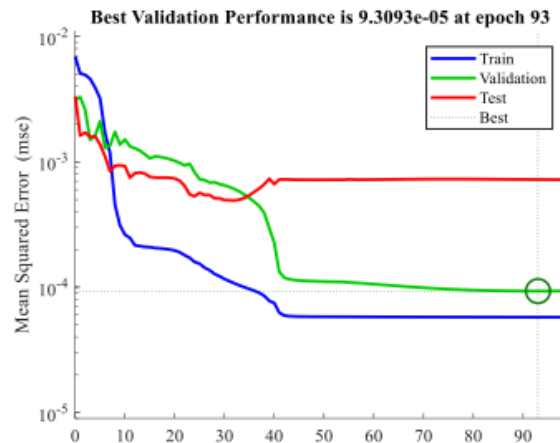


Figure 76: Illustration of the ANN2 performance diagram results in the KB-ANN network.

The ANN2 model for the thickness had the best validation performance of  $9.3 \times 10^{-5}$  at the 93rd iteration **Figure 76**. The amount iteration for this network is showing the training continues with a higher iteration. However, the test curve is not reducing as much as the test and the validation. This is due to the number of data points. Therefore, it is vital to keep in mind the training of the network though it has better results in comparison to the traditional network but still it requires more data points in order to generalize in a broader range of data domain.

The ANN3 best fit is obtained at the 4th iteration with the best validation performance of  $1.41 \times 10^{-4}$  **Figure 77**. This fit on training is smooth, but higher data points are required to reduce the training error lower than the current version.

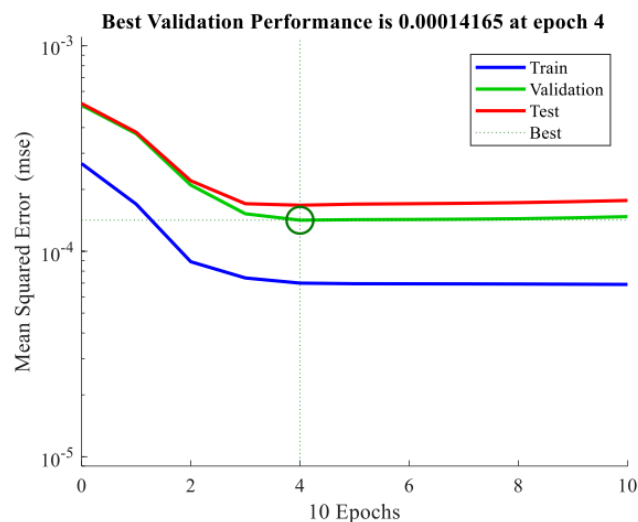


Figure 77: Illustration of the ANN3 performance diagram results in the KB-ANN network.

The result of the ANN4 which is constructing the model for the mass (Mt) is illustrated. The best validation for this network is obtained at 23<sup>rd</sup> iteration with the value of the 2.54e-04. Similar to the curve fitting performances of the previous case study this model is also following a downward trend Figure 78.

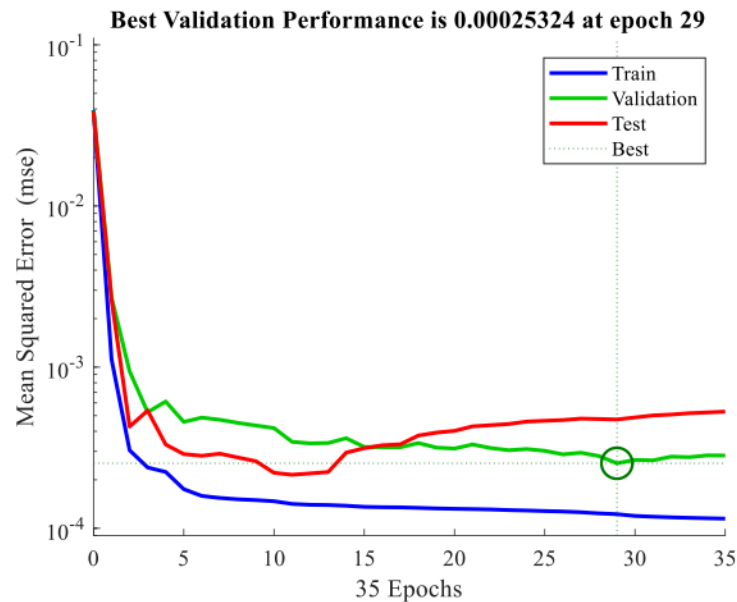


Figure 78: Illustration of the ANN4 performance curves in KB-ANN network.

In order to check the generalization of the models with random data points, an experimental test was conducted with different input values to validate the results and the final quality of the networks. Therefore, the standard error for the ANNs are obtained as follows:

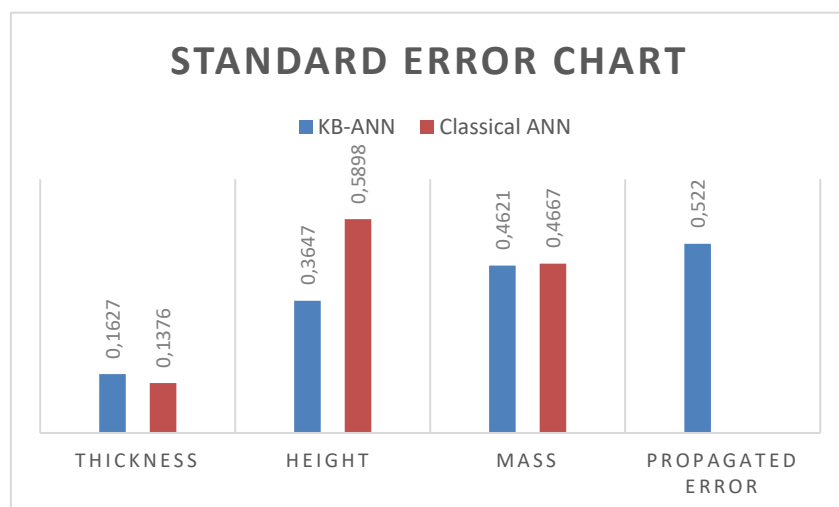


Figure 79: Illustration of generalization test and the amount of standard error for classical and KB-ANN networks.

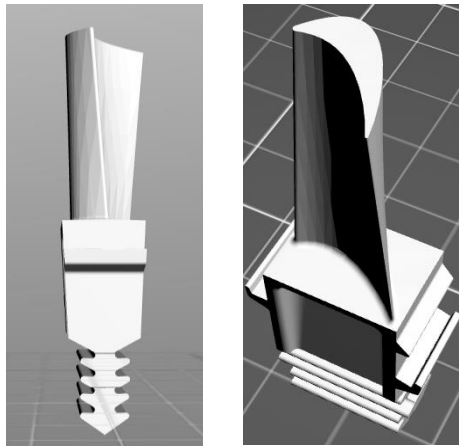
Comparison of two models in **Figure 79** indicates the performance of the prediction for both networks for the thickness is quite close where the classical model has shown better performance. However, this case is different from the height model prediction. The KB-ANN topology has comparably performed better with a lower amount of error. This case is also showing better results for the mass model. Although both networks have performed better in terms of standard error, the KB-ANN is showing slightly better results. One important thing also to notice is that the KB-ANN is going to have a global error meaning that the error of the ANN1 can propagate in the ANN2 and present finally in the output (illustrated in the last bar in **Figure 79**).

With consideration of this fact it is possible to conclude that the KB-ANN network with infusing the expert and existing knowledge is showing better performance. Perhaps in the KB-ANN network and the ANN1 case, the lack of supremacy of performance to classical approach could be the reason of lack of enough knowledge acquisition and infusion of this knowledge into the topology. Nonetheless, in the overall evaluation of the developed approach, the model is equipped with a hybrid tool that both benefits from the empirical data and knowledge in order to predict and obtain the controlling factors in additive manufacturing and especially, in this case, the fused deposition modeling.

#### Case study 4: Bead segmentation using clustering approach

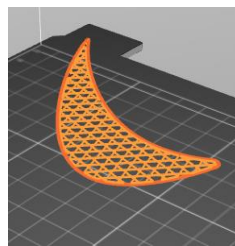
This case studies the application of machine learning in the absence of a labeled training data and uses a non-supervised learning approach for segmentation of the bead borders and in order to apply and separate the involving parameters in each zone and finally escalate the qualitative features of printing in the wire-arc additive manufacturing.

With contemplation of the topic discussed earlier for the WAAM process, if we consider a case of a turbine blade as a CAD model (**Figure 80**), and try to generate this CAD model using wire arc additive manufacturing, it is required to have a perfect nozzle deposition that delivers the material of print precisely. However, in practice, this deposition is not as ideal as it is in for instance in the FDM processes and this case is different for not the same type of material such as steel or aluminum. The quality of the material deposition depends significantly upon several factors especially if this is done automatically.



*Figure 80: Illustration of a turbine blade of a gas compressor.*

Through the process of printing, the CAD model is sliced into layers of surfaces. Moreover, after that it is printed line by line by line and layers of the CAD model is formed consequently. Meanwhile, this line or bead has various quality height, width, and length. This is because of the parameters involved in the quality of the beads. Hence, not only it is required to study the effect of the parameters, but also it is required to know which of the parameters have the most impact on each location or zones and to separate them accordingly.



*Figure 81: Illustration of a cross section or slice surface of the CAD model.*

The quality of the design is a function of the quality of the lines constructing the CAD model in WAAM process. In the illustrated **Figure 82** below the quality of lines of material which is deposited into the plate shows the difference between each bead and their corresponding in-consistency. This inconsistency is due to alteration of different parameters in the welding equipment.



*Figure 82: Illustration of bead with different qualities.*

Then, the design and parameter selection is conducted using factorial design, and it was used to create a tentative plan for modeling part accuracy in the WAAM process. Part quality in WAAM is affected by several input parameters such as current that is delivered to the material like the ignition, working, and ending current.

Also, the voltage delivered to the material and the plate such as ignition, working, and ending voltage has an impact on this. Another critical parameter is the wire feed and the amount of the rate it is delivered to the plate. Moreover, the travel speed, gas flow, and its rate, arc length correction the torch angle are among the other parameters.

With a Taguchi's design of experiments, an entirely orthogonal or semi-orthogonal experiment plan was designed for the input parameters for the WAAM process. Here nine input parameters are chosen for the design of the experiment. The chosen parameters of the welding are listed in the following **Figure 83**:

Symbol	Definition
tign (ms)	Ignition time
lign (A)	Ignition current
ALC (%)	Arc Length Correction
WFR (m/min)	Wire Feed Rate
TS (m/min)	Travel Speed
SGFR (l/min)	Shielding Gas Flow Rate
TA (°)	Torch Angle
tend (ms)	Ending time
lend (A)	Ending current

*Figure 83: Illustration of the parameters used in the welding process.*

These parameters have a high impact on the part precision and totality of the deposited material. Ignition time, ignition current, ALC, WFR, TS, SGFR, end time, and end current are designed for four levels in the DOE plan. Torch angle was designed for three levels. The values for the three levels for the torch angle were chosen based on existing knowledge from literature. It is observed that three positions of the torch (front-hand, backhand, and normal) need to be evaluated to observe the influence of this parameter on the part accuracy. Thus, three angles (110°, 70°, and 90°) were chosen. Four levels for travel speed (0.4, 0.5, 0.6, and 0.7 m/min) were chosen from available literature. It was also observed that the ratio of WFR/TS has a big influence on completeness of a deposited bead. Thus, for each torch angle, the minimum value of WFR for the maximum value of TS was evaluated. The torch angle 70° performed as a restrictive parameter and had the largest ratio of WFR/TS. Based on this evaluation, the minimum value of WFR was chosen to be 3.1 m/min.

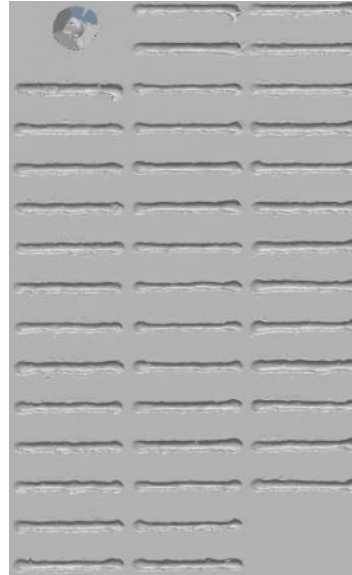
Similarly, it was observed from experiments that the ignition current and end current influence the shape of the bead in the start and end of the welding. From literature, it was seen that a lower value of ignition current ranging between 90 and 120 A (Amperes) need to be selected for the ignition region and to maintain uniformity of shape between starting and working region of the beads. Besides, it was seen from experiments that the ending region of the bead had an inverse influence as that of the ignition region. A higher end current value needs to be selected to avoid the bead from leveling at the end of the welding process. Thus, end current levels were chosen to be 125, 135, 145, and 155 A. Ignition time and ending time represent the time variable at which the ignition and ending parameters are active. The levels were chosen to be 10, 20, 30, and 40ms (millisecond).

The shielding gas flow rate as illustrated influences the quality of the deposited bead. A high flow rate creates a rich shielding zone and reduces the influence of the outside environmental improving quality. Thus, the flow rate was varied from standard/default value 14 l/min to a rich shielding value 20 l/min. Finally, the arc length correction value which marks the length of the arc produced during the welding, as well as the working distance of the torch and the substrate, are set in the range of range from (-30) to (+30).

In general, it is worth to mention that a mixed level orthogonal plan was preferred for the DOE. An L32 orthogonal design was chosen for individual torch angles. Thus, 96 experiments were designed. This procedure was repeated once to evaluate the repeatability of the WAAM process.

The total number of possible experiments according to the states chosen for the variables is  $N_{tot} = 3^1 * 4^8 = 196\ 608$  experiments. It significates that current experiments represent only 0.097% of the total space of experiments.

With conducting the experiments, a laser scanner was applied to scan the printed metal blocks with the purpose of measuring the bead's dimensional structures. The Polyworks 3D metrology software besides a laser scanner was used to obtain the model of the prints. It is worth to note that the maximum accuracy of the scanned part using the laser scanner is 13 microns. Surface scans were obtained for five metal plates containing the beads illustrated as similar to **Figure 84**.



*Figure 84: Illustration of the beads in Polyworks after being scanned using the laser scanner.*

As mentioned earlier the beads were printed on a rectangular metal plate with each plate housing 41 beads. The beads were printed into three rows with each row a distance of 50 mm. The beads are separated by a distance of 15 mm between each bead. The Poly-works 3D metrology software was used with a laser scanner to measure the bead's dimensional structures. The surface scans obtained for each rectangular plate was analyzed using the Polyworks inspector tool. The procedure was as such that the surface scans were imported into poly works inspector. After that, a new Cartesian coordinate system was created at the bottom left corner of the rectangular plate. In Polyworks, the surface scans do not align with the global coordinate system of the modeler. Consequently, a new coordinate system is required through rotation and translation of the global system on the surface scan.

In order to provide analytical data, it is required to create cross-sections separated by 0.1 mm along the X-axis for each bead. The cross-sections are designed to start at point  $(x_i, y_i, z_i)$  before the start of the bead and end at point  $(x_e, y_e, z_e)$  after the end of the bead. This technique is surveyed for the sides of the beads as well in the Y-axis. This is while it provides the option to generate complete cross-sections that run across the full width of the bead analyzed along the length. An example of a bead with cross-sections created using Polyworks is shown in **Figure 82**.



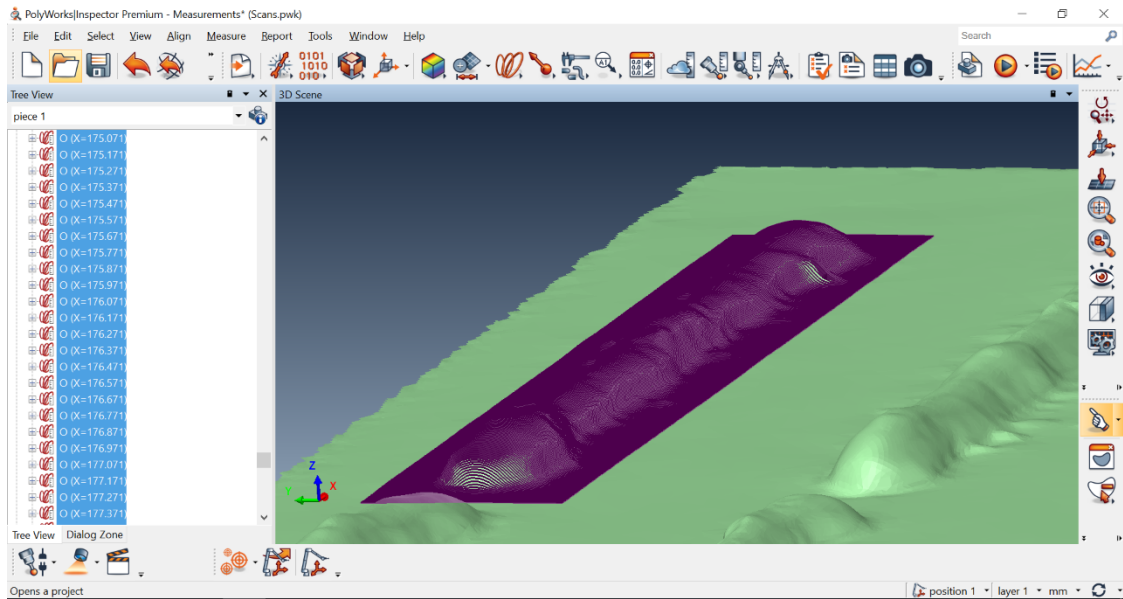


Figure 85: Illustration of a cross section conducted to provide the analytical data.

According to visualization of beads it is apparent that each bead is constructed of three zones this is perhaps as mentioned earlier is due to the involving parameters regarding these zones. In order to segment the zones for each bead and to perform a better analysis of the beads, it is necessary to illustrate the points and by using a non-supervised approach separate the zone's essential parameters.

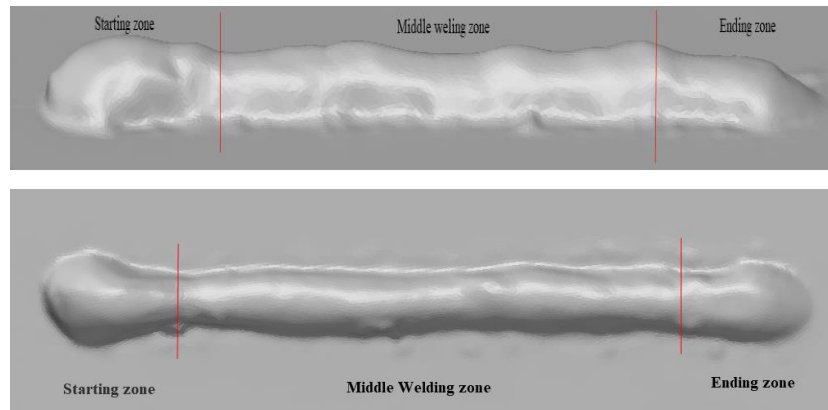


Figure 86: Illustration of a scanned bead with three corresponding zones.

With the use of data points and visualizing the corresponding 192 experiments the width and height of the beads are obtained with the maximum and minimum values, and it is possible to evaluate the quality of the beads with the correspondence of the amount of variation they have during the printing process. As it is observable in the following diagrams (**Figure 87**), it is seen that in both points of view there is a starting zone with a specific height and width (rise and fall) and there is also zones with ending.

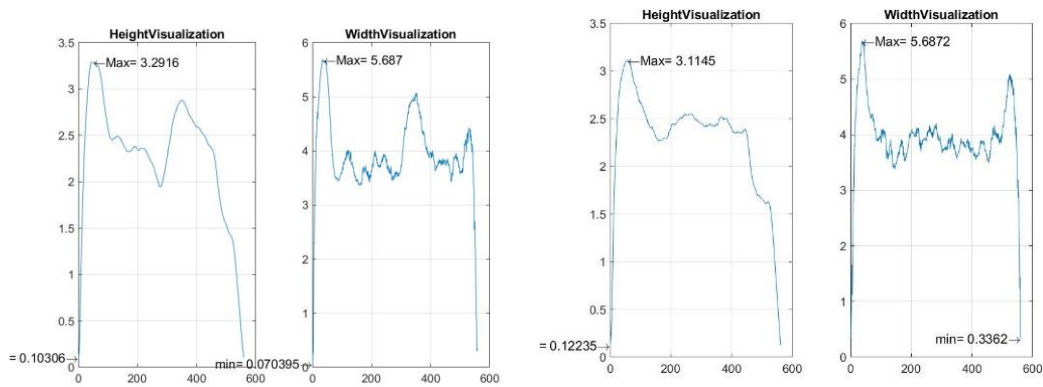


Figure 87: Illustration of bead 1 and 2 height and width data.

Among the clustering approaches the k-means is one of the most reliable and most famous technique. It uses the mean value of the data points and try to create centers and allocate each data point to a cluster based on the distance the data point has to the centers. This will allow the system to provide clusters of sections on the bead. However, one crucial factor is that in this approach the mean of the data is used as a cluster center. This is while this center can become imaginary meaning that it is not corresponding with any data points. This is while the K-medoid approach uses the closest data point that is close to the mean value or meaning it is getting the center of the cluster as the median of the values of the data points. With this now the center of the data is a real data point and it is not an imaginary data.

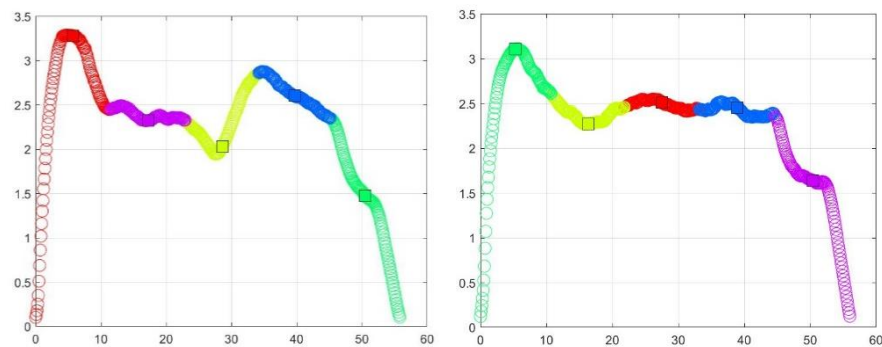


Figure 88: Illustration of the K-Medoid clustering technique for zone segmentation for a WAAM process.

```

%KMedoid
function [IDX,C]=KMedoid(X,k)

n=size(X,1); %size of data
p=size(X,2); %dimension

Xmin = min(X);
Xmax = max(X);
IDX = zeros(n,1);
%Initial the cluster centers
%   CenterIndex=randsample(n,k);
C=X(CenterIndex,:);

while true
    old_IDX =IDX;
    %distance to cluster center
    %   dcc=zeros(n,1);

    %Assginment
    for i=1:n
        xi=X(i,:);
        d=zeros(1,k);
        for j=1:k
            d(j)=norm(C(j,:)-xi);
        end
        %   [dcc(i), IDX(i)]=min(d);
        [~, IDX(i)]=min(d);
    end
    %Update cluster centers
    for j=1:k
        A=find(IDX==j);
        C(j,:)=mean(X(A,:));
        dcc=zeros(numel(A),1);
        for l=1:numel(A)
            i=A(l);
            dcc(l)=norm(C(j,:)-X(i,:));
        end

        % A=find(IDX==j);
        [~,l]=min(dcc);
        C(j,:)=X(A(l),:);
    end
    %Termination condition
    if all(IDX==old_IDX)
        break;
    end
end
end
end

```

Figure 89: Illustration of K-medoid clustering approach for bead segmentation in a WAAM process.

This segmentation is providing the ground for enhanced analysis of the beads and increasing the quality of the prints in a WAAM process. The segmented welding beads and their corresponding data points could be used to distinguish the influential parameters related to each zona. Besides, it assists to discover the optimal welding jobs and in order to decrease the standard deviation of height and the width in those zona. Finally, it helps to smoothen the welding beads and to increase the quality of the beads.

## Case study 5: Radial-Based Neural Networks

Through this case, the application of the RBF networks is investigated on the WAAM process and with the goal of estimation of the Height using the experimental data obtained in the previous sections and by using the DoE. For the implementation of this network, a Gaussian function is used. The centers included in the RBF networks are computed by utilizing K-Median algorithm similar to the bead segmentation case that was presented earlier. The RBF networks use a metric to compute the distance between the inputs and the centers.

For this purpose, the Euclidian equation is used as the metric in this network. The performance of this network is evaluated by measuring the MSE values. Multiple iterations are conducted as well to catch the best and minimal value for the MSEs. From the experimental results of DoEs, the nine parameters are used as inputs, and the Height is estimated using this RBF network as the output. The Error! Reference source not found. **Figure 83** in previous section illustrates the list of these parameters for the WAAM process. One of the most significant parameters which are required to set for the RBF networks is the location of the clusters or equal to the number of nodes for MLP network.

Accordingly, as it is discussed in [108] one technique is to allocate one Gaussian function kernel for each training samples. However, though this approach can provide a good fit with using the neighborhood concept, but in practice and presence of noise which is very common, the network is not able to generalize its decision making for the data samples that are outside of the training set. Consequently, it is common to select the centers randomly or to cluster the inputs and select those cluster centers as the kernels of the RBF network. In such cases also one common algorithm option is again the K-means or K-Medoid algorithm for clustering the inputs.

The **Figure 90** illustrates a typical topology of such networks with the radial kernels in the hidden layer section. This is similar to the MLP networks but instead of nodes the cluster centers or kernel centers have the control in approximation of the outputs. In this topology the  $x$  is used as the input variables and the  $F^*(s)$  would be considered as the estimated approximation of the  $F(x)$  output. Consequently, the difference of the  $F(x)$  and  $F^*(x)$  could be used for calculation of the standard error.

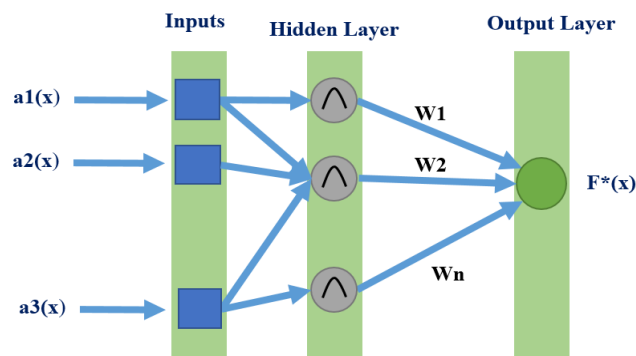


Figure 90: Illustration of the typical RBF network topology with radial function in the hidden layers.

This trainings in this case study are conducted with different learning rates and different number of the clusters. As how it is mentioned earlier in [108] the quality of training in terms of standard errors and generalization is very dependable on the selection and location of the centers while the learning rate has an excessive impact on the training quality. Therefore, in order to provide a better generalization, various type of center clusters is tested on the WAAM experimental data. In addition, the network performance is validated using different learning rates. If the learning rates decrease significantly it will impose a higher cost on the systems, sometimes it leads to better results with a higher number of iterations but it also leads to a lower precision too and the network can face the memorization of the samples and will lead to generalization issue eventually.

In the first training, the RBF network is trained with randomized selected training and testing data sets. The corresponding learning rate is selected as 0.0003 and the number of the Gaussian clusters is allocated as five. In the **Figure 91**, on top left corner the training and test curves are illustrated. This indicates the amount of MSE for the training and as it is seen the MSE is decreasing normally and it reaches the values less than 0.2 with the best training MSE=0.080704 after 77 iterations. The right top diagram indicates the error of the network between the target and the predicted values for the Height output variable. In this plot the horizontal axis stands for the index of the data samples that are used for test and vertical stem diagram shows the difference of the target and predicted values.

Moreover, the lower bottom diagram shows the distribution of the 60 test data samples and their predicted and actual values. This figure illustrates the quality of the fit to the test data. One important thing to mention is that the amount of data samples is very limited however, the RBF network is able to provide a decent regression fit in 77 iterations and with best test MSE of 0.055453.



Figure 91: Illustration of the result of RBF training with learning rate of 0.0003 and 5 cluster centers.

In the next training the learning rate is decreased to the 0.00003 and the amount of centers is limited to 4. With the decrease of the learning rate the number of iterations increases to 531 iterations. The training MSE get lower to the final value of the 0.086682. This is while in the right top corner of the diagram the error of t the target minus the actual values is shown in stem format. The test MSE for this network is equal to 0.084103.

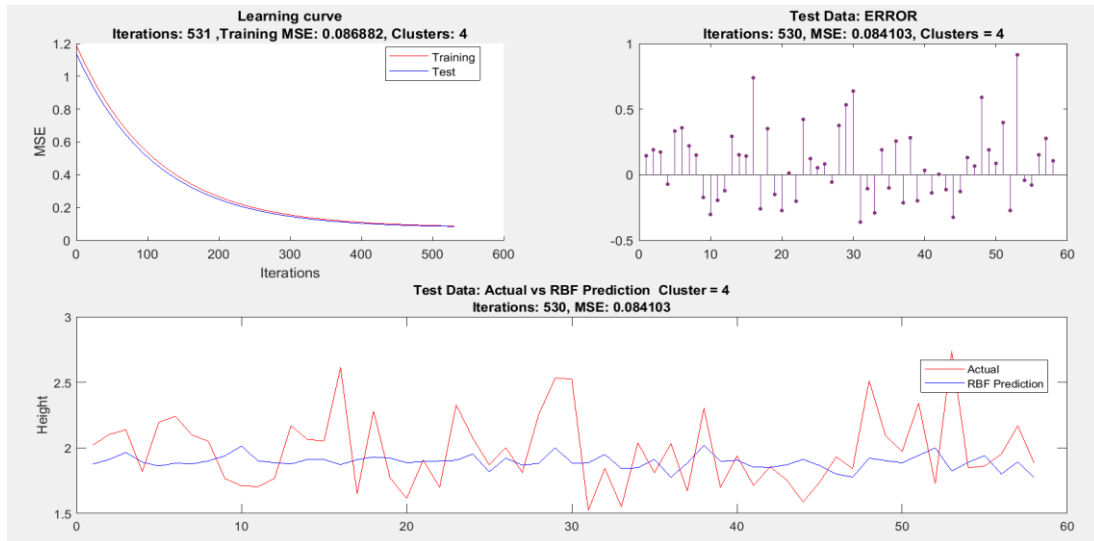


Figure 92: Illustration of the RBF network results with learning rate of the 0.00003 and 4 clusters.

If we change the learning rate to 0.00008 and the number of centers to 4 the corresponding result is obtained as in Figure 93 with an MSE=0.061946 and 4 clusters.

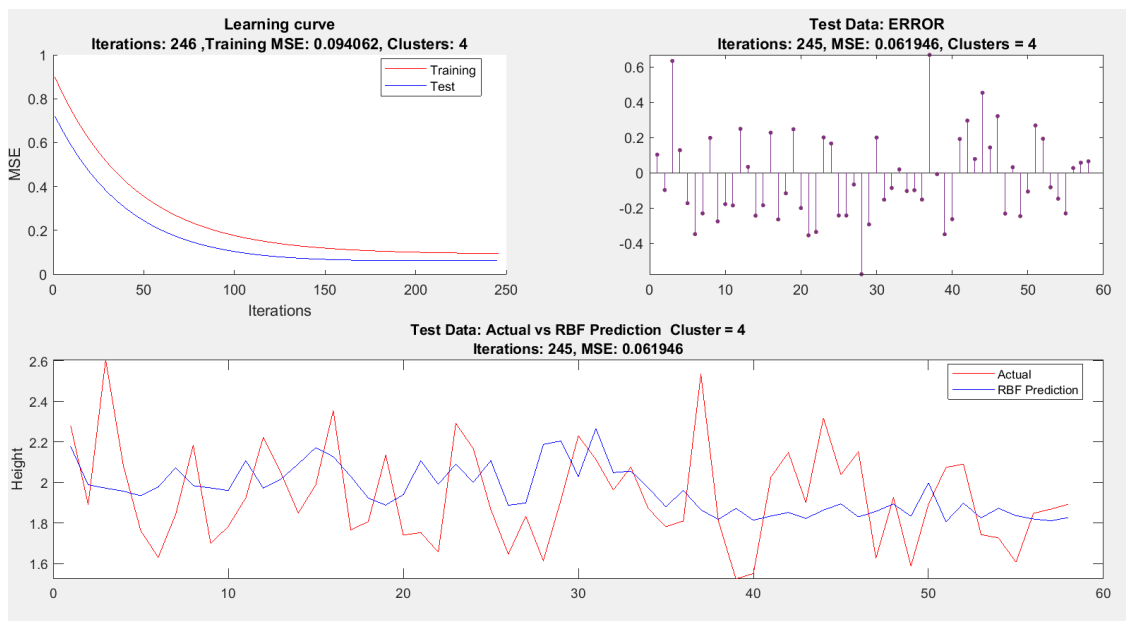


Figure 93: Illustration of results of the RBF network for the learning rate of 0.00008 and 4 number of centers.

Different networks with initialization must be tested and validated in terms of quality (MSE, generalizations, under fitting or overfitting) so that the best descriptive model to be generated. The reason for this is due to the data sets which is used for training and test randomly. Meanwhile, the increase in learning rate will lead to a lower amount of learning interactions and put the network at risk of under fitting. Whereas, a high learning rate would cause a higher number of iterations and puts the network into the risk of overfitting and sometimes would lead to a corrupted estimating model.

In order to illustrate the validation of the RBF results and its estimator it is worth to provide a model for a known relation. This is used as comparison model case and tends to studies the known behavior function  $y=\sin(x)$  plus an additional random noise and to show how the amount and quality of the data can assist in the training procedure. As it is shown and test with a higher number of data points, the quality of the approximation gets better. The test MSE for this network is equal to 0.011877 with six centers, and the network is trained with only 44 iterations. In comparison to the WAAM data set it is shown that with a range of 5 or 6 cluster centers it is possible to provide a respectable approximating model however in terms of the number of data point since the  $\sin(x)$  function benefits more from higher amount of data a higher learning rate is used a good fit is provided. While for the WAAM process the learning rate is adjusted accordingly to compensate this lack and to obtain an excellent fitting model.

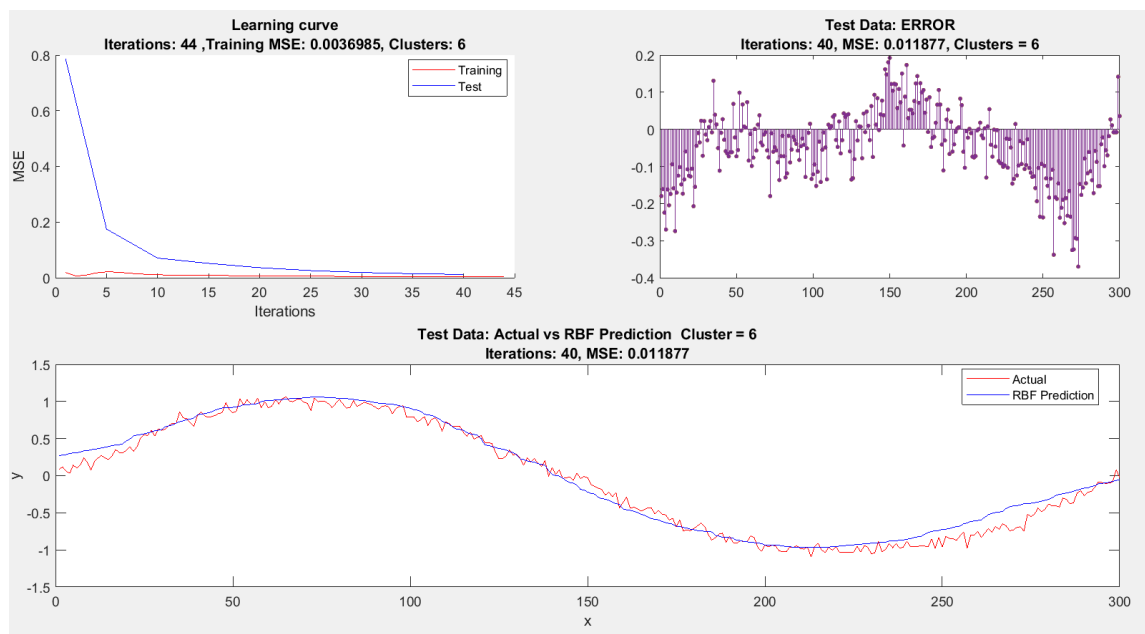


Figure 94: Illustration of an approximate using the RBF networks for a known  $\sin(x)$  function plus a random white noise. The best fit is obtained for a six cluster centers and with a learning rate of 0.03

## Case study 6: NeuroFuzzy

In this case, rules and knowledge are extracted from the real data, and an inference engine evaluates the accuracy of the welding through the association and amount of membership the value of the variables have in a specific set. The objective of this case is to seek the optimal parameter discovery in wire arc additive manufacturing. As mentioned earlier, this type of approximation is running based on the parameter optimization in WAAM process and has the goal of the error reduction and to obtain a weld with a uniformed shape. The network engine which is used for this case is using the fuzzy inference processing algorithm and similarly simulates the controlling procedure of the welding parameters. The inference engine is using the experimental data that is acquired through a DoE which has been mentioned briefly in the previous case study.

One of the advantages of the fuzzy inference system is that it is implementable as a suitable controller in additive manufacturing processes and is used usually for the systems which have high non-linearity [109][110]. One of the challenges that additive manufacturing optimization is facing is the requirement for a high number of data points. However, a controlling system similar to fuzzy systems, by using a knowledge base reduces the necessity for the amount of experimental training data. As illustrated in previous sections and in the **Figure 83: Illustration of the parameters used in the welding process.** the parameters that have the most influence on the welding process are listed. These parameters are selected either by using an expert knowledge or with primary statistical analysis. For instance, most of the parameters which are involved in the WAAM process and for the 3D metal welding is listed in the **Table 4** and a possible causal relation is illustrated similar to the **Figure 95.**

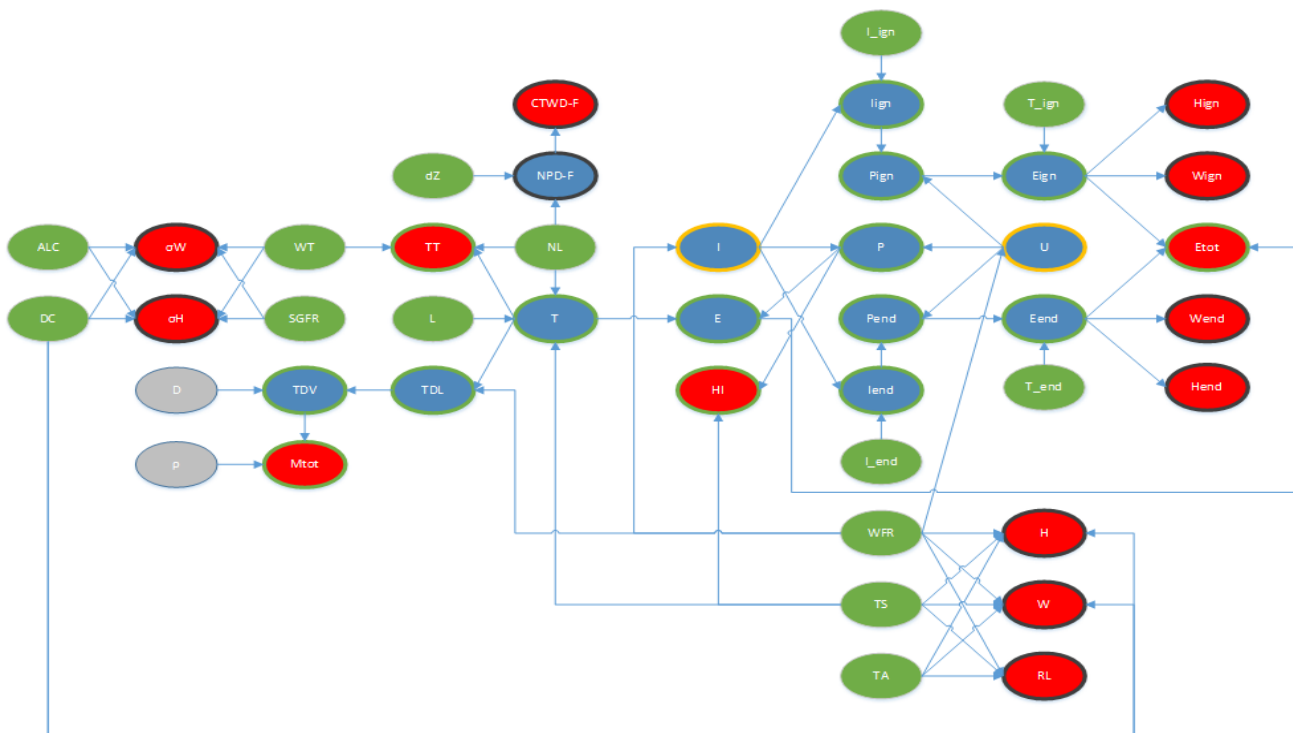


Figure 95: Illustration of a potential causal relation in the WAAM process.



Table 4: List of parameters involved in WAAM process.

Variables	Definition
ALC	Arc Length Correction (%)
DC	Dynamic Correction (#)
$\sigma_W$	Standard Deviation of width (mm)
$\sigma_H$	Standard Deviation of height (mm)
D	Wire Diameter (mm)
$\rho$	Volume Mass of the Wire (g/mm <sup>3</sup> )
WT	Waiting Time (s)
SGFR	Shielding Gas Flow rate (l/min)
TDV	Total deposited Volume (mm <sup>3</sup> )
Mtot	Total Deposited Mass (g)
dZ	Z-variation in RAPID code (mm)
TT	Total Time (s)
L	Length of the wall in RAPID code (mm)
TDL	Total Deposited Length of Wire (mm)
CTWD-F	Contact-Tip Working Distance Final (mm)
NPD-F	Nozzle-Plate Distance Final (mm)
NL	Number of Layers (#)
T	Welding Time only (s)
I	Average Current (A)
E	Average Energy (J)
HI	Average Heat Input (J/mm)
I_ign	Ignition current set in RCU5000i (%)
I <sub>ign</sub>	Ignition current computed (A)
P <sub>ign</sub>	Ignition power (W)
P	Welding power (W)
P <sub>end</sub>	Ending power (W)
I <sub>end</sub>	Ending current computed (A)
I <sub>end</sub>	Ending current set in RCU5000i (%)
WFR	Wire Feed Rate (m/min)
TS	Travel Speed (mm/s)
TA	Torch Angle (°)
T <sub>ign</sub>	Ignition time set in RCU5000i (s)
E <sub>ign</sub>	Ignition energy (J)
U	Average Voltage (V)
E <sub>end</sub>	Ending energy (J)
T <sub>end</sub>	Ending time set in RCU5000i (%)
H	Real height of the wall for welding zone (mm)
W	Real width of the wall for welding zone (mm)
RL	Real length of the wall for welding zone (mm)
H <sub>ign</sub>	Real height of the wall for ignition zone (mm)
W <sub>ign</sub>	Real width of the wall for ignition zone (mm)
E <sub>tot</sub>	Total energy of the process (J)
W <sub>end</sub>	Real width of the wall for ending zone (mm)
H <sub>end</sub>	Real height of the wall for ending zone (mm)

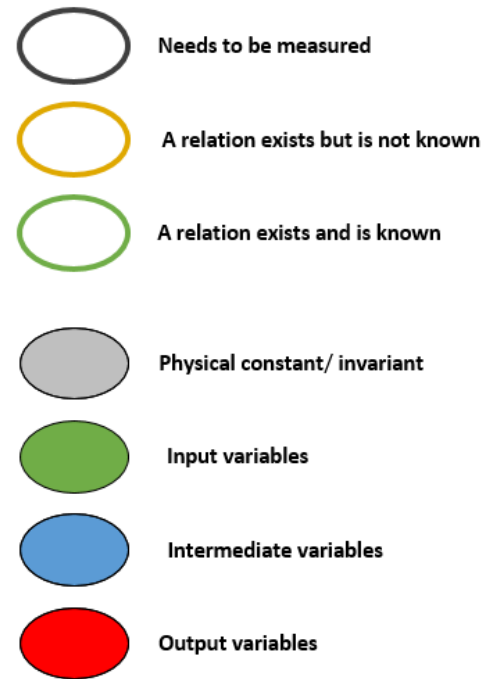


Figure 96: Definition of the causal graph.

The causal relation which is obtained in the diagram is mostly based on the prior expert knowledge and it is based on the observation conducted for 96 experiments in a WAAM process. With the selection of the nine variables, it is possible to validate this information and the prior knowledge by using the statistical models. Similarly, the integration plot of the Height as the output variable is checked along the other inputs. In the **Figure 100** the interaction between the variables is illustrated. If the interacting diagrams for the corresponding variables follow a parallel curve these two variables are not having any interaction or relation on the output. Meanwhile, if the curves intersect each other; this means that the two variables are having an interactive relation on the output.

One other important thing to notice is that these values must be analyzed with scrutinizing procedures in order to produce exact rules and models. This has the meaning that for a specific range of values the interaction between the parameters vanishes, but for some ranges of values, there is an interaction. Moreover, it also assists in the removal of any non-important interaction among the parameters and makes the designer focus only on the variables and values that are significant with regards of the interaction's effects. In the following figure the interaction of the Travel speed and the torch angle of welding device as one of the instances shows there no effect on the Height of the printed material. But in **Figure 100** for some of the variables there is a complete interaction for all the intervals and for some there only on specific intervals. Note that the values in **Figure 97** are using a normalized values of the parameters and experimental data.

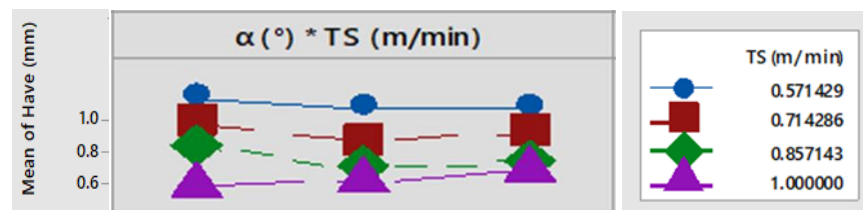
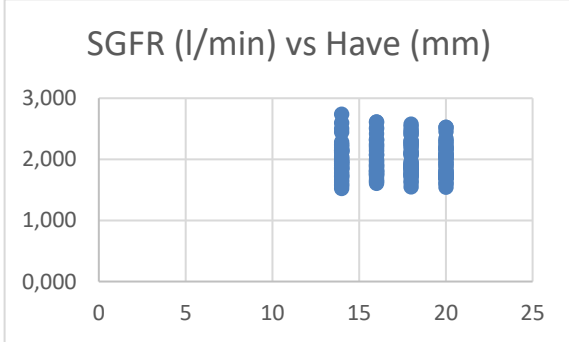
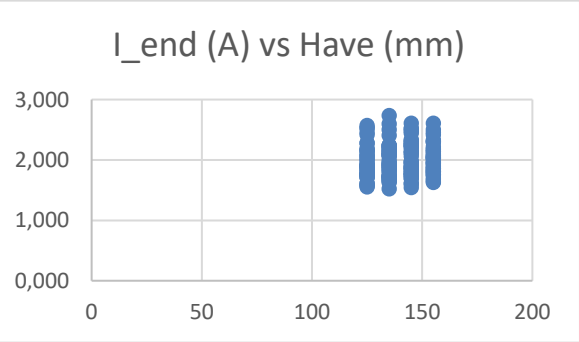
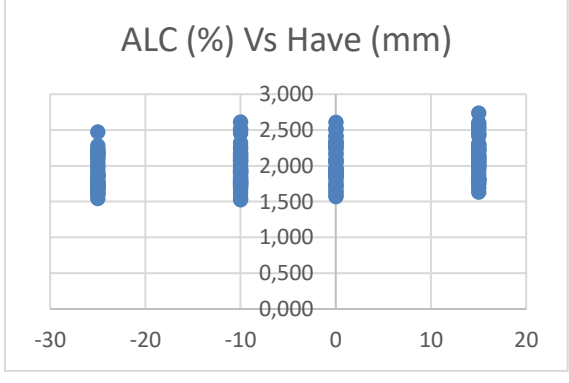
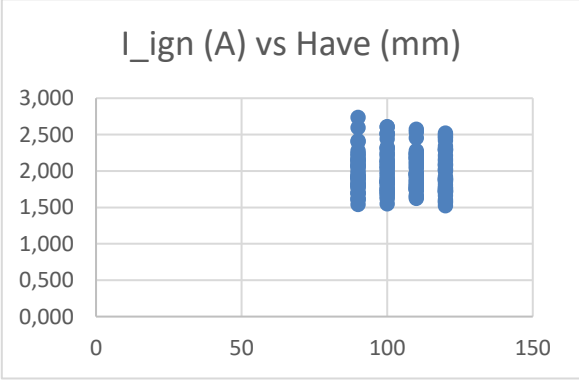
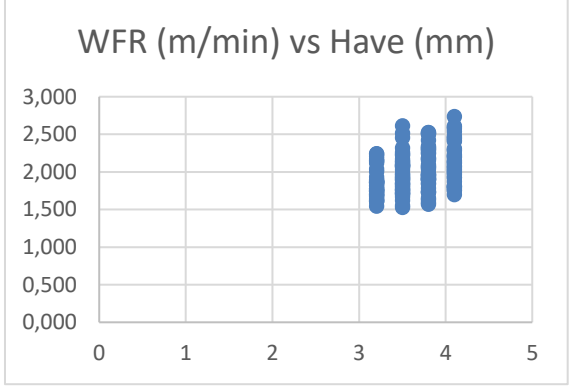
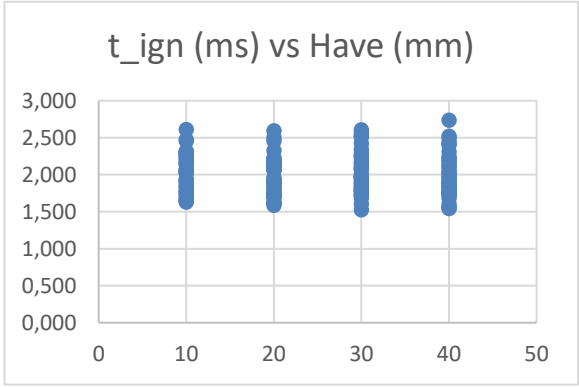
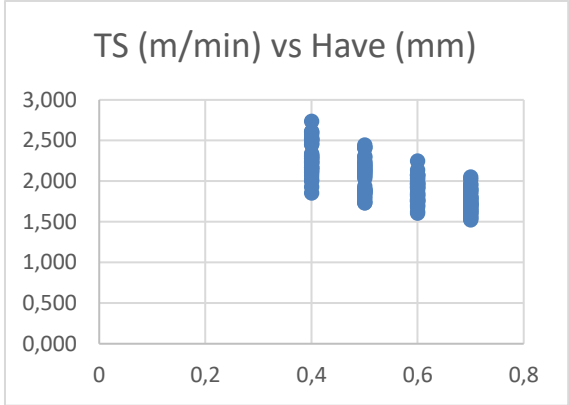
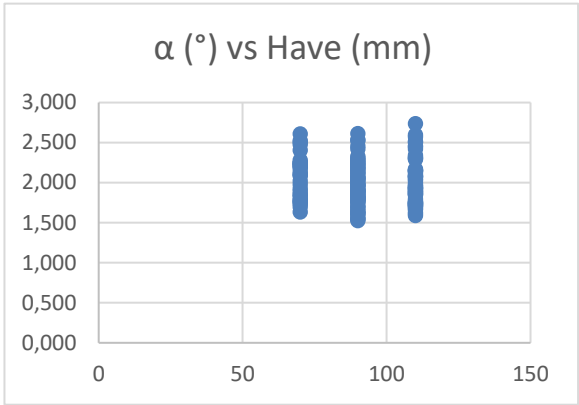


Figure 97: Illustration of the interaction among the parameters on the mean of average Height.

Another information or knowledge which can be extracted is by using the plot of the data versus the output variables and comparing the results with using the main effect plot of the variables. In this stage, one vital thing to notice when inspecting the one to one plot for the variables is that the interaction of the variables must be taken into consideration since the one to one diagrams is not showing the real behavior and effect of any change in those parameters on the output. However, this can be used to validate the interaction of the parameters when one exists in.



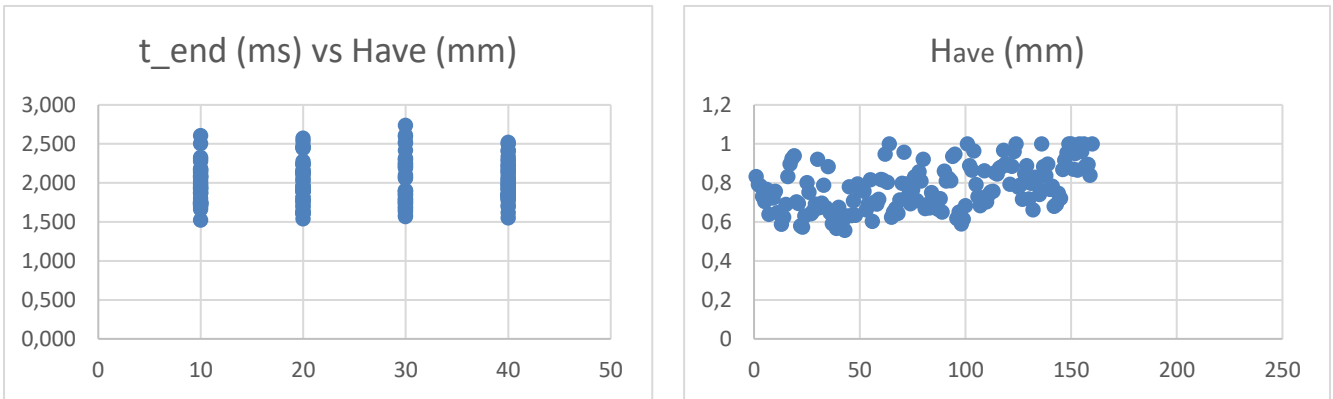


Figure 98: Illustration of the variables and their corresponding Height values.

With cross validating these diagrams **Figure 98** and the main effects plot **Figure 99** for the output variable; it is possible to assist design of rules immensely. It also helps the inference engines and decision making procedure to act flexible and robust.

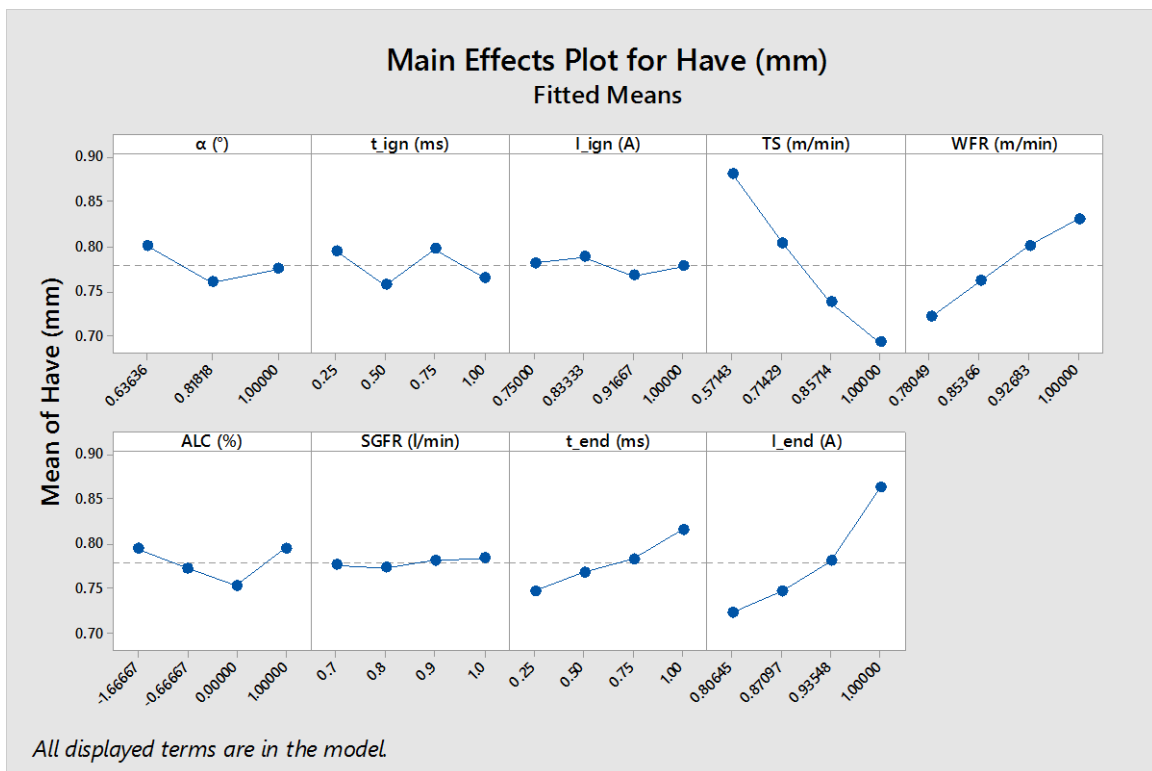


Figure 99: Illustration of the main effect plot for the output Height and the other main input variables.

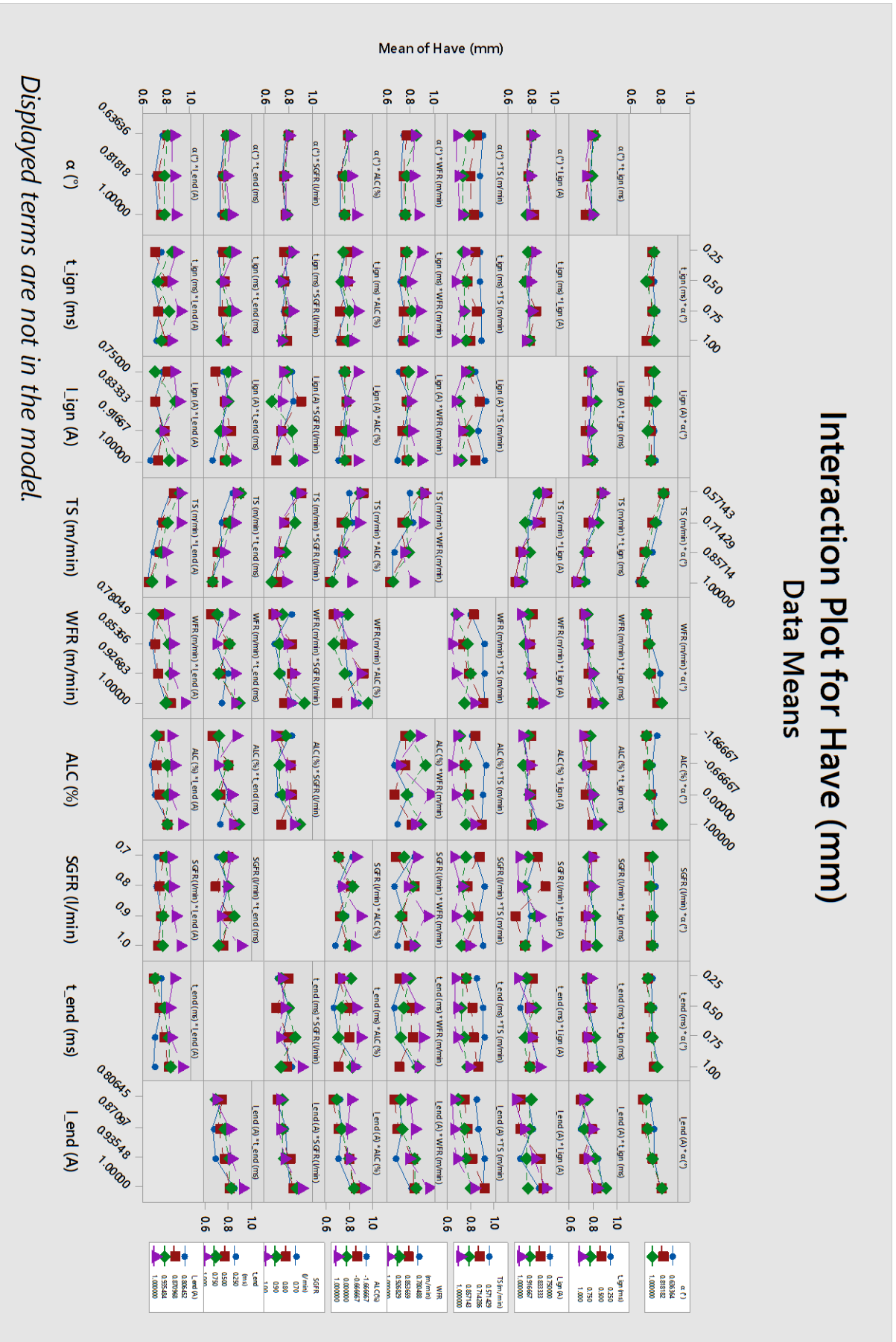


Figure 100: Interaction of the input variables in the WAAM process and the Height variable as the output.

Meanwhile, with normalizing the data points and conducting a two way ANOVA test which is accompanied by a hypothesis P-test; it is possible to discover the most influential parameters on the Height again on the output. The **Figure 101** lists all the variables factors with their different levels which is normalized.

Factor Information				
Factor	Type	Levels	Values	
$\alpha$ (°)	Fixed	3	0.63636, 0.81818, 1.00000	
t_ign (ms)	Fixed	4	0.25, 0.50, 0.75, 1.00	
I_ign (A)	Fixed	4	0.75000, 0.83333, 0.91667, 1.00000	
TS (m/min)	Fixed	4	0.57143, 0.71429, 0.85714, 1.00000	
WFR (m/min)	Fixed	4	0.78049, 0.85366, 0.92683, 1.00000	
ALC (%)	Fixed	4	-1.66667, -0.66667, 0.00000, 1.00000	
SGFR (l/min)	Fixed	4	0.7, 0.8, 0.9, 1.0	
t_end (ms)	Fixed	4	0.25, 0.50, 0.75, 1.00	
I_end (A)	Fixed	4	0.80645, 0.87097, 0.93548, 1.00000	

Figure 101: Illustration of the list of factors information after normalization for the Height output value.

The analysis of the variance indicates with the P-values that the difference of the WFR on the Height for example has a P-values of zero. In the P-Test if the value of the P become less than a specific range for instance the 0.05 or become less than 5 percent it indicates the null hypothesis of the effect of the variation is true and the variation the parameter has a significance effect on the output variation. Meanwhile if this value is high and usually greater than 0.05 this means that the variation is not very significant. According to the **Figure 102** the P-value for the WFR, TS, Torch Angle, Time of ignition, Time of weld ending and its final current I-end has a significant variation on the height. With comparing these values with the diagrams of the main effect plot in **Figure 99** these conclusion could be validated.

Analysis of Variance					
Source	DF	Adj SS	Adj MS	F-Value	P-Value
WFR (m/min)	3	0.10383	0.034612	11.27	0.000
TS (m/min)	3	0.81574	0.271914	88.56	0.000
$\alpha$ (°)	2	0.03459	0.017293	5.63	0.004
t_ign (ms)	3	0.04986	0.016619	5.41	0.002
I_ign (A)	3	0.00871	0.002903	0.95	0.421
ALC (%)	3	0.01820	0.006065	1.98	0.121
SGFR (l/min)	3	0.00268	0.000892	0.29	0.832
t_end (ms)	3	0.09948	0.033161	10.80	0.000
I_end (A)	3	0.45711	0.152371	49.62	0.000
Error	133	0.40837	0.003070		
Total	159	2.15085			

Figure 102: Illustration of the results of a TWO WAY ANOVA on the WAAM for estimation of the Height.

As listed in the **Figure 102** and with respect to the main effect plot **Figure 99: Illustration of the main effect plot for the output Height and the other main input variables**. the following variable variables has the significant impact on the height:

1. WFR(m/min)
2.  $\alpha$  (°)
3. TS (m/min)

But meanwhile for the ALC (%) the P-Value is 0.121. Though the Pearson correlation of the ALC (%) is quite significant **Figure 103** but in the **Figure 100** the matrix indicates the ALC (%) has interaction mostly with all the variables. It is possible to assume the main effect of ALC (%) and its role is neutralized. Here we assume that this hypothesis is correct for now and build a model only with other listed variables and effective variables. Also, based on the short come of the time this study is going to conducted analysis only middle zone of the beads to eliminate the corresponding effective variables related to the other zones **Figure 104** these parameters include the  $t_{ign}(ms)$ ,  $t_{end}(ms)$  and the  $I_{end}(A)$ .

	CoRel	Absolute	Relation
SGFR (l/min)	-0.011535553	0.011535553	-1
$I_{ign}$ (A)	-0.013897553	0.013897553	-1
$t_{ign}$ (ms)	-0.014999086	0.014999086	-1
$t_{end}$ (ms)	0.056932119	0.056932119	1
$\alpha$ (°)	-0.075017773	0.075017773	-1
$I_{end}$ (A)	0.081486845	0.081486845	1
ALC (%)	0.29797731	0.29797731	1
WFR (m/min)	0.368978811	0.368978811	1
TS (m/min)	-0.761025134	0.761025134	-1

Figure 103: Illustration of the correlation between input variables and the output Height.

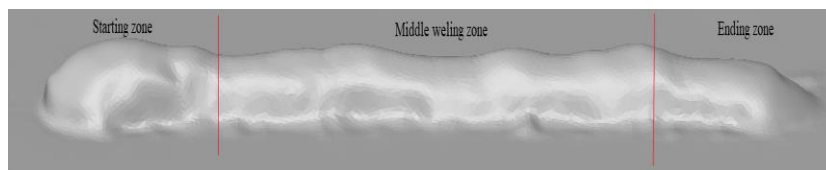


Figure 104: Bead segmentation to separate the bead zones.

In order to construct a fuzzy inference engine and later that to use this engine as a back bone of the neuro-fuzzy systems we would consider a model as follow:

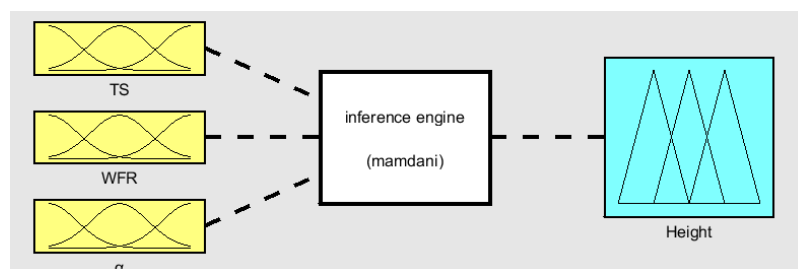


Figure 105: Illustration of the fuzzy inference engine with 3 inputs and 1 output variable.

The precision of the model is very dependable on the quality of the data and especially the rules which are define to estimate the output variable. For this purpose the variables are all normalized to attenuate the significate of very big values.

Based on the main effects plot for the  $H_{ave}$  **Figure 99**. It is possible to construct some rules in order to construct this type of model. The rules construction is based on membership's functions therefore, the definition and allocation of any type of the membership functions is defining the quality of the prints. For the TS and WFR intervals we can allocate four intervals as;

1. High
2. Medium High
3. Medium Low
4. Low

And for the torch angle it is possible to allocate the

1. High
2. Medium
3. Low.

Although there are nine different membership functions to test. In this case for the parameters trapezoidal-shaped and Gaussian membership functions are selected. The allocation of the values and membership function at the moment is manual. However, there are some approaches to allocate the rules automatically and based on the ranges.

The allocation of the membership function and the high, low, medium high, medium low is selected as follow. For the TS values before 0.57143 are considered as low travel speed, and the values until 0.71429 would be medium-low, the 0.85714 is medium high, and finally for the other values till TS=1 the membership would be considered as high **Figure 106**.

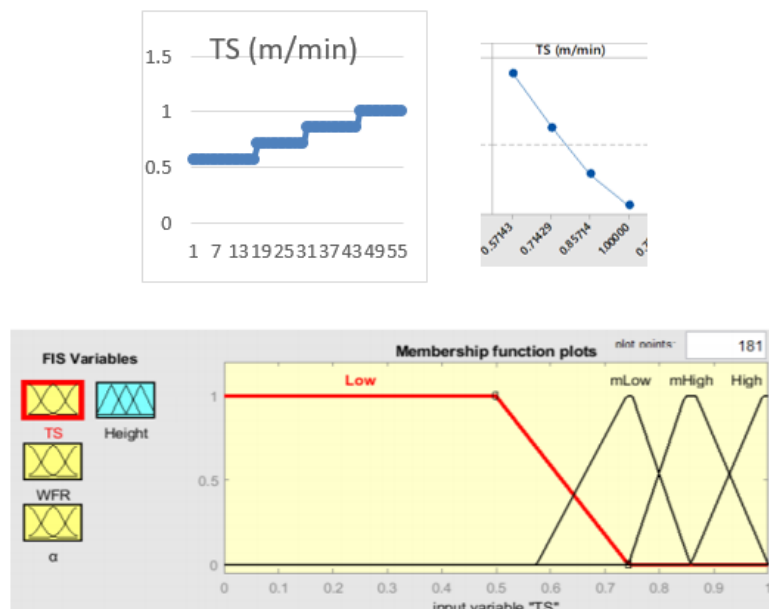


Figure 106: TS vs Height values (top) and Membership function for the TS (down).



For the WFR (m/min) parameter the membership functions are allocated as follow but with a positive correlation **Figure 107 (top right)**. With a similar procedure, the membership functions for this parameter is defined as illustrated in **Figure 107 (down)** which is also based on the value ranges of the WFR (m/min). This parameter is discretized into four categories of values like the TS parameter. One important thing to note is that these borders are primitive and they can be adjusted accordingly to provide the best prediction model.

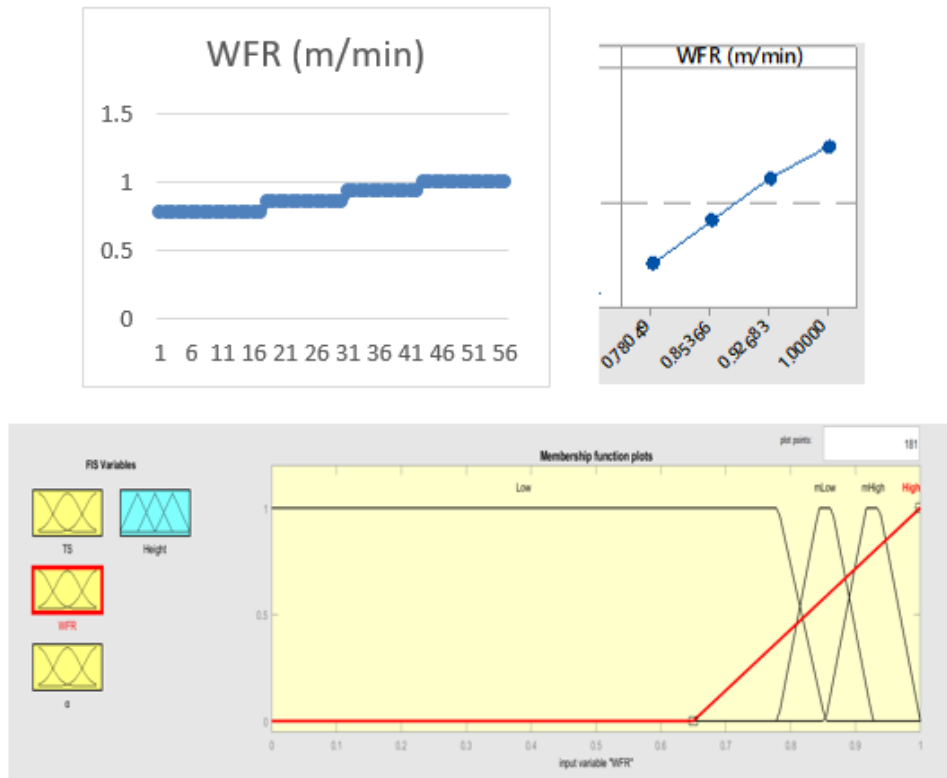


Figure 107: WFR (m/min) vs Height and corresponding WFR's range (top) and Illustration of the membership functions for the WFR.

For the torch angle  $\alpha$  ( $^{\circ}$ ) three membership functions are allocated as Medium, Low and High. In this case, we also use a trapezoidal membership function **Figure 108**. However, the study of the choosing and allocating the best membership function for the reason of optimization would be considered as the future works. But for now and for the torch angle  $\alpha$  ( $^{\circ}$ ) the membership functions are allocated as **Figure 109**.

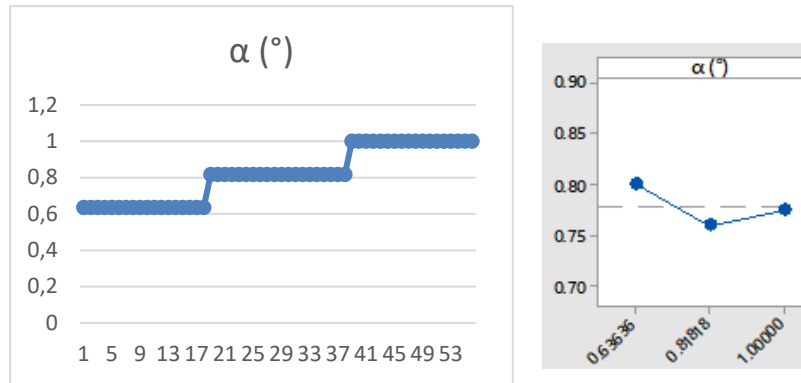


Figure 108:  $\alpha$  ( $^{\circ}$ ) vs height effect plot and the range of values.

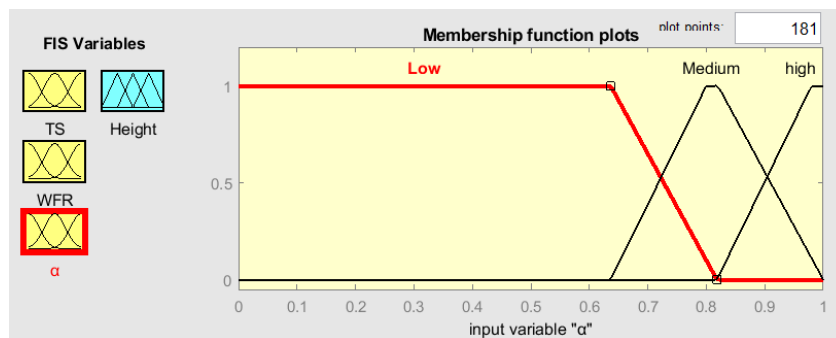


Figure 109: Illustration of the membership functions for the torch angle.

Finally, for the output it is required based on the Height and its ranges of values to allocate four levels of membership. Two triangular membership function is used for medium-low, and medium-high ranges, one trapezoidal for the low range and one Gaussian membership function for the high range of values **Figure 111**. The range of the average height (Have) normalized values changes from 0.5 to 1 which this is considered while allocating the membership functions in **Figure 110**.

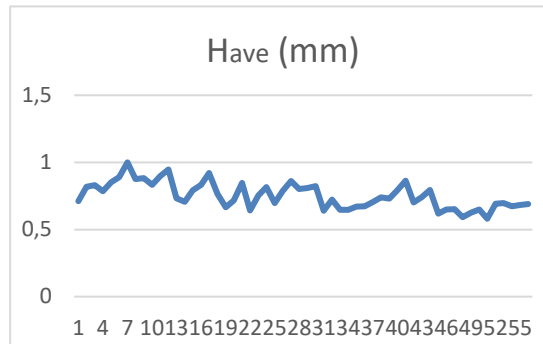


Figure 110: Illustration of the range of Height's values for 55 samples.

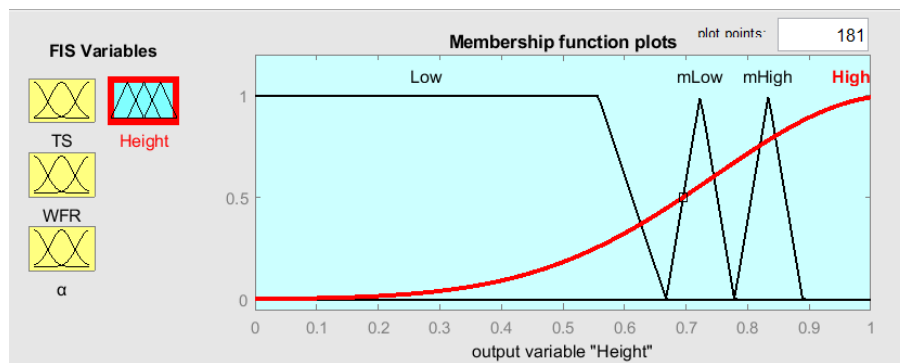


Figure 111: Illustration of the membership functions for the Height parameter.

As the membership functions for the input and output are obtained and allocated in order to conduct the inference, it is required to define the rules corresponding to the membership functions. The rules will be defined as follows but it is worth to mention here the quality of the model is very dependable also on the way of defining the rules. Therefore, in this case, due to the shortage of time, the modeling is studied with few rules but perhaps for the future works it is possible to investigate and increase the accuracy of the model correspondingly by adjusting and allocating the rules optimally.

Based on the interactions and the main effect plots in **Figure 99** the main three rules are constructed for estimation of the values which are falling in the high category of the Height:

1. **IF** the travel speed (TS) is medium high (mHigh) **OR** the wire feed rate (WFR) is medium high (mHigh) **OR** the torch angle  $\alpha$  ( $^{\circ}$ ) is low **THEN** the Height is medium high (mHigh).
2. **IF** the travel speed (TS) is medium high (mHigh) **OR** the wire feed rate (WFR) is medium low (mLow) **OR** the torch angle is medium **THEN** the Height is medium low (mLow)
3. **IF** the travel speed (TS) is low **OR** wire feed rate (WFR) is high **OR** the torch angle is high **THEN** the Height is high.

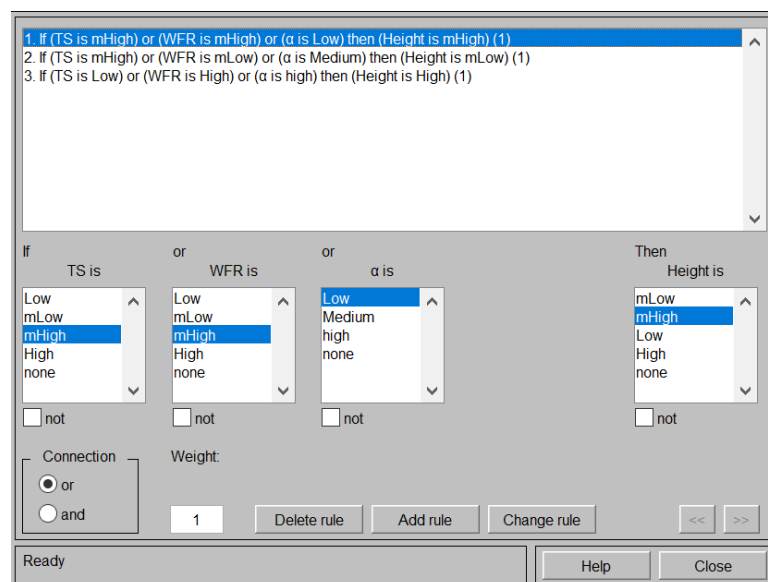


Figure 112: illustration of the rules definition in fuzzy toolbox in Matlab.

After definition of the rules in order to validate the quality of the inference engine some data points as the test has been given to the system and their estimated values are calculated **Figure 113**.

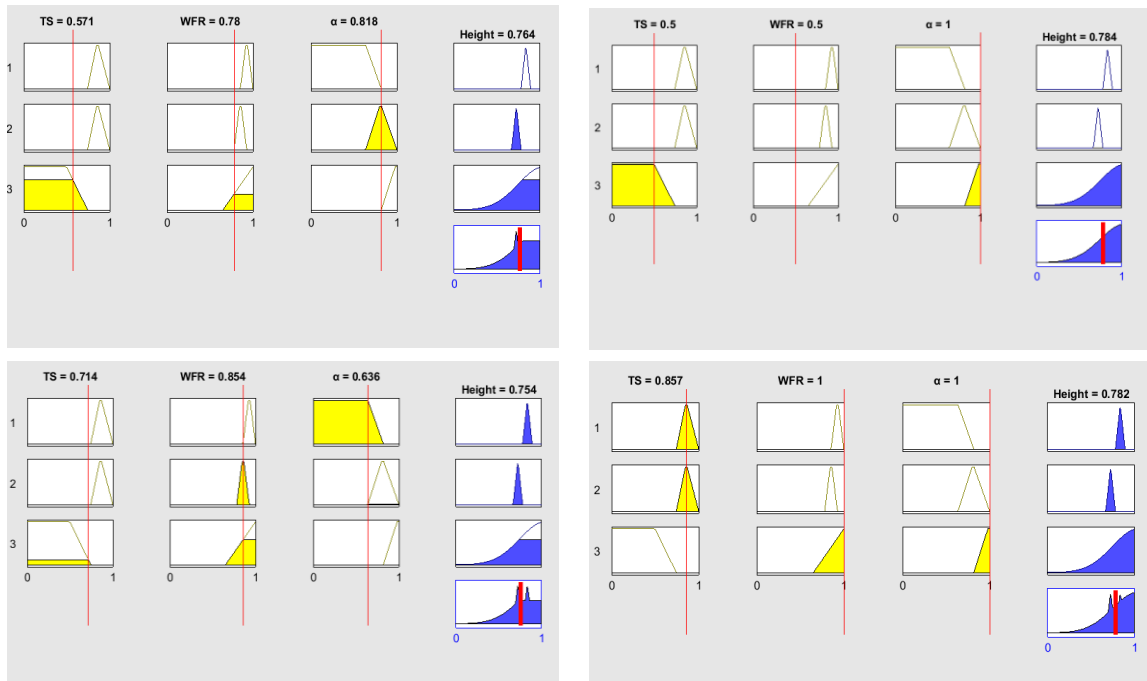


Figure 113: Illustration of the fuzzy inference with the corresponding rules define for estimation of the Height.

The **Figure 114** is presenting the actual test data for WAAM and the predicted values by the fuzzy inference engine.

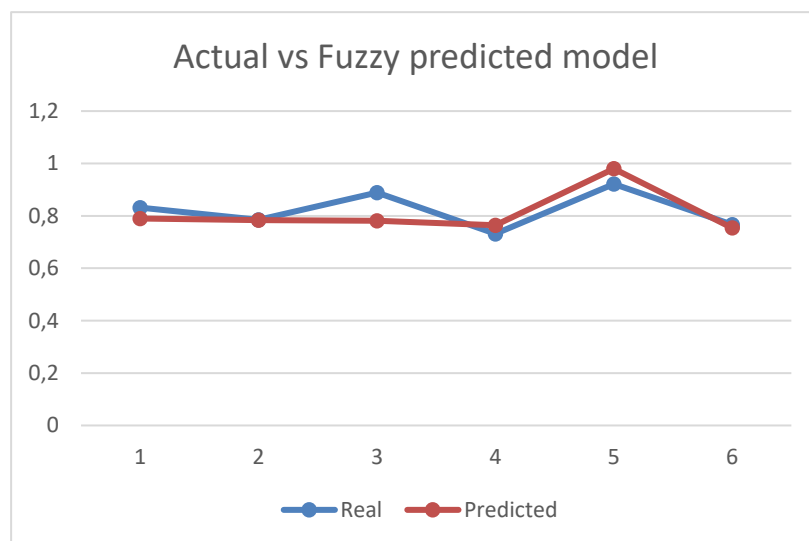


Figure 114: Illustration of comparison of the predicted values and actual values for the test data on for the WAAM.

In the **Table 5** the real data points and the estimated values are shown. This is while the error and residual values are calculated accordingly for this fuzzy inference engine. The mean squared error MSE is equal to 0.002974 and the standard error or root mean square error RMSE is equal to 0.05453. Almost 11 percent of the real data point are used for the tests.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Error_i)^2}{n}}$$

$$Error_i = Predicted\ value_i - Actual\ value_i$$

$$i = data\ point\ index$$

$$n = number\ of\ data\ points$$

Table 5: Table of prediction results.

Data point	Real	Predicted	Error	SE		
1	0.831387481	0.79	-0.041387481	0.001713		
2	0.78487825	0.784	-0.00087825	7.71E-07		
3	0.889333568	0.782	-0.107333568	0.01152		
4	0.731726961	0.764	0.032273039	0.001042		
5	0.921642741	0.98	0.058357259	0.003406		
6	0.766647703	0.754	-0.012647703	0.00016	MSE	RMSE
					0.002974	0.05453

In conclusion, the application of the fuzzy inference systems and the DACM method besides the ANOVA technique is presented in this case study. This is while the main influential parameters and their corresponding ranges after being obtained by the main effect plot and interaction plot are used in the construction of rules and allocation of membership functions. And finally, with the rules constructed, a predicting model is offered by the inference engine to estimate the high range values of the Height average in the WAAM process. In addition, it has to be mentioned that the rules and the membership functions are only allocated based on the obtained ranges while this is required to be investigated in future works on how to allocate the membership functions and to define the flexibility of these function on the ranges of values. On the other hand, the allocated rules must be investigated for other different ranges, and the best performance for the rules and their interaction must be defined. Since the rules have the option to receive weights based on the amount of validity they have within this system.

## CONCLUSIONS AND PROSPECTIVE WORKS

In this section the conclusion for the thesis is presented. Keynotes and results are addressed briefly and finally the future works and the implementations are discussed.

### Conclusion

As what has been undertaken so far, this thesis has addressed a detailed overview of the state of the arts and the existing technologies for the additive manufacturing. Also, after that two main processes as FDM and WAAM have been selected for the implementation of the methodologies existing for modeling and optimization of the processes. Subsequently, some details about these processes have been presented with the background and the description of techniques and the environment. Then, the existing methodologies have been presented for the artificial neural networks and other methods with an illustration of advantageous and disadvantageous.

Furthermore, six cases with multiple approximations and modeling techniques have been tested on the data which has been obtained through experimental plans, full factorial, and Tagouchi method. Multi-layer perceptron and back propagation method has been implemented in several cases with known relation and FDM processes and also a physical problem to check the accuracy of the estimations. Meanwhile, the results have been compared with some possible classical learning and approximation methods such as surface response technique.

In addition, the data have been tested with a radial based neural network due to the low amount of learning samples and the ability of these networks in modeling and dealing with such data sets. The best performance models have been obtained, and the results are compared with a known relation and approximation.

At the end of the last case study, one possible solution for knowledge extraction using statistical methods was presented, and the knowledge base has been used in a fuzzy inference system. Though the time was short to redo the experiments and implant the neuro-fuzzy networks; however, the core inference engine which is the application in the neuro-fuzzy networks too has been tested with the experimental data obtained for the WAAM process. The rules have been constructed using the knowledge base and infused for approximation in the inference engine. The results have shown an excellent approximation model on the test data. However, it still opens the discussion on further investigation and tests to generalize the results for all the possible situations.

This work can be concluded with the presentation of artificial neural networks and their ability to model the additive manufacturing processes. With the models obtained it is possible to predict the values during the process and providing the solution optimizing the quality and quantity of the printings.

## Prospective works

As what has been conducted in this work the methodologies have been tested based on the design of the experiments. These designs have been conducted using prior knowledge existing in the literature and so forth. However, as presented in the primary cases the quality of the models and their approximation is very dependent on the experimental data. For this case, the data which was at hand only had three to four level values. This is while higher number levels for the involving parameters can help the neural networks to perform a much better approximation and can reduce the amount of error due to lack of training data.

Also, in this work, only a technique has been tested for the clustering the bead zones. This while it is also possible to conduct the non-supervised techniques too and compare the results and level of accuracy too.

Meanwhile, in the last case, the application of the fuzzy systems has been tested for the WAAM process and the experimental data. In this case, the accuracy of the approximation in addition to generalization is depending on the knowledge which is extracted. This is why it keeps the discussion open for further investigation on this route and how to provide better solutions and techniques for defining the rules. Besides, in this case the rules had interaction on different levels which include the risk of having a weak approximation for some intervals. In this case it is possible to investigate more the logical interactions to generate better and concise rules [111].

Finally, another factor which defines the quality of the inference engine is the allocation of the membership functions. Further investigation of the impact of the allocation of these functions on the approximation of the outputs on the WAAM process is kept as an open question and it is possible to investigate the adaptive neuro-fuzzy inference systems on new and modified experimental data.



## REFERENCES

- [1] J. Liljemark Mattsson, “Simplifications of Simulations in Additive Manufacturing : Wire Feed Additive Manufacturing on Thin Structures,” 2015.
- [2] G. Coykendall, “Additive Manufacturing: Processes and Standard Terminology,” 2012.
- [3] S. W. Williams, F. Martina, A. C. Addison, J. Ding, G. Pardal, and P. Colegrove, “Wire Arc Additive Manufacturing.”
- [4] W. E. Frazier, “Metal Additive Manufacturing: A Review,” *J. Mater. Eng. Perform.*, vol. 23, no. 6, pp. 1917–1928, Jun. 2014.
- [5] C. Emmelmann, J. Kranz, D. Herzog, and E. Wycisk, “Laser Additive Manufacturing of Metals,” Springer, Berlin, Heidelberg, 2013, pp. 143–162.
- [6] P. Bartolo *et al.*, “Biomedical production of implants by additive electro-chemical and physical processes,” *CIRP Ann.*, vol. 61, no. 2, pp. 635–655, Jan. 2012.
- [7] E. G. (2015) B. E. CEIT, “Customer Case Study Medical Facts Cranial Implants a Perfect Fit Thanks to Additive Manufacturing.”
- [8] D. Herzog, V. Seyda, E. Wycisk, and C. Emmelmann, “Additive manufacturing of metals,” *Acta Mater.*, vol. 117, pp. 371–392, Sep. 2016.
- [9] Instrumentaria, “Customer Case Study Medical Facts Help Is Fast at Hand Thanks to Additive Manufacturing: Alphaform Produces a Hip Replacement Designed by Instrumentaria.”
- [10] J. Kranz, D. Herzog, and C. Emmelmann, “Design guidelines for laser additive manufacturing of lightweight structures in TiAl6V4,” *J. Laser Appl.*, vol. 27, no. S1, p. S14001, Feb. 2015.
- [11] D. Appleyard, “Powering up on powder technology,” *Met. Powder Rep.*, vol. 70, no. 6, pp. 285–289, Nov. 2015.
- [12] D. L. D. Bourell, J. J. Beaman, M. C. Leu, and D. W. Rosen, “A brief history of additive manufacturing and the 2009 roadmap for additive manufacturing: looking back and looking ahead,” *US-Turkey Work. ...*, no. 2, pp. 2005–2005, 2009.
- [13] I. Gibson, D. W. D. W. Rosen, and B. Stucker, *Additive Manufacturing Technologies: Rapid Prototyping to Direct Digital Manufacturing*, vol. 54. 2009.
- [14] M. Shellabear and O. Nyrrhilä, “DMLS-DEVELOPMENT HISTORY AND STATE OF THE ART.”
- [15] T. Wohlers and T. Gornet, “History of Additive Manufacturing,” 2014.
- [16] B. Vayre, F. Vignat, and F. Villeneuve, “Identification on some design key parameters for additive manufacturing: application on Electron Beam Melting Selection and/or peer-review

- under responsibility of Professor Pedro Filipe do Carmo Cunha,” *Procedia CIRP*, vol. 7, pp. 264–269, 2013.
- [17] D. T. Pham and S. S. Dimov, *Rapid Manufacturing : the Technologies and Applications of Rapid Prototyping and Rapid Tooling*. Springer London, 2001.
- [18] D. Bong Kim, P. Witherell, R. Lipman, and S. C. Feng, “Streamlining the additive manufacturing digital spectrum: A systems approach,” *Addit. Manuf.*, 2014.
- [19] O. Kerbrat, P. Mognol, and J.-Y. Hascoët, “A new DFM approach to combine machining and additive manufacturing Title of the paper: A new DFM approach to combine machining and additive manufacturing A new DFM approach to combine machining and additive manufacturing,” *Comput. Ind.*, pp. 684–692, 2011.
- [20] A. Mital, *Product development : a structured approach to consumer product development, design, and manufacture*. .
- [21] G. A. O. Adam and D. Zimmer, “Design for Additive Manufacturing-Element transitions and aggregated structures,” *CIRP J. Manuf. Sci. Technol.*, vol. 7, pp. 20–28, 2014.
- [22] ISO, “17296-2:2013, ISO: Additive Manufacturing — General principles — Part 2: Overview on process categories and feedstock,” vol. 2013, no. X, 2013.
- [23] M. K. Thompson *et al.*, “Design for Additive Manufacturing: Trends, opportunities, considerations, and constraints,” *CIRP Ann.*, vol. 65, no. 2, pp. 737–760, Jan. 2016.
- [24] R. Hague, I. Campbell, and P. Dickens, “Implications on design of rapid manufacturing,” *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.*, vol. 217, no. 1, pp. 25–30, Jan. 2003.
- [25] C. Lindemann, U. Jahnke, M. Moi, and R. Koch, “Analyzing Product Lifecycle Costs for a Better Understanding of Cost Drivers in Additive Manufacturing.”
- [26] I. Gibson, D. Rosen, and B. Stucker, *Additive Manufacturing Technologies*. 2015.
- [27] G. E. Hilmas, J. L. Lombardi, R. A. Hoffman, and K. Stuffle, “RECENT DEVELOPMENTS IN EXTRUSION FREEFORM FABRICATION (EFF) UTILIZING NON-AQUEOUS GEL CASTING FORMULATIONS.”
- [28] M. A. Yardimci, S. I. Guceri, M. Agarwala, and S. C. Danforth, “Part quality prediction tools for fused deposition processing,” 1996, pp. 539–548.
- [29] P. Witherell, A. Narayanan, and J. Lee, “Using Metamodels to Improve Product Models and Facilitate Inferencing,” in *2011 Fifth IEEE International Conference on Semantic Computing (ICSC)*, 2011, pp. 506–513.
- [30] J. L. J. L., *Welding principles and applications*. Delmar Cengage Learning, 2012.
- [31] K. Weman, G. Lindén, and M. Institute of Materials, *MIG welding guide*. Woodhead, 2006.
- [32] P. W. Muncaster, *Practical TIG (GTA) welding*. .

- [33] K. Gupta, *Advanced manufacturing technologies : modern machining, advanced joining, sustainable manufacturing.* .
- [34] Fronius, “Cold Metal Transfer.”
- [35] P. Podrzaj, B. Jerman, and D. Klobcar, “Welding defects at friction stir welding,” *Metalurgija*, vol. 54, no. 2, pp. 387–389, 2015.
- [36] R. Kumar, S. Chattopadhyaya, S. Hloch, G. Krolczyk, and S. Legutko, “Wear characteristics and defects analysis of friction stir welded joint of aluminium alloy,” vol. 18, no. 1, pp. 128–135, 2016.
- [37] J. Mirapeix, P. B. García-Allende, A. Cobo, O. M. Conde, and J. M. López-Higuera, “Real-time arc-welding defect detection and classification with principal component analysis and artificial neural networks,” *NDT E Int.*, vol. 40, no. 4, pp. 315–323, Jun. 2007.
- [38] J. Mirapeix, A. Cobo, O. M. Conde, C. Jaúregui, and J. M. López-Higuera, “Real-time arc welding defect detection technique by means of plasma spectrum optical analysis,” *NDT E Int.*, vol. 39, no. 5, pp. 356–360, Jul. 2006.
- [39] H. Luo, H. Zeng, L. Hu, X. Hu, and Z. Zhou, “Application of artificial neural network in laser welding defect diagnosis,” *J. Mater. Process. Technol.*, vol. 170, no. 1–2, pp. 403–411, Dec. 2005.
- [40] A. A. Robotics, “Robotics IRB 4600 Industrial robot ,datasheet,” 2018.
- [41] ABB Robotics, “Product specification IRB 4600,” p. 70, 2014.
- [42] “IRBP A Workpiece positioner, data sheet, PDF.”
- [43] R. L. Plackett and J. P. Burman, “THE DESIGN OF OPTIMUM MULTIFACTORIAL EXPERIMENTS,” *Biometrika*, vol. 33, no. 4, pp. 305–325, 1946.
- [44] S. Dowdy, S. Wearden, and D. Chilko, *Statistics for research*, vol. 512. John Wiley & Sons, 2011.
- [45] P. S. Levy and S. Lemeshow, *Sampling of populations: methods and applications*. John Wiley & Sons, 2013.
- [46] D. C. Montgomery, *Design and analysis of experiments*. John Wiley & Sons, 2017.
- [47] R. K. Roy, *Design of experiments using the Taguchi approach: 16 steps to product and process improvement*. John Wiley & Sons, 2001.
- [48] V. N. Nair *et al.*, “Taguchi’s parameter design: a panel discussion,” *Technometrics*, vol. 34, no. 2, pp. 127–161, 1992.
- [49] A. Garg, K. Tai, and M. M. Savalani, “Formulation of bead width model of an SLM prototype using modified multi-gene genetic programming approach,” *Int. J. Adv. Manuf. Technol.*, vol. 73, no. 1–4, pp. 375–388, Jul. 2014.

- [50] D. Ding, Z. Pan, D. Cuiuri, and H. Li, “A multi-bead overlapping model for robotic wire and arc additive manufacturing (WAAM),” *Robot. Comput. Integr. Manuf.*, vol. 31, pp. 101–110, Feb. 2015.
- [51] A. Garg and K. Tai, “Selection of a robust experimental design for the effective modeling of nonlinear systems using Genetic Programming,” in *2013 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, 2013, pp. 287–292.
- [52] J. Gardan, “Additive manufacturing technologies: State of the art and trends,” *Addit. Manuf. Handb. Prod. Dev. Def. Ind.*, vol. 7543, pp. 149–168, 2017.
- [53] V. Vijayaraghavan, A. Garg, J. S. L. Lam, B. Panda, and S. S. Mahapatra, “Process characterisation of 3D-printed FDM components using improved evolutionary computational approach,” *Int. J. Adv. Manuf. Technol.*, vol. 78, no. 5–8, pp. 781–793, May 2015.
- [54] E. Coatanéa, R. Roca, H. Mokhtarian, F. Mokammel, and K. Ikkala, “A conceptual modeling and simulation framework for system design,” *Comput. Sci. Eng.*, vol. 18, no. 4, pp. 42–52, 2016.
- [55] K. Efthymiou, K. Sipsas, D. Mourtzis, and G. Chryssolouris, “On knowledge reuse for manufacturing systems design and planning: A semantic technology approach,” *CIRP J. Manuf. Sci. Technol.*, vol. 8, pp. 1–11, 2015.
- [56] J. Hirtz, R. B. Stone, D. A. McAdams, S. Szykman, and K. L. Wood, “A functional basis for engineering design: reconciling and evolving previous efforts,” *Res. Eng. Des.*, vol. 13, no. 2, pp. 65–82, 2002.
- [57] D. C. Karnopp, D. L. Margolis, and R. C. Rosenberg, *System dynamics: modeling, simulation, and control of mechatronic systems*. John Wiley & Sons, 2012.
- [58] D. Margolis and T. Shim, “A bond graph model incorporating sensors, actuators, and vehicle dynamics for developing controllers for vehicle safety,” *J. Franklin Inst.*, vol. 338, no. 1, pp. 21–34, Jan. 2001.
- [59] T. Segreto, A. Simeone, and R. Teti, “Principal component analysis for feature extraction and NN pattern recognition in sensor monitoring of chip form during turning,” *CIRP J. Manuf. Sci. Technol.*, vol. 7, no. 3, pp. 202–209, 2014.
- [60] S. Rokka Chhetri, A. Canedo, and M. Abdullah Al Faruque, “KCAD: Kinetic Cyber-Attack Detection Method for Cyber-Physical Additive Manufacturing Systems.”
- [61] M. Correa, C. Bielza, and J. Pamies-Teixeira, “Comparison of Bayesian networks and artificial neural networks for quality detection in a machining process,” *Expert Syst. Appl.*, vol. 36, no. 3, pp. 7270–7279, Apr. 2009.
- [62] C. W. Dawson and R. L. Wilby, “Hydrological modelling using artificial neural networks,” *Prog. Phys. Geogr.*, vol. 25, no. 1, pp. 80–108, Mar. 2001.
- [63] J. V. Tu, “Advantages and disadvantages of using artificial neural networks versus logistic

- regression for predicting medical outcomes,” *J. Clin. Epidemiol.*, vol. 49, no. 11, pp. 1225–1231, Nov. 1996.
- [64] G. Gamst, L. S. Meyers, and A. J. Guarino, “ANOVA AND RESEARCH DESIGN,” in *Analysis of Variance Designs: A Conceptual and Computational Approach with SPSS and SAS*, Cambridge University Press, 2008, pp. 3–8.
- [65] L. S. Kao and C. E. Green, “Analysis of variance: is there a difference in means and what does it mean?,” *J. Surg. Res.*, vol. 144, no. 1, pp. 158–70, Jan. 2008.
- [66] I. . Basheer and M. Hajmeer, “Artificial neural networks: fundamentals, computing, design, and application,” *J. Microbiol. Methods*, vol. 43, no. 1, pp. 3–31, Dec. 2000.
- [67] A. K. Jain, Jianchang Mao, and K. M. Mohiuddin, “Artificial neural networks: a tutorial,” *Computer (Long. Beach. Calif.)*, vol. 29, no. 3, pp. 31–44, Mar. 1996.
- [68] R. Hecht-Nielsen, “ON THE ALGEBRAIC STRUCTURE OF FEEDFORWARD NETWORK WEIGHT SPACES,” *Adv. Neural Comput.*, pp. 129–135, Jan. 1990.
- [69] D. Poole, A. Mackworth, and R. Goebel, *Computational Intelligence: A Logical Approach*. 1998.
- [70] E. Alpaydin, *Introduction to Machine Learning Second Edition*. 2010.
- [71] I. H. (Ian H. . Witten, E. Frank, M. A. (Mark A. Hall, and C. J. Pal, *Data mining : practical machine learning tools and techniques*. .
- [72] S. J. Russell and P. Norvig, *Artificial intelligence: a modern approach*. Malaysia; Pearson Education Limited, 2016.
- [73] F. Solla, D. Bertonecelli, V. Rampal, A. Voury-Pons, and C. M. Bertonecelli, “Supervised Learning Methods to Predict Outcomes in Orthopaedic Research,” 2018.
- [74] T. Hastie, R. Tibshirani, and J. Friedman, “Unsupervised learning,” in *The elements of statistical learning*, Springer, 2009, pp. 485–585.
- [75] P. Witherell, A. Narayanan, and J. Lee, “Using Metamodels to Improve Product Models and Facilitate Inferencing,” in *2011 IEEE Fifth International Conference on Semantic Computing*, 2011, pp. 506–513.
- [76] R. S. Sutton and A. G. Barto, *Reinforcement learning : an introduction*. MIT Press, 1998.
- [77] M. Niknafs, “Neural Network Optimization.” 2016.
- [78] S. S. Haykin and Simon, *Neural networks : a comprehensive foundation*. Macmillan, 1994.
- [79] M. H. Hassoun, *Fundamentals of artificial neural networks*. MIT Press, 1995.
- [80] D. O Hebb, “The Organization of Behavior A NEUROPSYCHOLOGICAL THEORY.”
- [81] A. J. Maren, “A Logical Topology of Neural Networks,” 1991.

- [82] F. U. Dowla and L. L. Rogers, *Solving problems in environmental engineering and geosciences with artificial neural networks*. MIT Press, 1995.
- [83] K. Swingler, *Applying neural networks : a practical guide*. Academic Press, 1996.
- [84] T. Masters, *Practical neural network recipes in C++*. .
- [85] J. Hirtz, R. B. Stone, D. A. Mcadams, S. Szykman, and K. L. Wood, “A Functional Basis for Engineering Design: Reconciling and Evolving Previous Efforts.”
- [86] W. F. Schmidt, S. Raudys, M. A. Kraaijveld, M. Skurikhina, and R. P. W. Duin, “Initializations, back-propagation and generalization of feed-forward classifiers,” in *IEEE International Conference on Neural Networks*, pp. 598–604.
- [87] I. A. Basheer and M. Hajmeer, “Artificial neural networks: Fundamentals, computing, design, and application,” *J. Microbiol. Methods*, vol. 43, no. 1, pp. 3–31, 2000.
- [88] M. J. L. M. Orr, “Introduction to radial basis function networks,” *Univ. Edinbg.*, pp. 1–7, 1996.
- [89] D. Lowe and D. Broomhead, “Multivariable functional interpolation and adaptive networks,” *Complex Syst.*, vol. 2, no. 3, pp. 321–355, 1988.
- [90] A. Bors, *Introduction of the Radial Basis Function (RBF) Networks*, vol. 1. 2001.
- [91] R. J. Schalkoff, *Artificial neural networks*. McGraw-Hill, 1997.
- [92] S. Pal and P. K. Srimani, “Neurocomputing motivation, models, and hybridization,” 1996.
- [93] L. Ngai and Y. Wong, “A numerical test method of California bearing ratio on graded crushed rocks using particle flow modeling,” 2015.
- [94] R. F. Gunst, “Response Surface Methodology: Process and Product Optimization Using Designed Experiments,” *Technometrics*, vol. 38, no. 3, pp. 284–286, Aug. 1996.
- [95] G. E. P. Box, W. G. Hunter, J. S. Hunter, and others, *Statistics for experimenters*. John Wiley and sons New York, 1978.
- [96] J. Kaur and M. Mahajan, “Hybrid of Fuzzy Logic and Random Walker Method for Medical Image Segmentation,” *Int. J. Image, Graph. Signal Process.*, vol. 7, no. 2, pp. 23–29, Jan. 2015.
- [97] L. A. Zadeh, “Fuzzy sets,” *Inf. Control*, vol. 8, no. 3, pp. 338–353, Jun. 1965.
- [98] T. J. Ross, “Membership Functions, Fuzzification and Defuzzification,” *Physica, Heidelberg*, 2000, pp. 48–77.
- [99] T. Peikert, H. Garbe, and S. Potthast, “Fuzzy-Based Risk Analysis for IT-Systems and Their Infrastructure,” *IEEE Trans. Electromagn. Compat.*, vol. 59, no. 4, 2017.
- [100] M. Negnevitsky, *Artificial Intelligence A Guide to Intelligent Systems*. 2005.

- [101] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *Int. J. Man. Mach. Stud.*, vol. 7, no. 1, pp. 1–13, Jan. 1975.
- [102] D. J. C. MacKay, "Bayesian interpolation," *Neural Comput.*, vol. 4, no. 3, pp. 415–447, 1992.
- [103] F. D. Foresee and M. T. Hagan, "Gauss-Newton approximation to Bayesian learning," in *Proceedings of the 1997 international joint conference on neural networks*, 1997, vol. 3, pp. 1930–1935.
- [104] E. Coatanéa, V. Tsarkov, S. Modi, D. Wu, G. G. Wang, and H. Jafarian, "Knowledge-Based Artificial Neural Network (KB-ANN) in Engineering: Associating Functional Architecture Modeling, Dimensional Analysis and Causal Graphs to Produce Optimized Topologies for KB-ANNs," in *Volume 1B: 38th Computers and Information in Engineering Conference*, 2018, p. V01BT02A020.
- [105] E. Coatanea, R. Roca, H. Mokhtarian, F. Mokammel, and K. Ikkala, "A Conceptual Modeling and Simulation Framework for System Design," *Comput. Sci. Eng.*, vol. 18, no. 4, pp. 42–52, Jul. 2016.
- [106] J. D. Anderson and J. Wendt, *Computational fluid dynamics*, vol. 206. Springer, 1995.
- [107] B. B. Shahriar, C. France, N. Valerie, C. Arthur, and G. Christian, "Toward improvement of the properties of parts manufactured by FFF (fused filament fabrication) through understanding the influence of temperature and rheological behaviour on the coalescence phenomenon," in *AIP Conference Proceedings*, 2017, vol. 1896, no. 1, p. 040008.
- [108] R. Yousef and K. Hindi, "Training radial basis function networks using reduced sets as center points," *Int. J. Inf. Technol.*, vol. 2, no. 1, pp. 21–35, 2005.
- [109] Mo-yuen Chow and Hahn Tram, "Application of fuzzy logic technology for spatial load forecasting," in *Proceedings of 1996 Transmission and Distribution Conference and Exposition*, pp. 608–614.
- [110] P. J. King and E. H. Mamdani, "The application of fuzzy control systems to industrial processes," *Automatica*, vol. 13, no. 3, pp. 235–242, May 1977.
- [111] Michel Lesty, "La nouvelle approche dans le choix des régresseurs de la régression multiple en présence d'interactions et de colinéarités," *MODULAD*, vol. Modulad 22, p. pp.41-77, 1999.