



**Titre:** Statistical Analysis of Machining Processes of Composite Materials  
Title:

**Auteur:** Randa Seif  
Author:

**Date:** 2014

**Type:** Mémoire ou thèse / Dissertation or Thesis

**Référence:** Seif, R. (2014). Statistical Analysis of Machining Processes of Composite Materials  
Citation: [Mémoire de maîtrise, École Polytechnique de Montréal]. PolyPublie.  
<https://publications.polymtl.ca/1606/>

 **Document en libre accès dans PolyPublie**  
Open Access document in PolyPublie

**URL de PolyPublie:** <https://publications.polymtl.ca/1606/>  
PolyPublie URL:

**Directeurs de  
recherche:** Soumaya Yacout, & Helmi Attia  
Advisors:

**Programme:** Génie industriel  
Program:

UNIVERSITÉ DE MONTRÉAL

STATISTICAL ANALYSIS OF MACHINING PROCESSES OF COMPOSITE MATERIALS

RANDA SEIF

DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION

DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES

(GÉNIE INDUSTRIEL)

OCTOBRE 2014

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

STATISTICAL ANALYSIS OF MACHINING PROCESSES OF COMPOSITE MATERIALS

présenté par : SEIF Randa

en vue de l'obtention du diplôme de : Maîtrise ès science appliquées

a été dûment accepté par le jury d'examen constitué de :

M. ADJENGUE Luc-Désiré, Ph. D., président

Mme YACOUT Soumaya, D. Sc., membre et directeur de recherche

M. ATTIA Helmi, Ph. D., membre et codirecteur de recherche

M. OUALI Mohamed-Salah, Doctorat, membre

## ACKNOWLEDGEMENTS

As I prepare this thesis for submission, I intermittently glance over the television coverage of the search for flight Malaysian Air MH370 suspected to have gone down over the Pacific Ocean and I can't help but remark how far the minds of engineers, the perseverance of researchers and the ideas of visionaries have come together to allow a machine, like the immense Boeing 777, to carry us from one place on earth to another in a relatively very short period of time and connect civilizations and humanities across our planet.

However, the hardships faced in the search mission continuously remind us that we have so much more to explore and focuses our innovative minds to new ideas and directions.

With this thought in mind, I dedicate this thesis to all engineers, academics and visionaries that spend each day looking for new ways to bettering our lives and society.

I would also like to thank my supervisors, Dr Soumaya Yacout and Dr Helmi Attia, for their support, ideas, understanding, and dedication to my success. Also, I am grateful to Dr Mouhab Meshreki for his continuous technical support during the whole process.

I also would like to acknowledge the polytechnic school of engineering, its entire academic, non-academic staff, and its students for providing me with the resources to complete this work and be recognized within my field.

For my family and friends, I would like to dedicate this work for the support at home and their patience throughout this process.

To my loving husband, Dr. Roland Nassim, the publication of these findings are a true testament to the passion to research in academia that we both share in life.

This work may have looked at a very specific area within the engineering field, but the implications can and will save lives down the road. And for that, I will be forever proud.

## RÉSUMÉ

Le perçage dans les matériaux composites constitue un prérequis essentiel pour faciliter leur assemblage. L'un des principaux défis en matière de perçage est de fournir une excellente finition du produit et de minimiser les coûts de production.

L'objectif de ce mémoire est d'étudier la relation entre la vitesse d'avance/vitesse de rotation et les caractéristiques de qualité du processus de perçage définies par: le délaminage à l'entrée et la sortie, la rugosité de la surface, l'erreur du diamètre à l'entrée et la sortie, et la circularité à l'entrée et la sortie. En outre, trois variables mesurables, non contrôlables par l'opérateur, (force de poussée, force de découpe et moment de torsion) sont analysées afin de comprendre comment ils réagissent au changement de la vitesse d'avance et de rotation, ainsi que la façon dont ils affectent les caractéristiques de qualité. Les méthodes de régression linéaire, linéaire multiple et non linéaire sont développées pour comprendre l'effet et l'importance de chaque variable d'entrée sur les caractéristiques de qualité du processus. Ce mémoire présente la collecte de données faite durant le perçage, leur modélisation mathématique et leur analyse statistique. En plus, le rôle et l'implication de chaque caractéristique de qualité dans le processus de perçage sont documentés.

La méthode de régression linéaire multiple a démontré des bons résultats dans le cas des variables suivantes: force de poussée, force de découpe, moment de torsion, délaminage à l'entrée et sortie. Cette application n'est pas recommandée pour l'analyse ainsi que la prédiction des autres variables. De plus, les résultats indiquent d'une part que la vitesse d'avance a énormément d'impact sur toutes les variables de sorties étudiées, y compris les variables mesurables, à l'exception de la circularité à l'entrée. D'autre part, l'impact de la vitesse de rotation s'est avéré significatif sur les variables suivantes: la force de poussée, le délaminage à la sortie, l'erreur de diamètre à l'entrée et la sortie et la circularité à la sortie. Ainsi, dans le but de combiner les variables d'entrées avec les variables mesurables, l'application de l'analyse de covariance est développée, mais les résultats se sont avérés non concluants en raison de la forte corrélation présente entre les variables d'entrée. Par conséquent, la méthode de covariance est incohérente et non valable dans le processus du perçage. En conclusion, les résultats établissent un modèle de prédiction mathématique qui explique et surtout quantifier l'influence des variables d'entrée sur les caractéristiques de qualité du processus de perçage. Cette étude démontre également comment ces variables sont reliées et corrélées entre eux.

## ABSTRACT

Hole drilling in composite materials is an essential requirement in facilitating their assembly. One of the main challenges in drilling is providing an excellent product finish and achieving cost effectiveness.

The purpose of this dissertation is to investigate the relationship between the feed rate/spindle speed and the drilling quality characteristics as defined by: delamination at entry, delamination at exit, surface roughness, diameter at entry, diameter at exit, circularity at entry and circularity at exit. Also, three measurable variables (thrust force, cutting force and torque) are analyzed to understand how they relate to the feed rate and the spindle speed as well as how they affect the other quality characteristics. Multiple linear and nonlinear regression techniques are used to understand the effect and the importance of each input variable on the quality outputs. Data collection, mathematical modeling and statistical analysis were utilized in this dissertation. Furthermore, the role of each quality characteristics in drilling process is documented as well as the measurable variables.

Results show that the multiple regression models showed significance in a subset of the outputs (thrust force, cutting force, torque, delamination at entry and delamination at exit) and proved insignificant on the study of the other drilling quality characteristics. The results also indicate that of all the variables studied including the measurable variables, the feed rate appeared to be the most significant on the outputs except for the circularity at entry. On the other hand, the spindle speed was shown to impact significantly the following variables: thrust force, delamination at exit, diameter error at exit and entry, and circularity at exit. Furthermore, the application ANCOVA methodology is investigated on the drilling outputs but it provided inconclusive results due to strong correlation among the input variables. Therefore, the ANCOVA methodology was deemed not suitable in the study of this drilling process.

In conclusion, the results establish a mathematical prediction model that explains and more importantly quantifies the influence of the various variables (inputs) on the quality characteristics of the drilling process (outputs). This novel model also demonstrates how these variables are related amongst each other.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	iii
RÉSUMÉ.....	iv
ABSTRACT .....	v
LIST OF TABLES .....	x
LIST OF FIGURES .....	xvi
LIST OF ABBREVIATIONS.....	xxi
LIST OF ANNEXES .....	xxiii
CHAPTER 1 INTRODUCTION .....	1
A. Organization of the dissertation.....	2
B. Statement of the problem and process parameters .....	2
C. Description of the tool .....	5
D. Description of the machine .....	6
E. Description of the material .....	6
F. Research objectives.....	6
CHAPTER 2 LITERATURE REVIEW .....	8
CHAPTER 3 REGRESSION ANALYSIS TECHNIQUE WITH CONTROLLABLE VARIABLES AS INPUTS.....	18
3.1 Thrust force analysis.....	27
3.1.1 Thrust force distribution study over feed rate .....	27
3.1.2 Thrust force distribution study over spindle speed.....	28
3.1.3 Thrust force multiple linear regression analysis .....	30
3.1.4 Thrust force multiple linear regression analysis with interaction .....	35
3.1.5 Thrust force data transformation.....	38
3.1.6 Comparison of the methods used for thrust force .....	42

3.2 Cutting force analysis.....	42
3.2.1 Cutting force distribution study over feed rate .....	42
3.2.2 Cutting force distribution study over spindle speed.....	43
3.2.3 Cutting force multiple linear regression analysis.....	44
3.2.4 Cutting force multiple linear regression with interaction.....	48
3.2.5 Comparison of the methods used for cutting force data .....	51
3.3 Torque analysis .....	52
3.3.1 Torque distribution study over feed rate.....	52
3.3.2 Torque distribution study over spindle speed .....	53
3.3.3 Torque multiple linear regression analysis .....	53
3.3.4 Torque multiple linear regression with interaction .....	57
3.3.5 Comparison of the methods used for torque.....	60
3.4 Delamination at entry analysis.....	60
3.4.1 Delamination at entry distribution study over feed rate.....	60
3.4.2 Delamination at entry distribution study over spindle speed .....	61
3.4.3 Delamination at entry multiple linear regression analysis .....	62
3.4.4 Delamination at entry multiple linear regression with interaction.....	66
3.4.5 Delamination at entry nonlinear regression.....	69
3.5 Delamination at exit analysis .....	72
3.5.1 Delamination at exit distribution study over feed rate .....	72
3.5.2 Delamination at exit distribution study over spindle speed .....	73
3.5.3 Delamination at exit multiple regression analysis .....	74
3.5.4 Delamination at exit multiple linear regression analysis with interaction.....	77
3.6 Surface roughness analysis .....	81
3.6.1 Surface roughness distribution study over feed rate .....	81

3.6.2 Surface roughness distribution study over spindle speed .....	82
3.6.3 Surface roughness multiple linear regression analysis .....	82
3.6.4 Surface roughness multiple linear regression analysis with interaction.....	86
3.7 Diameter error at exit analysis .....	93
3.7.1 Diameter error at exit distribution study over feed rate .....	93
3.7.2 Diameter error at exit distribution study over spindle speed.....	94
3.7.3 Diameter error at exit multiple linear regression analysis.....	95
3.7.4 Diameter error at exit transformation .....	97
3.8 Diameter error at entry analysis .....	99
3.8.1 Diameter error at entry distribution study over feed rate .....	99
3.8.2 Diameter error at entry distribution study over spindle speed .....	100
3.8.3 Hole diameter error at entry multiple linear regression analysis.....	101
3.8.4 Diameter error at entry transformation .....	103
3.9 Circularity at exit analysis.....	104
3.9.1 Circularity at exit distribution study over feed rate.....	104
3.9.2 Circularity at exit distribution study over spindle speed .....	105
3.9.3 Circularity at exit multiple linear regression.....	106
3.10 Circularity at entry analysis .....	113
3.10.1 Circularity at entry distribution study over feed rate.....	113
3.10.2 Circularity at entry distribution study over spindle speed .....	114
3.10.3 Circularity at entry multiple linear regression.....	115
3.11 Delamination at exit analysis versus thrust force and cutting force .....	119
3.11.1 Delamination at exit distribution study over thrust force.....	119
3.11.2 Delamination at exit distribution study over cutting force .....	122
3.11.3 Delamination at exit distribution study over thrust force and cutting force .....	126

3.12 Delamination at entry analysis versus thrust force and cutting force .....	129
3.12.1 Delamination at entry distribution study over thrust force .....	129
3.12.2 Delamination at entry distribution study over cutting force.....	134
CHAPTER 4 CONCLUSION .....	138
REFERENCES.....	142

## LIST OF TABLES

Table 2.1: Experiment setup with 35 runs.....	12
Table 2.2: Experiment setup with 18 runs.....	12
Table 2.3: Correlation table between all variables for ANCOVA.....	17
Table 3.1: The relationship between the null hypothesis and the type of statistical errors.....	21
Table 3.2: Controllable factors involved in the drilling process.....	23
Table 3.3: Parameter estimates and ANOVA results of thrust force against feed rate.....	28
Table 3.4: Parameter estimates and ANOVA results of thrust force against spindle speed.....	29
Table 3.5: Parameter estimates and ANOVA results of the thrust force using MLR.....	31
Table 3.6: Shapiro-Wilk Test for thrust force normality using MLR.....	33
Table 3.7: Illustration of the assumption validation results for thrust force MLR.....	35
Table 3.8: Parameter estimates and ANOVA results for thrust force using MLR-interaction.....	36
Table 3.9: Shapiro-Wilk Test for Thrust Force Normality for MLR-interaction.....	37
Table 3.10: Illustration of the assumption validation results.....	38
Table 3.11: Goodness of fit of thrust force transformations.....	39
Table 3.12: Parameter estimates and ANOVA results of log-thrust force.....	39
Table 3.13: Shapiro-Wilk Test for log thrust force.....	40
Table 3.14: Illustration of the assumption validation results for log thrust force model.....	45
Table 3.15: Comparison of all methods used for the thrust force analysis.....	42
Table 3.16: Parameter estimates and ANOVA results of cutting force against feed rate.....	43
Table 3.17: Parameter estimates and ANOVA results of cutting force over spindle speed.....	43
Table 3.18: Parameter estimates and ANOVA results for cutting force using MLR.....	44
Table 3.19: Shapiro-Wilk test for cutting force using MLR.....	46
Table 3.20: Illustration of the assumption validation results for cutting force MLR.....	48
Table 3.21: Parameter estimates and ANOVA results for cutting force using MLR-interaction.....	48
Table 3.22: Shapiro-test for cutting force using MLR-interaction.....	50
Table 3.23: Illustration of the assumption validation results for cutting force using MLR- interaction.....	50
Table 3.24: Akaike's results for Thrust force model with multiple regression--no interaction.....	51
Table 3.25: Akaike's results for Thrust force model with multiple regression--with interaction.....	51
Table 3.26: Comparison of main methods used for cutting force analysis.....	51

Table 3.27: Parameter estimates and ANOVA results for the torque against feed rate.....	52
Table 3.28: Parameter estimates and ANOVA results for the torque against spindle speed.....	53
Table 3.29: Parameter estimates and ANOVA tables for the torque using MLR.....	54
Table 3.30: Shapiro-Wilk test for normal data for torque using MLR.....	56
Table 3.31: Illustration of the assumption validation results for torque using MLR.....	57
Table 3.32: Shapiro-Wilk test for torque using MLR-interaction.....	59
Table 3.33: Illustration of the assumption validation results for the torque MLR-interaction.....	59
Table 3.34: Comparison of all methods used for torque analysis force analysis.....	60
Table 3.35: Parameter estimates and ANOVA results for the delamination at entry against feed rate.....	61
Table 3.36: Parameter estimates and ANOVA results for the delamination at entry against spindle speed.....	62
Table 3.37: Parameter estimates and ANOVA results for delamination at entry using MLR.....	63
Table 3.38: Shapiro-Wilk test for normal data for delamination at entry.....	65
Table 3.39: Illustration of the assumption validation results.....	66
Table 3.40: Parameter estimates and ANOVA results for delamination at entry using MLR- interaction.....	67
Table 3.41: Shapiro-Wilk test for normal data for delamination at entry.....	68
Table 3.42: Illustration of the assumption validation results of delamination at entry using MLR- interaction.....	69
Table 3.43: Goodness of fit of the delamination at entry transformations.....	70
Table 3.44: Parameter estimates and ANOVA results of cubic transformation of the delamination at entry.....	70
Table 3.45: AIC of the squared transformation of the delamination at entry.....	70
Table 3.46: AIC of the MLR model of the delamination at entry.....	70
Table 3.47: AIC of the MLR model with interaction of the delamination at entry.....	71
Table 3.48: Critical values of the models developed for delamination at entry.....	71
Table 3.49: Comparison of all methods used for delamination at entry analysis.....	71
Table 3.50: Parameter estimates and ANOVA results for the delamination at exit against feed rate.....	72

Table 3.51: Parameter estimates and ANOVA results for the delamination at exit against spindle speed.....	73
Table 3.52: Parameter estimates and ANOVA results for delamination at exit using MLR.....	74
Table 3.53: Shapiro-Wilk test for delamination at exit using MLR.....	76
Table 3.54: Illustration of the assumption validation results of the delamination at exit using MLR.....	77
Table 3.55: Parameter estimates and ANOVA results for delamination at exit using MLR- interaction.....	78
Table 3.56: Shapiro-Wilk test for normal data for delamination at exit.....	79
Table 3.57: Illustration of the assumption validation results for delamination at exit MLR- interaction.....	80
Table 3.58: Goodness of fit of the delamination at exit transformations.....	80
Table 3.59: AIC of the MLR model of the delamination at exit.....	80
Table 3.60: AIC of the MLR model with interaction of the delamination at exit.....	80
Table 3.61: Comparison of all methods used for delamination at exit analysis.....	80
Table 3.62: Parameter estimates and ANOVA results of the surface roughness over feed rate....	81
Table 3.63: Parameter estimates and ANOVA results of the surface roughness over spindle speed.....	82
Table 3.64: Parameter estimates and ANOVA results for surface roughness using MLR.....	83
Table 3.65: Shapiro-Wilk test for normal data for surface roughness.....	85
Table 3.66: Illustration of the assumption validation results for the surface roughness using MLR.....	86
Table 3.67: Parameter estimates and ANOVA results for surface roughness using MLR- interaction.....	87
Table 3.68: Shapiro-Wilk test for surface roughness using MLR-interaction.....	88
Table 3.69: Illustration of the assumption validation results for surface roughness with MLR- interaction.....	89
Table 3.70: Outlier in surface roughness transformation.....	89
Table 3.71: DFBETA of feed rate and spindle speed in surface roughness transformation.....	89
Table 3.72: Parameter estimates and ANOVA results for the surface roughness with n=34.....	90
Table 3.73: Shapiro-Wilk Test for surface roughness with n=34.....	91

Table 3.74: Akaike's results for surface roughness model with MLR-interaction.....	92
Table 3.75: Akaike's results for surface roughness model with MLR n=34.....	92
Table 3.76: Comparison of all methods used for surface roughness.....	93
Table 3.77: Parameter estimates and ANOVA results for the diameter error at exit against feed rate.....	94
Table 3.78: Parameter estimates and ANOVA results for the diameter error at exit against spindle speed.....	95
Table 3.79: Parameter estimates and ANOVA results for diameter error at exit using MLR.....	96
Table 3.80: Testing results of the transformations of the diameter error at exit.....	97
Table 3.81: Test of the transformation of the feed rate.....	98
Table 3.82: Parameter estimates and ANOVA results for diameter error at exit.....	98
Table 3.83: Comparison of all methods used for the diameter error at exit.....	99
Table 3.84: Parameter estimates and ANOVA results for diameter error at entry against feed rate.....	100
Table 3.85: Parameter estimates and ANOVA results for diameter error at entry against spindle speed.....	101
Table 3.86: ANOVA and parameter estimates tables for the diameter error at entry using MLR.....	102
Table 3.87: Results for the transformation of the diameter error at entry.....	103
Table 3.88: BFBETAs for diameter error at entry.....	104
Table 3.89: Comparison of all methods used for the diameter error at entry analysis.....	104
Table 3.90: Parameter estimates and ANOVA results for circularity at exit against feed rate...	105
Table 3.91: Parameter estimates and ANOVA results for circularity at exit against spindle speed.....	106
Table 3.92: Parameter estimates and ANOVA results for the diameter error at exit MLR.....	106
Table 3.93: Shapiro-Wilk test for normal data of the circularity at exit data using MLR.....	108
Table 3.94: Illustration of the assumption validation results for circularity at exit using MLR..	109
Table 3.95: Values of DFBETA of feed rate and spindle speed for circularity at exit.....	110
Table 3.96: Parameter estimates and ANOVA results for hole circularity at exit with n=34.....	111
Table 3.97: Akaike's results for circularity at exit model with MLR.....	112
Table 3.98: Akaike's results for circularity at exit model with multiple regression-n=34.....	112

Table 3.99: Comparison of all methods used for the circularity at exit analysis.....	113
Table 3.100: Parameter estimates and ANOVA results for circularity at entry over feed rate...	114
Table 3.101: Parameter estimates and ANOVA results for circularity at entry over spindle speed.....	115
Table 3.102: Parameter estimates and ANOVA results for circularity at entry using MLR.....	115
Table 3.103: Shapiro-Wilk test for the circularity at entry.....	117
Table 3.104: Illustration of the assumption validation results for circularity at exit.....	118
Table 3.105: Comparison of all methods used for the circularity at entry analysis.....	118
Table 3.106: Parameter estimates and ANOVA results of delamination at exit against thrust force.....	119
Table 3.107: Shapiro-Wilk Test for delamination at exit normality delamination at exit against thrust force.....	120
Table 3.108: Illustration of the assumption validation results for delamination at exit-simple regression.....	121
Table 3.109: Parameter estimates and ANOVA results of cutting force against delamination at exit.....	122
Table 3.110: Shapiro-Wilk Test for cutting force normality using simple regression.....	123
Table 3.111: Illustration of the assumption validation results for delamination at exit versus cutting force using simple regression.....	124
Table 3.112: Parameter estimates and ANOVA results of delamination at exit against cutting force using nonlinear regression.....	125
Table 3.113: Parameter estimates and ANOVA results of thrust force and cutting force against delamination at exit.....	126
Table 3.114: Shapiro-Wilk Test for delamination at exit normality using MLR.....	127
Table 3.115: Parameter estimates and ANOVA results of delamination at entry against thrust force.....	129
Table 3.116: Shapiro-Wilk Test for delamination at entry normality using simple regression...	130
Table 3.117: Illustration of the assumption validation results for delamination at entry with simple regression.....	131
Table 3.118: Parameter estimates and ANOVA results of transformed thrust force against delamination at entry.....	132

Table 3.119: Shapiro-Wilk test for normal data for delamination at entry.....	133
Table 3.120: Illustration of the assumption validation results.....	134
Table 3.121: Parameter estimates and ANOVA results of cutting force against delamination at entry.....	134
Table 3.122: Shapiro-Wilk Test for cutting force normality using simple regression.....	135
Table 3.123: Illustration of the assumption validation results for delamination at entry vs cutting force using simple regression.....	136
Table 3.124: Presentation of all analysis completed for the delamination at entry against thrust force and cutting force.....	137
Table 4.1: Coded values for the feed rate and spindle speed settings.....	139
Table 4.2: Relationship between independent variables and outputs.....	139
Table 4.3: Models for the outputs studied.....	140

## LIST OF FIGURES

Figure 1.1: Representation of the delamination factor and mechanism .....	3
Figure 1.2: Surface Roughness illustration.....	3
Figure 1.3: Representation of the hole circularity errors.....	4
Figure 1.4: Representation of the hole diameter error.....	4
Figure 1.5: Two Flute Standard Point Solid Carbide Twist Drill of 5 mm diameter .....	6
Figure 1.6: Makino A88.....	6
Figure 2.1: The basic geometric representation for a box-behnken design for three factors.....	10
Figure 3.1: Thrust force boxplot against feed rate.....	28
Figure 3.2: Thrust force boxplot against spindle speed.....	29
Figure 3.3: Thrust force plot fitted against observed values using MLR.....	32
Figure 3.4: Plot of residuals of the thrust force data versus run order.....	32
Figure 3.5: Normal probability plot for the thrust force data using MLR.....	33
Figure 3.6: Residual plot for the thrust force data using multiple regression.....	34
Figure 3.7: Scatterplots of thrust force versus feed rate and spindle speed.....	34
Figure 3.8: Thrust force plot fitted against observed values using MLR-interaction.....	36
Figure 3.9: Normal probability plot of standardized residuals for the thrust force with MLR- interaction.....	37
Figure 3.10: Thrust force fitted values against standardized residuals.....	38
Figure 3.11: Plot of the log thrust force fitted model against the feed rate and spindle speed.....	40
Figure 3.12: Plot of residual vs fitted values for log thrust force transformation.....	41
Figure 3.13: Scatterplot of feed rate and spindle speed against log thrust force.....	41
Figure 3.14: Boxplot of feed rate against cutting force.....	42
Figure 3.15: Boxplot of cutting force against spindle speed.....	44
Figure 3.16: Plot of the observed vs predicted values for cutting force using MLR.....	45
Figure 3.17: Plot of standardized residuals of the cutting force versus case number.....	45
Figure 3.18: Normality plot of the standardized residuals of the cutting force data.....	46
Figure 3.19: Cutting force fitted values versus residuals.....	47
Figure 3.20: Scatterplots feed rate and spindle speed against cutting force.....	47
Figure 3.21: Observed vs predicted values of the cutting force using MLR-interaction.....	49
Figure 3.22: Normality plot for cutting force residuals using MLR-interaction.....	49

Figure 3.23: Plot of fitted against residual values for cutting force with MLR-interaction.....	50
Figure 3.24: Boxplot of the torque against the feed rate.....	52
Figure 3.25: Boxplot of the torque against the spindle speed.....	53
Figure 3.26: Plot of observed versus predicted values for torque using MLR.....	55
Figure 3.27: Plot of the standardized residuals of the fitted torque against the run order of the data collection.....	55
Figure 3.28: Normality plot for the torque experimental data.....	56
Figure 3.29: Observed vs residual values plot of the torque data using MLR.....	56
Figure 3.30: Scatterplots feed rate and spindle speed against torque.....	57
Figure 3.31: Plot of torque observed against predicted values using MLR-interaction.....	58
Figure 3.32: Normal probability plot of the torque using the MLR-interaction.....	59
Figure 3.33: Predicted torque values against residuals using MLR-interaction.....	59
Figure 3.34: Boxplot of delamination at entry against feed rate.....	61
Figure 3.35: Boxplot of delamination at entry against spindle speed.....	62
Figure 3.36: Delamination at entry observed values against predicted using MLR.....	64
Figure 3.37: Plot of the standardized residuals of the fitted delamination at entry against the run order of the data collection.....	64
Figure 3.38: Normality plot for the delamination at entry.....	65
Figure 3.39: Predicted vs standardized residual values plot of the delamination at entry.....	65
Figure 3.40: Scatterplots feed rate and spindle speed vs delamination at entry.....	66
Figure 3.41: Delamination at entry observed against predicted values using MLR- interaction.....	67
Figure 3.42: Normality plot for delamination at entry using MLR-interaction.....	68
Figure 3.43: Residual plot versus fitted data of delamination at entry using MLR-interaction....	68
Figure 3.44: Boxplot of delamination at exit against feed rate.....	72
Figure 3.45: Boxplot of delamination at exit against spindle speed.....	74
Figure 3.46: Delamination at exit estimated plot using MLR.....	75
Figure 3.47: Plot of the standardized residuals of the fitted Delamination at exit against the run order of the data collection.....	75
Figure 3.48: Normality plot for the delamination at exit with MLR.....	76
Figure 3.49: Predicted vs standardized residual values plot of the delamination at exit.....	76

Figure 3.50: Scatterplot feed rate and spindle speed vs delamination at exit.....	77
Figure 3.51: Predicted vs standardized residual values plot of the delamination at exit experimental data.....	78
Figure 3.52: Normality plot for the delamination at exit with MLR-interaction.....	79
Figure 3.53: Predicted vs standardized residual values plot of the delamination at exit using MLR-interaction.....	79
Figure 3.54: Boxplot of surface roughness against feed rate.....	81
Figure 3.55: Boxplot of surface roughness against spindle speed.....	82
Figure 3.56: Surface roughness observed values against predicted using MLR.....	84
Figure 3.57: Plot of the standardized residuals of the fitted surface roughness against the run order of the data collection.....	84
Figure 3.58: Normality plot for the surface roughness experimental data.....	85
Figure 3.59: Predicted vs standardized residual values plot of the delamination at exit.....	85
Figure 3.60: Scatterplots feed rate and spindle speed vs hole surface roughness.....	86
Figure 3.61: Surface roughness observed values against predicted using MLR-interaction.....	87
Figure 3.62: Normality plot of the surface roughness data using MLR-interaction.....	88
Figure 3.63: Predicted vs standardized residual values plot of the surface roughness using MLR- interaction.....	88
Figure 3.64: Surface roughness observed values against predicted using MLR with n=34.....	90
Figure 3.65: Plot of normality of the surface roughness with n=34.....	91
Figure 3.66: Residual plot of the surface roughness experimental data with n=34.....	92
Figure 3.67: Boxplot of diameter error at exit against feed rate.....	93
Figure 3.68: Boxplot of diameter error at exit against spindle speed.....	94
Figure 3.69: Predicted vs standardized residual values plot of the diameter error at exit.....	97
Figure 3.70: Residual plot for the diameter error at exit experimental data.....	99
Figure 3.71: Diameter error at entry boxplot against the feed rate.....	100
Figure 3.72: Diameter error at entry boxplot with the spindle speed.....	101
Figure 3.73: Residual plot for diameter error at entry using MLR.....	103
Figure 3.74: Boxplot of the circularity at exit over feed rate.....	104
Figure 3.75: Boxplot of the circularity at exit over spindle speed.....	105
Figure 3.76: Plot of the circularity at exit fitted vs observed values using MLR.....	107

Figure 3.77: Plot of the standardized residuals of the circularity at exit data against the number of cases.....	107
Figure 3.78: Normal plot of the standardized residuals of the circularity at exit data.....	108
Figure 3.79: Residuals plot for circularity at exit.....	108
Figure 3.80: Scatterplots feed rate and spindle speed against circularity at exit.....	109
Figure 3.81: Plot of the leverage against normalized residual squared of the circularity at exit.....	109
Figure 3.82: Plot of the circularity at exit fitted vs observed values n=34.....	112
Figure 3.83: Boxplot of circularity at entry over the feed rate.....	113
Figure 3.84: Boxplot of circularity at entry over the spindle speed.....	114
Figure 3.85: Plot of the circularity at entry fitted vs observed values using MLR.....	116
Figure 3.86: Plot of standardized residuals of the circularity at entry against case number.....	116
Figure 3.87: Plot of standardized residuals of the circularity at entry.....	117
Figure 3.88: Scatterplots feed rate and spindle speed against circularity at entry.....	117
Figure 3.89: Residuals plot for circularity error at entry.....	118
Figure 3.90: Plot of the thrust force against the delamination at exit.....	120
Figure 3.91: Normal probability plot for the delamination at exit simple regression.....	120
Figure 3.92: Residual plot for the delamination at exit data using simple regression.....	122
Figure 3.93: Scatterplot of thrust force versus feed rate.....	122
Figure 3.94: Plot of the cutting force against the delamination at exit.....	123
Figure 3.95: Normal probability plot for the cutting force data against the cutting force.....	123
Figure 3.96: Residual plot for the cutting force versus delamination at exit data using simple linear regression.....	124
Figure 3.97: Scatterplot of cutting force against delamination at exit.....	124
Figure 3.98: Plot of the quadratic delamination at exit fitted model.....	126
Figure 3.99: Plot of the delamination at exit using MLR.....	127
Figure 3.100: Normal probability plot for the delamination at exit using MLR.....	127
Figure 3.101: Normal probability plot for the delamination at exit using MLR.....	127
Figure 3.102: Residual plot for the delamination at exit data using MLR.....	128
Figure 3.103: Test of collinearity.....	128
Figure 3.104: Plot of the thrust force against the delamination at entry.....	130

Figure 3.105: Normal probability plot for the delamination at entry using simple regression with thrust force.....	130
Figure 3.106: Plot of the delamination at entry observed vs predicted values against thrust force transformation.....	132
Figure 3.107: Normal plot of the standardized residuals of the delamination at entry.....	133
Figure 3.108: Scatterplot of the delamination at entry against the log (thrust force).....	133
Figure 3.109: Residuals plot for circularity at exit.....	134
Figure 3.110: Plot of the cutting force against the delamination at entry.....	135
Figure 3.111: Normal probability plot for the delamination at entry vs cutting force data using simple regression.....	135
Figure 3.112: Residual plot for the delamination at entry versus cutting force data using simple regression.....	136
Figure 3.113: Scatterplot of cutting force against delamination at entry.....	136

## LIST OF ABBREVIATIONS

MLR	Multiple Linear Regression
ANOVA	Analysis of variance
RSM	Response Surface Model
NRC	National Research Council of Canada
CFRP	Carbon Fiber Reinforced Polymer
BBD	Box-Behnken Design
ANN	Artificial neural network
SVR	Support Vector Regression
FP	False Positive
FN	False Negative
SRM	Structural Risk Minimization
SVM	Support Vector Machines
$R^2_{adj}$	Adjusted coefficient of determination
Ln	Logarithmic scale
ANCOVA	Analysis of covariance
Tf	Thrust force
Cf	Cutting force
T	Torque
Den	Delamination at entry
Dex	Delamination at exit
SR	Surface Roughness
Drx	Diameter Error at Exit
Dre	Diameter Error at Entry
Cex	Circularity at Exit
Cen	Circularity at Entry
$d_{max}$	Maximum diameter
d	Hole diameter
SS	Sum Square
df	Degrees of Freedom
MS	Mean Square

AIC	Akaike Information criterion
BPCW	Breusch-Pagan/Cook-Weisberg

**LIST OF ANNEXES**

ANNEX A.....	146
ANNEX B.....	147
ANNEX C.....	149

## CHAPTER 1 INTRODUCTION

Over the years, composites were found to be one of the most important materials used in many critical industries: infrastructure, aerospace, and military applications because of its several structural advantages and qualities: rigidity, durability, lightness, corrosion resistance and hardness. That being said, the cost remains relatively high. Composites are products made from two major constituents: fiber and matrix whereas the matrix provides protection of the fiber. This combination provides the final product built with the composites material higher and surpassing performance than the starting materials. The structural behavior depends on the properties of the fibers (their amount and orientations) and the matrix, and is manufactured in distinct layers. These different layers or plies bonded together form a laminate [1].

The fabrication methods in these different industries range from very simple to complex processes with expensive operations equipment. Machining of composites materials is one of the critical processes. The most common machining processes are drilling, turning, cutting and milling. The process selected to fabricate the end product using the composite part is dependent upon factors such as the design requirements, part complexity and surface finish and appearance. The manufacturing method analyzed in this research is drilling.

Drilling is the most popular conventional machining process and one of the most essential metal-cutting operations, comprising approximately 33% [2] of all metal-cutting operations. The drilling process is mainly characterized by the feed rate, the cutting speed, the tool (coated/uncoated), the laminate design and the cooling strategy. These factors are known to be controllable because they are defined by the process experts and can be regulated by the operator. Also, other uncontrollable factors, which can only be measured during the running of the process, are involved. These factors are the thrust force, the torque, the cutting force, the cutting temperature, and the tool wear. At the process monitor and control level, most of the hole's defects observed are surface delamination, circularity errors, surface roughness and diameter error. These undesirable imperfections need to be investigated to avoid the failure of the material structure. In an aircraft construction for example, the holes must be drilled with keen attention to ensure minimum defects based on pre-defined design and manufacturing requirements, and error tolerance. Failure to do so can result in the parts becoming scrapped and major disaster of the whole product dysfunction. This study addresses the problems related to the characterizations

presented above to develop an approach for damage-free drilling composites by predicting their comportment.

## **A. Organization of the dissertation**

Feed rate and spindle speed are the only two parameters considered in this study as the controllable and explanatory variables. Accordingly, in **chapter 1**, the quality characteristics of the delamination at entry, the delamination at exit, the hole surface roughness (in microns), the hole diameter error at exit (in %), the hole diameter error at entry (in %), the hole circularity at exit (in %) and the hole circularity at entry (in %) are explained to understand their effect on the drilling process performance. Also, the uncontrollable variables (thrust force, cutting force and torque) impacts are explored.

In **chapter 2**, a literature review of past studies, facts and experiments are presented to identify the researcher's methodologies and works in modeling drilling operation.

In **chapter 3**, analytical models of the uncontrollable variables and the quality characteristics are developed using the regression method. A detailed presentation and analysis of this technique are underlined. As a result, the accuracy of each model will be investigated for future usage.

## **B. Statement of the problem and process parameters**

During this drilling process, seven hole quality characteristics outputs are measured: the delamination at entry, the delamination at exit, the surface roughness, the diameter error at exit, the diameter error at entry, the circularity at exit and the circularity at entry. Their definition and impact is presented below.

The **delamination** is one of the major concerns in drilling. It is considered a severe damage since it reduces the service life of the material. It is caused by the acting between the drill feed motion and the thrust force that leads to cracks between the plies in the drilled hole which may result in deterioration of its mechanical performance (durability and strength). The surface delamination is defined by the separation of the plies where the cutter enters and exits the composite materials. Therefore, two types of delamination are differentiated: the push-out at entry and the peel-up at exit. **Figure 1.1** displays a representation of the delamination factor and its mechanism.

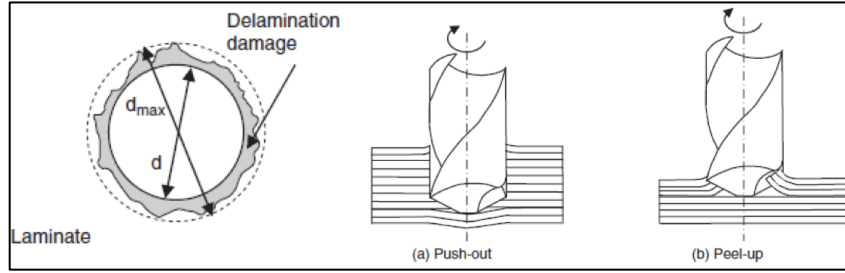


Figure 1.1: Representation of the delamination factor and mechanism [3]

The delamination is mostly affected by the feed rate, spindle speed, drill diameter, drill point design, and material configuration [3]. The delamination on the outer surface plies generally increases with the rise of the feed rate and spindle speed. Also, it is noticed that the exit delamination is highly correlated with the thrust force which is dependent on the drill point geometry [3]; the push-out delamination is reduced by lowering the thrust force. Also, it has been found that the push-out delamination is more severe than the peel-up [4]. The delamination could be measured by different practices: digital image processing [5], ultrasound [6], x-ray [7], and laser-based imaging [8]. In machinability, an improved delamination is obtained by machining at high spindle speeds and low feed rates. A delamination factor ( $F$ ) is defined as the ratio of the maximum diameter ( $d_{max}$ ) of the damage zone to the hole diameter ( $d$ ) [4]:  $F = \frac{d_{max}}{d}$ .

Another essential quality characteristic to be precisely controlled and monitored by the experts is the **surface roughness** which corresponds to the finer surface irregularities on the surface's texture. **Figure 1.2** illustrates the surface roughness as a result of the manufacturing process [8]. Surface finish in drilling composite materials have been found to be influenced by the feed rate, cutting speed, drill geometry, tool wear and tool material [9].

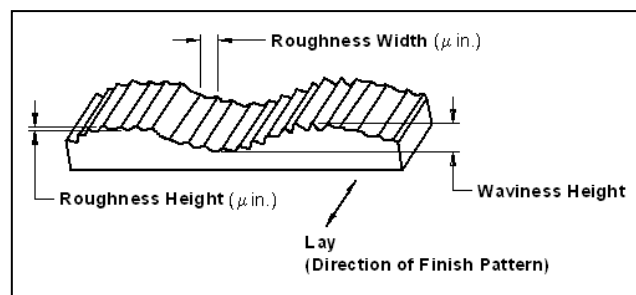


Figure 1.2: Surface Roughness illustration [10]

Another important quality attribute in drilling is the **circularity**; it is measured at entry and exit. A large circularity value is problematic for parts with relative motion because it induces vibration and heat. **Figure 1.3** represents a sketch of the circularity error. The circularity of the microhole essentially reflects the surface finish at the rim of the hole machined and is measured from the difference between two concentric boundaries: maximum and minimum height of the irregularities at the rim [11].

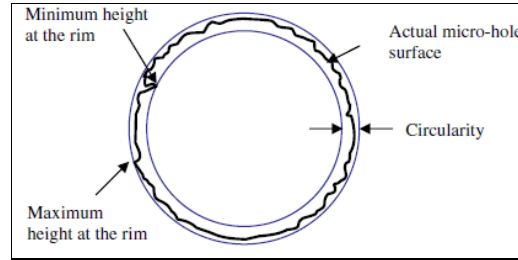


Figure 1.3: Representation of the hole circularity errors [11]

An additional output of interest for the hole quality is the **diameter error**; it is measured at entry and exit using a coordinate measuring machine. In drilling, it is important to produce accurate diameter within pre-defined tolerances. The difference between the measured diameter and the designed diameter is the diameter error. As a result, a positive error indicates over-sizing of the holes. **Figure 1.4** illustrates the diameter error.

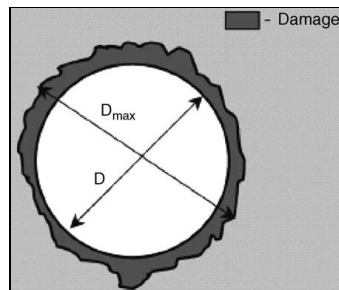


Figure 1.4: Representation of the hole diameter error [12]

During this drilling process, three uncontrollable variables (thrust force, cutting force and torque) are measured to study their behavior when varying the feed rate and the spindle speed at different levels. The definition and the main interests of these variables are presented below.

The **thrust force** is generated by the cutting action of the cutting edges and the chisel edge. For metal, using conventional twist drill, the thrust force could be correlated with the feed rate and the drill diameter by the empirical relationship below [4].

$$F_A = d^2 H_B [ K_1 \frac{f^{0.8}}{d^{1.2}} + K_2 (\frac{c}{d})^2 ]$$

Where:

*H<sub>B</sub> is the work piece Brinell hardness in kg/mm<sup>2</sup>*

*f is the feed rate in mm/rev*

*c is the chisel edge length*

*K<sub>1</sub> and K<sub>2</sub> are constants that depend on the work piece material, thickness and drill point geometry*

The **torque** is caused by the cutting force couple (F<sub>c</sub>) acting on the major cutting edges and its magnitude is defined by the magnitude of the cutting force and the drill diameter (d) [4].

Mathematically, the torque is represented by:  $M = F_c \frac{d}{2}$ .

The force and torque signals were measured using a Kistler-four component piezoelectric dynamometer model 9272. The thrust force and the torque are influenced by the feed rate, the cutting speed and the drill geometry. They both increase significantly with the rise of the feed rate due to its direct influence on uncut chip size. Also, it has been observed that the interaction combining the effect of the feed rate and the drill diameter on the thrust force and the torque is found to be more significant than the separate effect of either one of the variables. However, the cutting speed effect could have less significant effect on the thrust force and the torque [4]. The cutting force is dependent on the feed rate, and it is generally directly proportional to it [4]. In this experiment, the cutting force was derived from the torque.

Therefore, a big challenge is facing the industry of drilling manufacturing: how can they produce higher quality holes with minimum damages. To understand the effects and relations between the factors under study and the quality characteristics while drilling a composite material, several studies are conducted using the following statistical methods: multiple linear regression (MLR), analysis of variance (ANOVA) and response surface model (RSM).

### C. Description of the tool

The tool material used in the present investigation is a standard carbide twist drill (M43236), a product of Kennametal Inc. This drill is a 2- flute, right hand spiral, right-hand cut drill with a 30° helix angle and 118° point angle. The carbide grade is ISO K10 - K20 with approximately 7% cobalt as binder. This drill is shown in **Figure 1.5**.

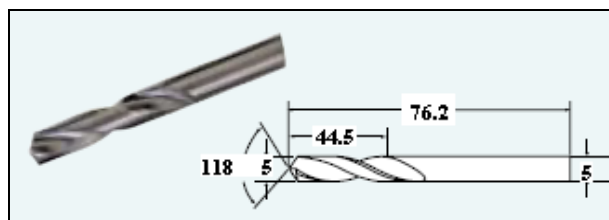


Figure 1.5: Two Flute Standard Point Solid Carbide Twist Drill of 5 mm diameter

#### D. Description of the machine

The drilling experiments were carried out on a 5-axis high-speed, high-power horizontal machining centre Makino A88 (shown in **Figure 1.6**). It has the following characteristics: 50 kW spindle power, 3 linear and 2 rotary axes, maximum spindle speed of 18,000 rpm, maximum feed rate of 50 m/min, minimum feed setting unit of 1  $\mu$ m, tool clamping force of 19.6 kN and HSK 100A spindle adapter.



Figure 1.6: Makino A88

#### E. Description of the material

The work piece comprised of woven carbon fibre as the reinforcements and epoxy as the matrix material. The woven prepreg, L-930 HT 139, used for manufacturing the laminate was supplied by J.D. Lincoln Inc. The woven prepreg was a plain weave fabricated out of T300 graphite fibers each having 3000 filaments.

#### F. Research objectives

The objective of this research is to formulate analytical models to predict the quality characteristics under study and the uncontrollable variables during the drilling of composite material using the regression method.

Therefore, the following steps will be used:

- a) Use a full factorial design with two factors at appropriate levels of the independent variables.
- b) Analyze the data distribution of each output to understand the existing relationship with the input variables.
- c) Develop a mathematical model of each output using the multiple linear regression method.
- d) Add an interaction effect of the independent variables to the model and refit the data to a new multiple linear regression model.
- e) If an accurate model couldn't be found through the multiple regression methodology, attempt a new fitting with the nonlinear regression technique or different types of transformation.
- f) For each developed model, check the model adequacy by demonstrating how well the model fits the observed data and how well the regression model predicts new observations.

## CHAPTER 2 LITERATURE REVIEW

In the literature, several experiments were carried out on drilling of CFRP composites. S. Jayabal & U (2010) used a full factorial design with three factors at three levels each ( $3^3$  equal 27 experimental runs) to evaluate the mechanical and machinability characteristics of hybrid composites [13]. The measured responses were the thrust force, the torque and the tool wear and the process control factors were the bit-drill diameter, the spindle speed and the feed rate. The multiple regression technique was used to identify the mathematical model of the interaction of the specified main effects in the drilling process. The machinability study determined that the feed rate has the most significant role on the machining characteristics. To conclude their experiment, the authors defined the best combination of the drilling parameters to minimize the effects of the thrust force, the torque and the tool wear. Using this design, the average absolute errors for thrust force (2.56%), torque (2.91%) and tool wear (2.93%) are between 2.5% and 3%. The variability of the model was 92.96% for thrust force, 89.32% for the torque and 96.72% for the tool wear which is considered very good estimates.

In another paper, S. Jayabal & U. Natarajan (2011) carried out a statistical modeling to develop mathematical models to relate few outputs (thrust force, torque and tool wear) to three inputs (drill-bit diameter, feed rate and spindle speed) through the multiple linear regression technique [14]. The ANOVA was performed to test the significance of the obtained coefficients at one per cent level of significance. The developed models were verified by eight sets of experiments. They used a Box-Behnken Design (BBD) with three factors at three levels each to study the effects of each factor and their interactions on predefined hole's characteristics with a total of 17 experiments. To analyze the collected data, a quadratic design was chosen for the thrust force and the torque models. However, linear terms were used to define the model of the tool wear. The RSM was used to define the optimal responses (thrust force, torque and tool wear) for each input (drill bit diameter, spindle speed and feed rate). In that study, the feed rate and the drill bit diameter were found to be the most significant factors affecting the thrust force. To confirm the accuracy of the obtained results, the percentage of the error between the observed and predicted values was calculated using the following equation:

$$\% \text{ of error} = \frac{\text{predicted value} - \text{experimental value}}{\text{predicted value}}$$

Using the developed mathematical model, the percentage of error calculated was within the acceptable ranges which confirm the adequacy of the results. In fact, the average absolute errors for the thrust force (0.4%), the torque (0.08%) and the tool wear (0.57%) were less than 1%. The variability of the model was 99.88% for the thrust force, 99.88% for the torque and 96.22% for the tool wear.

The difference between the previous two papers presented above is that the authors chose to use a full experiment design in the first versus a BBD in the second. In general, every machine used in a production process allows its operators to adjust various settings affecting the quality characteristics of the manufactured product. Experimentation and testing allow the manufacturing engineer to learn which factors have the highest impact on the resultant quality characteristics by adjusting the settings of the machine in a methodical manner. Using this information, the settings can be regularly enhanced until optimum features are achieved. Moreover, it is important to know what to change in order to produce a better product at minimum cost. At first, experimenters considered three factors (bit-drill diameter, spindle speed and feed rate) affecting the production process at three levels to determine whether any of these changes would affect the outputs under study (thrust force, torque and tool wear). The most instinctive approach to study those factors would be to vary the three factors (thrust force, torque and tool wear) in a full factorial design; to try all possible combinations of settings. This method is acceptable when the number of factors under study and the settings are small. In fact, the number of necessary observations (runs in the experiment) increases when more factors and settings are involved. For example, to study six factors, the necessary number of runs in the experiment would be  $2^6 = 64$  observations and for 10 factors, it's  $2^{10} = 1024$  observations. Every observation involves time and cost to set and reset the machine. In a production environment, it's usually not feasible to run a high volume of a set of production for the experiment. In these conditions, fractional factorials (reduced number of observations) are used to "sacrifice" interaction effects so that the main effects may still be computed correctly. By running fewer experiments, the same results and conclusions may be pulled to determine the optimum values of the process and to analyze the behavior of each factor. In the second paper, the authors selected an alternate design of experiment to reveal significant interactions and find the optimum operating conditions for a high-quality drilling: the BBD. This design is specially made for factors with three levels coded as (-1, 0, +1). BBD is an independent

quadratic design and not a fractional factorial design. Consequently, with the same 3 factors (bit-drill diameter, spindle speed and feed rate), 17 runs are completed:

- Four points in the center of each face which makes a total of 12 points (refer to **figure 2.1** for a basic illustration), and
- Five replicates of the center points.

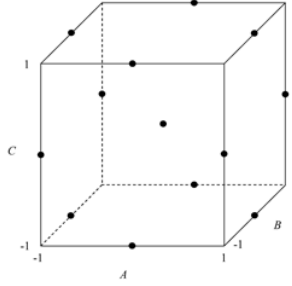


Figure 2.1: The basic geometric representation [15] for a BBD for three factors

The use of +1 and -1 for the variable settings is called “coding the data”. In this case, the authors used three coded levels: -1 for the higher value, 0 for the center point and +1 for the lower value.

This specific methodology was chosen to fit a second-order response surfaces for the three factors under study. In both papers, the results are showing that the most important factor affecting the thrust force, torque and tool wear is the feed rate. In fact, the BBD methodology compared to the full factorial experimental design saved the researchers 10 experiments; a significant difference in time, material consumption and cost. The lesson to learn is that BBD can efficiently be applied for modeling drilling factors (bit-drill diameter, spindle speed and feed rate) in an economical way of obtaining the information with least number of runs. Refer to **annex B note 5** for the experimental design matrix in terms of coded factor levels for a BBD for three factors.

Coding is a simple linear transformation of the original measurement scale. In the real scale, the highest value is  $X_h$  and lowest is  $X_L$ . The scaling transformation takes any original  $X$  value and converts it to:

$$TR(X) = \frac{(X - A)}{B}$$

Where:

$$A = \frac{X_h + X_L}{2}; B = \frac{X_h - X_L}{2}$$

- If  $X = X_h$  then,  $TR(X_h) = (X_h - X_L) / (X_h - X_L) = 1$ .
- If  $X = X_L$  then,  $TR(X_L) = - (X_h - X_L) / (X_h - X_L) = -1$ .
- If  $X = \text{average of the settings} = A$  then,  $TR(X_{avg}) = [(X_h + X_L) - (X_h + X_L)] / (X_h - X_L) = 0$ .

To transform back a coded value to its original measurement scale, multiply the coded value by  $B$  then, add  $A$ :  $X = (\text{coded value}) * B + A$ . As an example, if the variable  $X$  is pressure and the high setting is 100 psi and the low setting is 20 psi then,  $A = (100 + 20)/2 = 60$  psi and  $B = (100 - 20)/2 = 40$  psi. The real value of the coded value of the center point 0 has a temperature of:  $X = 0*(40) + 60 = 60$  psi.

Before presenting further studies, the design of experiments (DOE) concept is explained. The DOE screens a large number of factors with a minimum sample size. From a cost-effective and time-reduction point of view, engineers and physicists can no longer afford to experiment in a trial-and-error manner testing by changing one factor at a time. Consequently, a more effective method known by design of experiments, has been conceived. It's an efficient systematic approach and powerful tool based on a stochastic search technique for solving optimization problems, which has been widely applied in many scientific and engineering fields for process efficiency and product quality. The type of design is highly dependent on the number of factors studied. This method:

- Considers all factors simultaneously;
- Provides an effective way to solve serious operational and production problems; and
- Reveals information about the interaction of factors and the way the whole system works (even the interaction between the factors), a fact not obtainable through testing one factor at a time.

The input factors are independent variables that affect the responses and outputs under study. Each factor has a set of settings defined by the experimenter. This approach tackles the drilling quality problems by performing the minimum number of experiments needed. Then, develop analytical equations that express all the important and significant factors which can be used, depending on the drilling conditions to predict the desired outputs. Therefore, the constructed model is able to identify the significant factors affecting one or multiple outputs, achieve an optimal process output (combination for best quality) and reduce variability. Refer to **annex A** to understand in details how to build a DOE.

In this research, the experts completed 35 runs which are considered a full factorial plan that combines all combinations with no replications. Other plans can be built to provide more adequate results. Here are two recommended designs that can be used:

- With the same number of runs:  $2^2+1$  with 7 replicates equal to 35 runs. In this case, two levels (minimum and maximum) are defined for each input variables: feed rate (20, 800) and spindle speed (1500, 15000). The “+1” represents the center point (200, 8500). In this design, the observations in **table 2.1** are replicated 7 times.

Table 2.1: Experiment setup with 35 runs

Run #	Feed rate	Spindle speed
1	20	1500
2	20	15000
3	800	1500
4	800	15000
5	200	8500

- With less number of runs:  $2^3$  with two replications equal to 18 runs. In this case, three levels (minimum, center and maximum) are defined for each input variables: feed rate (20, 200 and 800) and spindle speed (1500, 8500 and 15000). This design represents half of the current experiments done and will provide good results. In this design, the observations in **table 2.2** are replicated twice.

Table 2.2: Experiment setup with 18 runs

Run #	Feed rate	Spindle speed
1	20	1500
2	20	8500
3	20	15000
4	200	1500
5	200	8500
6	200	15000
7	800	1500
8	800	8500
9	800	15000

Replications are added in both plans to determine the experimental error  $\epsilon_i$ .

In another study, the authors [16] performed an analysis of the thrust force in drilling of glass fiber-reinforced plastic. During this investigation, the spindle speed, the feed rate, and the drill diameter were considered as machining input parameters at three levels each. The authors

selected “Tagushi’s  $L_{27}$ ” experimental design with three repetitions (81 runs in total) to examine the relations between the inputs (spindle speed, feed rate, and drill diameter) and the output (thrust force). For a first estimation, Pareto ANOVA (a graphical method to understand the overall relationships but not very exact whereas no error terms are considered) was employed to determine the significant factors and interactions. For further analysis, the ANOVA was exploited at 95% confidence level to understand the effects of every factor. Using the RSM, a mathematical model was deduced and the correlation verified between the spindle speed, feed rate, and drill diameter, and the thrust force. As per the authors, the obtained results are only near optimal.

As reported in many papers, to analyze the process and to find the optimum response, one may turn to the RSM methodology. The RSM is the collection of mathematical and statistical techniques that are useful for the modeling and analysis of problems in which a response of interest is influenced by several variables and the objective is to optimize that response [18]. RSM also quantifies the relationship between the controllable input parameters and the obtained response surfaces [19]. This technique helps experts to define the best settings combination for the factors under study to provide the most appropriate values for the desired responses.

An alternative approach [17] involving the design of experiments has been used to select the optimal cutting parameters of carbon fiber reinforced thermosets. The input variables under study were cutting speed, feed rate and tool point angle at five levels each. The outputs to be optimized for enhanced hole quality characteristics are the thrust force, the delamination, the damage width and the surface roughness. The methodology used in that study combines tagushi’s technique, RSM and ANOVA.

Another related research involves the use of the regression technique is presented. In this paper [20], the authors employed regression and artificial neural network (ANN) to predict tool wear in end milling. The conducted experiments to measure tool wear are based on the DOE of five levels of four factors full factorial technique. The input variables under study are the feed rate, the helix angle, the spindle speed and the depth of cut against the tool wear being the only output. The ANN techniques provided a higher accuracy rate (average error < 2%) models compared to the results obtained from the classic statistical regression model (average error < 5%). As a first step, the authors normalized all the raw data of input values to develop an appropriate model via the regression model. Then, they used the data generated from the previous model to demonstrate a

new one through the ANN method to predict the minimum value of tool wear and estimate the best combination values of the process.

During this exploration, some researchers used the ANN and the support vector regression (SVR) methods to predict the drilling and cutting output values for better sensitivity and specificity; statistical measures of the performance of the diagnostic. The sensitivity is the capability to correctly identify the output with the identified problem; however test specificity is to incorrectly identify those with the known problem. There are two types of errors intrinsic in every technique [21]:

- Sensitivity = error Type I (FP - False Positive): error of deciding when an action should be taken while the system is in a normal condition and,
- Specificity = error Type II (FN - False Negative): error of deciding that the system is in a normal condition while it is not.

There are four possible outcomes of the simulation:

- True negative (a) = correctly rejected
- False negative (b) = incorrectly rejected
- False positive (c) = incorrectly identified
- True positive (d) = correctly identified

From a mathematical point of view, the sensitivity and specificity are defined by:

- Sensitivity = Probability of FP (p1):  $p1 = \frac{c}{a+c}$
- Specificity = Probability of FN (p2):  $p2 = \frac{b}{b+d}$

Then, the probability of correct decision will be defined by (Pc):  $Pc = \frac{a+d}{a+b+c+d}$ .

The ANN technique is usually used when the researchers want to search for certain patterns in the data or if they face complex relationships between inputs and outputs. Artificial neural networks implement the empirical risk minimization principle to minimize the error on the training data, while SVR adheres to the structural risk minimization (SRM) principle seeking to set up an upper bound of the generalization error (Vapnik et al., 1996). In fact, Jixin Li performed some experiments to understand the performance of both methods and he concluded that both have similar performance in binary classification, but support vector machines (SVM) outperformed ANN in multi-class classification [22]. Refer to **annex B note 3** for the ANN process description.

Another author [23] developed empirical models to predict surface roughness and tool wear in the cutting process. He used three different techniques: RSM, ANN and SVR. The process parameters were the cutting speed, the feed rate and the cutting time. The RSM method was used to estimate the response value based on a full quadratic regression model. Then, the author relied on the ANOVA to justify the goodness of fit for the developed RSM models. Afterwards, the ANN and SVR were applied for the same purpose. After comparing and evaluating the results given by the models, it has been found that ANN and SVR models are much better than RSM models for predicting surface roughness, tool wear and power.

Some of the quality characteristics of the drilling process under study may not be modeled by the multiple linear regression technique. In this case, a different approach is applied: the nonlinear regression. The fundamental idea of nonlinear regression is the same as the linear regression. It is characterized by the fact that the prediction equation depends nonlinearly on one or more unknown variables. Whereas linear regression is often used for building a purely empirical model, nonlinear regression usually arises when there are physical reasons for believing that the relationship between the response and the predictors follows a particular functional form. In the general normal nonlinear regression model, the function relating the response to the predictors is not necessarily linear:

$$y_i = f(\beta, X_i) + \epsilon_i$$

Where:

*X<sub>i</sub> is a vector of predictors for the *i*th of *n* observations*

*i = 1, 2, ..., n*

*β is the vector of regression parameters to be estimated*

*ε<sub>i</sub> is the random error*

The use of nonlinear regression is seen in many applied sciences, ranging from biology to engineering to medicine to agriculture. From the examined articles, few articles are found to use the nonlinear regression method to extract the prediction model. In one of the studies, the authors [24] explored two techniques which are the multiple regression analysis and the artificial neural networks to study the influence of the cutting speed, feed, and volume fraction of the reinforcement particles (inputs) on the thrust force and torque (outputs) in the drilling process of self-lubricated hybrid composite materials. They also compared both prediction models to examine the prediction accuracy of each one. The results showed that the linear regression model

works well on the thrust force  $R^2_{adj} = 92.2\%$ . The  $R^2_{adj}$  is the adjusted coefficient of determination indicating how well the data values fit the developed statistical model; a detailed explanation is presented in **chapter 3**. The closer the  $R^2_{adj}$  is to 100%, the better the model is. However, the prediction model for the torque was determined to be ineffective with  $R^2_{adj} = 76.6\%$ . To improve on the variability, they transformed the outputs (thrust force and torque) into a logarithmic scale (ln) to create a linear relationship between the input and the output. In fact, this action was very successful because higher values of  $R^2_{adj}$  were obtained: 95.46 % for ln(thrust force) and 92.65% for ln(torque). The challenge of the nonlinear regression technique is to find out what should be used as the basis of transformation. Usually, this is determined by the examination of scatter plots or from the understanding of the underlying process itself. If a transformation is not possible, the use of different nonlinear functions can be attempted. The common models used [25] are:

- Least squares estimates;
- The Box–Cox Method [18].
- Logarithmic regression:  $f(x, \theta) = \theta_1 \exp(-\theta_2 x)$

Where:

*x is the predictor variable*

*$\theta$  is the parameter vector (estimated from the data by minimizing a suitable goodness-of-fit expression with respect to  $\theta$ )*

A transformation is used for three purposes [18]:

- Stabilize response variance,
- Make the distribution of the response variable closer to the normal distribution, and
- Improve the fit of the model to the data.

Very few researchers put together the controllable and uncontrollable variables in the same model in metal drilling and cutting processes. One of the known methods is the analysis of covariance (ANCOVA) which is used to evaluate the collected data and to predict an accurate mathematical model of outputs. This technique is usually selected to improve the precision of an experiment and to evaluate how strong the relationship is between two variables and there are five assumptions that underlie the use of the ANCOVA model [26]. One of these assumptions is that the variables must not be correlated.

In fact, this methodology has not been applied in modeling of the drilling process because one of the main issues is that the uncontrollable (thrust force, cutting force and torque) and the controllable variables (feed rate and spindle speed) are highly correlated among each other. In fact, is considered a major violation of one of the assumptions. **Table 2.3** indicates that all variables, but spindle speed are strongly correlated. Consequently, combining those variables (controllable and uncontrollable) is not recommended to determine an appropriate covariance model. In presence of correlation between the variables, it is difficult to distinguish if the variation on the dependent is due to one or the other predictor. The uncontrollable variables (thrust force, cutting force and torque) are extremely dependent on the feed rate. Also, the uncontrollable variables are correlated among each other.

Table 2.3: Correlation table between all variables for ANCOVA

	feed	spindle~d	thrust~e	cuttin~e	torque
feed	1.0000				
spindlespeed	0.0000	1.0000			
thrustforce	0.9330	-0.2731	1.0000		
cuttingforce	0.9485	-0.2591	0.9840	1.0000	
torque	0.9485	-0.2591	0.9840	1.0000	1.0000

In this literature survey, the MLR was only used to investigate few drilling characteristic outputs: surface roughness or delamination. In this research, this method will be applied over the seven different characteristic outputs described previously. The next section will present the regression analysis methodology of all the desired outputs. Also, at the end of **chapter 3**, the thrust force and the cutting force will be analysed as inputs to predict the delamination at entry and exit.

## CHAPTER 3 REGRESSION ANALYSIS TECHNIQUE WITH CONTROLLABLE VARIABLES AS INPUTS

The regression analysis was developed by Sir Francis Galton in the latter part of the 19<sup>th</sup> century and to this day, the term regression describes the existing statistical relations between variables. This relationship is defined by a regression model and may differ in the form of the regression function (linear, curvilinear, etc.). It may contain more than one predictor variable. Here are the steps to construct a regression model [27]:

- Select the major inputs that contribute the most to reduce the variation on the output,
- Attempt the approximation of a complex regression function by a linear regression function, and
- Restrict the coverage of the model to region of values of the predictor variables.

A basic regression model is known to be “simple linear” and this model is written as [27]:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

Where:

*$Y_i$  is the value of the response variable in the  $i$ th trial*

*$\beta_0$  and  $\beta_1$  are parameters or regression coefficients*

*$X_i$  is a known constant, the value of the predictor variable in the  $i$ th trial*

*$\varepsilon_i$  is a random error term with mean  $E\{\varepsilon_i\}=0$  and variance  $\sigma^2 \{\varepsilon_i\} = \sigma^2$ ;  $\varepsilon_i$  and  $\varepsilon_j$  are uncorrelated so that their covariance is zero*

*$i = 1, \dots, n$*

The difference between simple and linear terminologies is defined in **annex B note 1**. Experts can be facing a model which is not simple and has more than one predictor variable. In this case, the multiple linear regression is used. Therefore, suppose there are two predictor variables, and the relationship is linear between these variables and the dependent, the model will be [27]:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i$$

*(Multiple Linear Regression model with two independent variables)*

Where:

*$X_{i2}$  is the value of the second predictor variables in the  $i$ th trial*

For some cases, the data won't be fitted by a linear simple regression. In this case, another type of fitting may be used known as "Transformed Variables". It may be represented for example by the following model [27]:

$$Y_i = \log(Y_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \varepsilon_i$$

(Transformed Variables Regression model with three independent variables)

Where:

$Y_i$  is the transformed response

In this model, the response is transformed to a logarithmic function but it can be any other mathematical function.

To better understand the variability of the response when analyzing the variables, an **"Interaction effect"** is involved. It represents the combined effect of the independent variables. When an interaction is present, the impact of one variable depends on the level and intensity of the other variable [27]. The model corresponds to:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1} X_{i2} + \varepsilon_i$$

(Interaction effect with two independent variables)

If the data cannot be fitted with the linear or the transformed regression model, the nonlinear regression method is developed. An example of this technique would be:

$$Y_i = \beta_0 \exp(\beta_1 X_i) + \varepsilon_i$$

(Non-linear regression with one independent variable)

The regression coefficients interpretation depends on the:

- Magnitude of the coefficient for each independent variable which provides the size of the effect that the variable have on the dependent variable, and
- Sign of the coefficient (negative or positive) withdraws the trend of the effect.

The coefficient helps experts understand how much the dependent variable is expected to increase/decrease when that independent variable increases/decreases by one, holding all the other independent variables constant.

When conducting any statistical analysis, it is important to evaluate how well the model fits the data and that the data meet the assumptions of the model. To determine if a linear regression model is appropriate, the following assumptions shall be verified:

- Independency: the response variables  $Y_i$  are independent,

- Linearity-Outliers: the relationship between Y and X is a straight line,
- Normality:  $Y_i$  are normally distributed, and
- Homoscedasticity-variability:  $Y_i$  have all the same variance  $\sigma^2$ .

The assumptions on the residuals are the same as the response variables [18]. As a result, if the assumptions on the random errors are satisfied thereby, the assumptions on the response variable are validated. Residuals ( $e$ ) are defined to be the random errors, the differences between the observed values ( $Y_i$ ) and the values fitted ( $\hat{Y}_i$ ) by the model [27]. The residual is denoted by:

$$e_i = Y_i - \hat{Y}_i \text{ where } i = 1, \dots, n$$

**Independence of the response variables** is subject to the design of the study and the way the data have been collected. The errors associated to different observations must be statistically independent from each other including the predicted response. The residuals in time order of data collection is plotted to detect any substantial correlation [18]; a tendency to have runs of positive and negative residuals indicates positive correlation. Ideally, most of the residual should fall within the 95% confidence bands around zero, which are located at roughly  $\pm 2/\sqrt{n}$  where  $n$  is the sample size. If there is significant **correlation**, it indicates a fundamental structural problem in the model and can interfere with the analysis of variance. If the plot does not reveal any pattern, the independence assumption is satisfied. A common pattern is that the residuals increase as the fitted values increase.

To investigate the **normality structure of the distribution**, a normal probability plot of the residuals is plotted [18]. The plot shows the residuals against a theoretical normal distribution. The normality is confirmed if the values lay on a straight line. When examining this plot, emphasis will be on the fitting of the central values than on the extremes. Therefore, retreats from this straight line indicate retreats from the normality distribution. Lack of normality assumption may provide misleading interpretations of the results.

To verify the **linearity** (random errors have zero means) and **homoscedasticity** (constant variation), a scatterplot of the standardized residuals versus the fitted values is plotted. The standardized residuals are the deviations of the observations away from the predicted values. The variability in the residuals should not depend in any way on the predicted value. If both assumptions are satisfied, the residuals are expected to vary randomly around zero and the spread

of the residuals to be about the same throughout the plot. Any observed pattern in the residual plot suggests a violation of one or more assumptions or it's an indication that there may be a need for data transformation. In this study, for a better understanding of the data behavior, the scatterplots of the output against the independent variable under study is plotted to verify the linearity assumption. Also, in some cases when verifying the homoscedasticity, the plot of the standardized residuals versus the fitted values may not be convincing; subsequently, the Breusch-Pagan test may be used to confirm the constancy of error variance [18].

In each output study, the regression model is presented then, the assumptions are verified to conclude if the model is valid for future use or not.

Assume that two independent variables are involved in the process: the feed rate and the spindle speed, and the output variable is the thrust force. The global F-test in the ANOVA table tests the null hypothesis:

$$\begin{aligned} &H_0: \beta_1 = \beta_2 = 0 \\ &\text{vs. the alternative hypothesis} \\ &H_1: \beta_1 \neq 0 \text{ or } \beta_2 \neq 0 \end{aligned}$$

If  $H_0$  is accepted, it indicates that the feed rate and spindle speed are non-significant variables on the thrust force. However, if  $H_0$  is rejected, it means that at least one of the variables has an impact on the behavior of the thrust force. However, it doesn't necessarily suggest that the found model is the best. For more accurate models, researches can attempt to add more independent variables or higher order terms. On the other hand, experts must be careful not to over-fit the model to the sample points. Errors can occur in every statistical process.  $\alpha$  is the Type I error that occurs when that the process is declared to be out of control when in fact it is not,  $H_0$  is rejected when in fact it is true. In a production setting, this sort of error is extremely costly. That's the reason why  $\alpha$  is defined to be the smallest possible. However,  $H_0$  should not be retained if there is no enough evidence; in this case, the process is out of control and consequently, some results are missed. **Table 3.1** explains the relationship between the null hypothesis and the type of errors.

Table 3.1: The relationship between the null hypothesis and the type of statistical errors

Statistical Decision	True State of $H_0$	
	$H_0$ is true	$H_0$ is false
Reject $H_0$	Type I error	Correct
Do not reject $H_0$	Correct	Type II error

In this research, the main testing will be on the significance of the sample coefficients of the predictor variables. There is little relevance in testing the significance of the estimate of the y-intercept. Thus, it doesn't provide anything about the usefulness of the model.

Analysis of variance known by ANOVA or Fisher's ANOVA was developed in the 1920s by R.A. Fisher (1890-1962). ANOVA helps determine whether there are any differences among the population output means when the input values are changed, and specifically which pairs of means are significantly different ( $H_0: \mu_1 = \mu_2 = \dots = \mu_a$  /  $H_1: \mu_1 \neq \mu_2 \neq \dots \neq \mu_a$ ). This technique performs a series of F-tests (designed to test if two population variances are equal by comparing the ratio of two variances), one for each pair of means. In an F-test, an F-statistic is computed and compared to a critical value. Alternatively, the p-value approach could be used for decision-making [26]. The p-value is a probability ranging between zero and one. It represents the confidence that the null hypothesis is right by proving that all of the alternatives are wrong. The lower the p-value is the more significant the results are. If p-value for the F-statistic is less than the predefined level of significance (0.01, 0.05 or 0.1),  $H_0$  is rejected.

It has become a scientific standard to say that p-values exceeding 0.05 aren't strong enough to be the sole evidence that factors being studied differ in their effect. When the p-value is lower than this conventional reference point, it indicates that the factor is significant and has an imperative influence on the output results.

After each ANOVA analysis, the  $R^2$  known as the coefficient of multiple determination is studied to assess the usefulness of the regression model developed. Many authors discuss the importance and how to calculate this value. The  $R^2$  is defined as the proportion of variation in the response that is explained by the regression model [28]. The value of  $R^2$  varies between 0 and 1. An  $R^2$  that is near 0 indicates that the regression model does not fit the data well and there is no vital relationship between the data. While an  $R^2$  closer to 1 indicates that the regression line fits the data very well and most of the variability is included in the model to predict the outcome accurately.

The formula of the coefficient of determination is:

$$R^2 = \frac{\text{Total Variation} - \text{Unexplained Variation}}{\text{Total Variation}}$$

In Mathematical term:

$$R^2 = \frac{\Sigma(Y - \bar{Y})^2 - \Sigma(Y - \hat{Y})^2}{\Sigma(Y - \bar{Y})^2}$$

Where:

$\bar{Y}$  is the calculated mean value of the variable

$\hat{Y}$  is the value predicted of the variable by the regression model

To balance the effect that the number of independent variables has on the coefficient of the multiple determinations, the adjusted  $R^2$  is investigated in our analysis, instead of the  $R^2$ . It is calculated by:

$$R_{adj}^2 = 1 - \frac{n - 1}{n - k - 1} (1 - R^2)$$

Where:

$n$  is the sample size

$k$  is the number of independent variables

In this study, a mathematical model is developed based on the mechanics and dynamics of the drilling process to predict the influence of the feed rate and the spindle speed on the measured holes quality characteristics: delamination, surface roughness, circularity and diameter error. Using these models and with set values of the feed rate and the spindle speed, the trend of each output can be predicted. However, with the regression technique, the effect of the uncontrollable variables cannot be combined to the independent variables due to high correlation (refer to **table 2.3**). Thus, to better understand the interactions between all variables, two steps are required:

1. Consider the uncontrollable variables which are measured during the experiments as the outputs and construct a prediction model using the controllable factors as inputs; and
2. Build up a regression model of each of the seven characteristic outputs (e.g. delamination at entry) and the controllable inputs (feed rate and spindle speed).

The drilling process involves two factors which are independent and controllable (refer to **table 3.2**). The levels are the settings of each factor in the study chosen by the implicated experts.

Table 3.2: Controllable factors involved in the drilling process

	Factors	Units	Levels
<b>A</b>	Feed (F)	Micron/rev	20 – 60 – 100 – 200 – 400 – 600 – 800 (7 levels)
<b>B</b>	Spindle Speed	Rpm	1500 – 5000 – 8500 – 12000 – 15000 (5 levels)

In this chapter, the following variables are not controlled during this process, but they are measured within the experiment to determine how they would affect the final product's quality:

1. Thrust force (Tf)
2. Cutting force (Cf)
3. Torque (T)

These variables are analyzed as outputs to study their variability when the feed rate and spindle speed vary within the specified levels in **table 3.2**.

The quality outputs of the experiments are measurable outcomes potentially affected by the independent factors at their respective levels. The dependents described below are analyzed to determine the significant factors and their modelling fit:

1. Hole Delamination at entry (Den)
2. Hole Delamination at exit (Dex)
3. Hole Surface Roughness (SR)
4. Hole Diameter Error at Exit (Drx)
5. Hole Diameter Error at Entry (Dre)
6. Hole Circularity at Exit (Cex)
7. Hole Circularity at Entry (Cen)

A series of experiments are conducted using a full factorial plan that includes all the combinations of the controllable factors levels: 35 runs (7 levels feed rate x 5 levels spindle speed). **Table C.1 in annex C** lists the data collected for each factor and measured outputs.

Because the regression analysis entails a lot of calculations, the data analysis is conducted using the statistical program **Stata** (Version 13, StataCorp, College Station, Texas) to perform the parameter estimates, ANOVA and coefficient of determination results, and all graphical illustrations using the experimental data in **table A1.1 in annex A**.

In this chapter, the analysis of the ten variables mentioned previously is presented using the simple linear and multiple linear regression applications. To better understand the distribution of the output data compared to the independent variables, the box-and-whisker is plotted, also known by the boxplot. This procedure illustrates the symmetry, tail sizes, and median value of the sample as well as indicating the possible existence of outliers and inhomogeneous data [29]. Also, this scheme has lines extending vertically from the boxes which indicate the variability

outside the upper and lower quartiles. The line from the end of the box to the largest and smallest values that are not outliers is called whiskers. If the upper whisker is much longer than the lower whisker, it gives the impression of positive skewness. Refer to **annex B note 2** for a visual representation of the boxplot. After analyzing the boxplot, three different tables below are presented to investigate the effects of the independent variables and the model's variability:

1. **ANOVA** table consists of:

a) SS: three types of sum square are important to the regression model:

- i.  $SS_{\text{model}}$  is the sum of squares due to treatments
- ii.  $SS_{\text{Error}}$  is the sum of squares due to error (the squared differences between the actual value  $y$  and the prediction of  $y$ )
- iii.  $SS_{\text{total}} = SS_{\text{model}} + SS_{\text{Error}}$  [26]

b) Degrees of Freedom (df) [26]:

- i.  $df_{\text{Intercept}} = 1$
- ii.  $df_{\text{model}} = v1 = a - 1$  (where  $a$  refers to the number of treatments)
- iii.  $df_{\text{error}} = v2 = N - a$ , (where  $N$  is the total number of experiments).
- iv. Total =  $N - 1$

In this case, there are 35 observations thus,  $N = 35$  and  $a = 7$  (7 levels of the feed rate used); the degrees of freedom will be:

- $df_{\text{Intercept}} = 1$
- $df_{\text{model}} = 6$
- $df_{\text{error}} = 35 - 7 = 28$
- Total =  $35 - 1 = 34$

c) MS: is the mean square and there is two mean squares that are important:

- i.  $MS_{\text{model}}$ : is the explained variability and it's the mean square of the model equal:  

$$MS_{\text{model}} = (\text{Sum of the square})_{\text{model}} / df_{\text{model}}$$
- ii.  $MS_{\text{error}}$ : is the unexplained variability and it's the mean square of the error:  $MS_{\text{error}}$   

$$= (\text{Sum of the square})_{\text{error}} / df_{\text{error}}$$

The significance tests associated with the ANOVA are based on F-statistic. It tests the null hypothesis that all the regression coefficients are equal to zero; the alternative hypothesis is that at least one of the coefficients is non-zero. If  $F$  is large, then the null hypothesis can be rejected.

2. **Parameter estimates table** includes:

- a) The variable Parameter: they are the coefficients of the model used to predict the response.
- b) The standard error: is the estimate of the variability of the coefficient; a measure of the precision with which the regression coefficient is measured. The smaller the standard error, the more representative the regression coefficient will be.
- c) The variable t-value: it's the variable "Parameter" divided by its standard error.
- d) The variable p-value: test if our parameter is significantly different from 0.
- e) Conf. Lmt (-95%) and (+95%): it's the confidence interval of the regression coefficients values. Approximately 95% of the data are expected to fall within two standard deviations of the regression line. For example, if a  $\beta$  coefficient is equal to zero, it means that this variable has no influence on the dependent variable. In this case, if the 95% CL contains the zero, the p-value of this variable is expected to be greater than 0.05. Also, the other important information to extract from the 95%CL is the precision. The 95%CL is a function of the standard deviation (measure of how the numbers are spread out). Thus, a wider 95%CL indicates a poor precision.

3. **Coefficient of determination table** contains:

- a) Number of observations
- b) F-value: is the ratio of the mean regression sum of squares divided by the mean error sum of squares ( $MS_{\text{model}}/MS_{\text{error}}$ ). Its value will range from zero to an arbitrarily large number. The larger the F is, the more useful the model.
- c) P-value: is  $\Pr(F > F_{1, n-2})$ .
- d)  $R^2$
- e) Adjusted  $R^2$
- f) Root MSE

After each ANOVA, a fitted model is constructed and a plot of the estimated regression model is generated to identify any abnormal fit or outliers. On the graphical illustration, the confidence interval is drawn. The confidence band enables the experts to see the region in which the entire regression line lies to determine the appropriateness of the fitted function (how well the data define the best-fit curve) [27]. The distance between the observed values and the estimated regression equation is the error (residual) of each observation. While analyzing the data, multiple

well-fitted models for the same output under study may be found. In this case, the tool to be used to compare the models is called Akaike Information Criterion (AIC).

This method helps selecting the more appropriate model from a set of models to predict the chosen output. The AIC approach yields consistent results and is independent of the order in which the models are computed [30]. The smallest the AIC is, the better the model is. Mathematically, it is defined by:

$$AIC = n * \log((RSS/n)^2) + 2 * K$$

Where:

*n is the number of observations*

*K is the number of parameters in the model +1*

*RSS is the residual sum of squares*

All the following analysis will be developed using the same steps to be consistent with the conclusions. The dependent variable experimental data will be fitted through:

1. A first order regression analysis against the feed rate,
2. A first order regression analysis against the spindle speed,
3. A multiple linear regression model without and with an interaction effect, and
4. The nonlinear regression or transformation, if applicable.

### **3.1 Thrust force analysis**

#### **3.1.1 Thrust force distribution study over feed rate**

In this section, the feed rate is analyzed as the only effect on the thrust force. This process is the simple linear regression because only one independent variable is involved. As a start, the parameter estimates, ANOVA and coefficient of determination results are generated in **table 3.3** to understand the effect of the feed rate on the thrust force. This table indicates that the calculated p-value is near 0, lower than 0.05 which implies that the feed rate has a significant effect on the thrust force. The adjusted  $R^2$  corresponds to 0.6792 which indicates that 67.93% of the variability is explained only by the feed rate. Therefore, the feed rate has a major influence on the thrust force.

Table 3.4: Parameter estimates and ANOVA results of thrust force against feed rate

Source	SS	df	MS	Number of obs = 35		
Model	373941.645	1	373941.645	F( 1, 33) =	72.98	
Residual	169080.126	33	5123.64019	Prob > F =	0.0000	
				R-squared =	0.6886	
				Adj R-squared =	0.6792	
Total	543021.771	34	15971.2286	Root MSE =	71.58	

thrustforce	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.3738568	.0437615	8.54	0.000	.2848232	.4628903
_cons	28.65518	18.22439	1.57	0.125	-8.422616	65.73298

The boxplot in **figure 3.1** demonstrates that the thrust force is higher as the feed rate increases. Thus, the feed rate is directly proportional to the thrust force. The thrust force reaches its maximum value when the feed rate is at its maximum as well. An obvious difference is seen between the boxes at 20 to 400 and 600 to 800:

- For feed rate from levels 20 to 400, the boxes are comparatively short. This suggests that the thrust force observations are contained within a very small segment of the sample.
- At level feed rate 600 and 800, the boxes are tall which indicates that the thrust force varies much more within these samples.

Also, at feed rate equal to 100, an outlier outside the upper boundary is observed which may affect the fitted model.

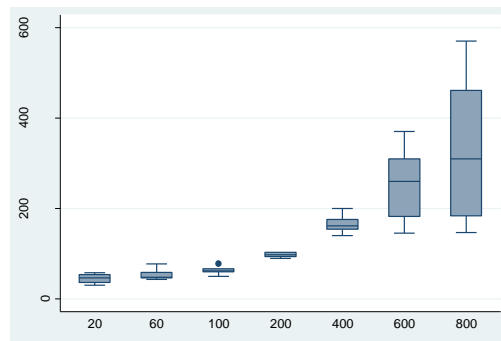


Figure 3.1: Thrust force boxplot against feed rate

### 3.1.2 Thrust force distribution study over spindle speed

In this section, the spindle speed is analyzed over the thrust force to understand their relationship and dependency. As shown in **table 3.4**, the p-value of the spindle speed is at 0.050, equal to the conventional reference point ( $\alpha = 0.05$ ), which indicates that the spindle speed does have a

limited effect on the thrust force trend. In fact, to perform this analysis, the most significant factor, the feed rate, is removed and its effect was added to the residual which makes the spindle speed effect irrelevant. The fact to pull a variable or an interaction from the ANOVA is known by “Pooling” [31]. In fact, the interaction effect and other unknown variables’ effects that do not appear in the model are added into the residual sum of squares and the model is estimated. To investigate the pooling effect, the adjusted coefficient of determination  $R^2$  must be analyzed. The outcome seen here is the “pooling” effect. The adjusted  $R^2$  is equal to 0.0846 which indicates that only 8.46% of the variability of the thrust force data is explained by the spindle speed. This indicator is very low and endorses the results given by the ANOVA that the spindle speed has very limited influence on the thrust force behavior.

Table 3.4: Parameter estimates and ANOVA results of thrust force against spindle speed

Source	SS	df	MS	Number of obs = 35		
Model	60585.2134	1	60585.2134	F( 1, 33) =	4.14	
Residual	482436.558	33	14619.2896	Prob > F =	0.0499	
				R-squared =	0.1116	
				Adj R-squared =	0.0846	
Total	543021.771	34	15971.2286	Root MSE =	120.91	

thrustforce	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
spindlespeed	-.008649	.0042486	-2.04	0.050	-.0172929	-5.16e-06
_cons	217.7367	41.12607	5.29	0.000	134.0651	301.4084

In **figure 3.2**, a slight decrease of the thrust force is observed with the increase of the spindle speed. Accordingly, the thrust force is inversely proportional to the spindle speed. Also, it is noticed that the variability of the thrust force within each level of the spindle speed is very wide. The overall data is positively skewed because the majority of the whiskers are pushed up.

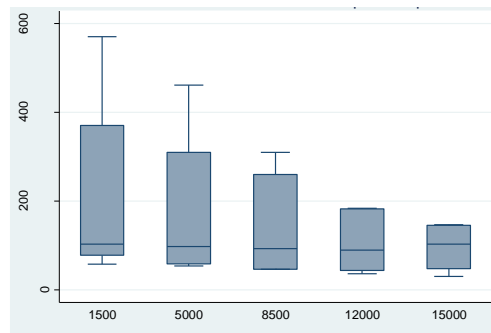


Figure 3.2: Thrust force boxplot against spindle speed

### 3.1.3 Thrust force multiple linear regression analysis

In this section, the MLR is developed which is an extension of simple linear regression in which more than one independent variable (feed rate and spindle speed) is used to predict a single dependent variable (thrust force). The predicted value of the thrust force is a linear transformation of the independent variables. From a mathematical point of view, the difference is that the simple linear regression only expresses two weights (the intercept  $\beta_0$  and the slope  $\beta_1$ ); while in this case, three weights ( $\beta_0$ ,  $\beta_1$ , and  $\beta_2$ ) are estimated.

As mentioned in previous sections, the parameter estimates, ANOVA and coefficient of determination results are generated to perform the MLR analysis. The p-value of each variable is shown in **table 3.6**. The feed rate and the spindle speed have both a p-value near 0, lower than the conventional reference point ( $\alpha = 0.05$ ). There is enough evidence to reject the null hypothesis at  $\alpha = 0.05$ . Therefore, by examining the p-value of each variable, the feed rate and spindle speed are significant at  $\alpha = 0.05$  confidence level. Both variables are important to the prediction of the thrust force. A remarkable change in significance is observed with the spindle speed compared to what has been observed in section 3.1.2.

As demonstrated in sections 3.1.1 and 3.1.2, while compiling the linear regression method, the feed rate contributed significantly to the thrust force outcome, while the spindle speed did not. In the multiple regression method and in the presence of the feed rate, the spindle speed is strongly significant as indicated by the near-zero p-values for both variables from **table 3.5**. This occurs because the spindle speed included in the model significantly reduces the sum of squares for the error and consequently, the mean square error; which is the denominator of the F-test, for testing hypotheses about the terms of the model. When this indicator is small, it will make the test statistic larger which allows experts to detect important differences such as significant variables. Also, the loss of degrees of freedom for estimating the error does not compensate the reduction in the sum of squares of the error. The adjusted  $R^2$  of the MLR of the thrust force data is 0.7877. This model fits the data well. There is good concurrence between the experimental and predicted values: 78.77% of the variability is explained by the feed rate and the spindle speed. The remaining 21.23% are attributed to unknown variability and noise (variables that are not being controlled by the researches). In fact, this model has no interaction effect which may cause this unknown variability. Since the p-value of this model is lower than 5% ( $0.00000 < 0.05$ ),  $H_0$  is

rejected and this regression model is judged to be useful. Based on the parameter estimates and ANOVA tables, this model is found to be acceptable with an explained variability of 78.77%, and the significant factors on the thrust force trend are the feed rate and the spindle speed.

Table 3.5: Parameter estimates and ANOVA results of the thrust force using MLR

Source	SS	df	MS	Number of obs = 35		
Model	434526.859	2	217263.429	F( 2, 32) =	64.08	
Residual	108494.913	32	3390.46603	Prob > F =	0.0000	
				R-squared =	0.8002	
				Adj R-squared =	0.7877	
Total	543021.771	34	15971.2286	Root MSE =	58.228	

thrustforce	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.3738568	.0355986	10.50	0.000	.3013448	.4463688
spindlespeed	-.008649	.002046	-4.23	0.000	-.0128167	-.0044814
_cons	101.3071	22.6972	4.46	0.000	55.07438	147.5398

Based on **table 3.5**, the fitted regression model of the thrust force is given by:

$$Tf = 101.307 + 0.3739 \cdot F - 0.0087 \cdot SS$$

### Interpretation of the coefficients

1. Constant ( $\beta_0$ ): when the feed rate and spindle speed are equal to zero, the expected mean of the thrust force is 101.3.
2. Feed rate ( $\beta_1$ ): for each change of one unit of the feed rate, the thrust force will increase by 0.037 when the spindle speed is held constant.
3. Spindle speed ( $\beta_2$ ): for each change of one unit of spindle speed, the thrust force will decrease by 0.0087 when the feed rate is held constant.

MLR attempts to model the relationship between the feed rate and the spindle speed by fitting a linear equation to the thrust force observed data. Therefore, it is interesting to review the plot of the predicted versus the observed of the thrust force values. A straight line is drawn to provide the best estimate of the observed trend. If the model is appropriate for the data, the points are expected to follow the straight line. However, if this plot displays a non-linear pattern, it indicates that the model doesn't fit the data properly. The distance between the predicted data and the fitted line are the residuals. **Figure 3.3** is a graphical illustration of the thrust force observed values against the predicted ones using the multiple linear regression method. This figure displays that:

- The regression fitted-model represented by the straight line,
- The values around the fitted line are the predicted values, and

- The interval of confidence at 95% is represented by the dashed lines.

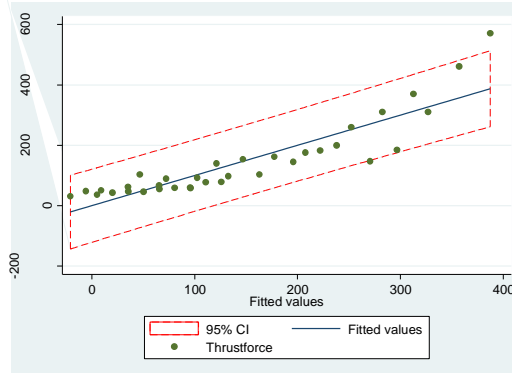


Figure 3.3: Thrust force plot fitted against observed values using MLR

By investigating the shape of the values around this line, it indicates a linear fit and the scatter of points lie roughly close to the regression line and are within the 95% confidence of interval of the predicted values. The overall trend tends to rise. However, some outliers are identified to be indicating a lack fit of the model. Further investigations are needed into the assumptions to check the validity of this model. Refer to **table C.2 in Annex C** for the predicted and observed values.

Before presenting the conclusions for this model-fitting, the appropriateness of the regression model for the thrust force data is examined.

**Independency assumption:** Figure 3.4 displays the graphical representation of the residuals against the run order of the data collection of the thrust force. This plot shows a random pattern of residuals, and does not show any recognizable patterns. Thus, the observations are independent.

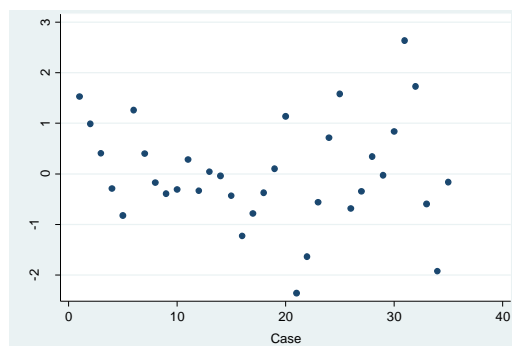


Figure 3.4: Plot of residuals of the thrust force data versus run order

**Normality assumption:** The normal probability plot below shows a slight deviation of the straight line at the lower and upper tail which could be potentially due to the influence of outliers.

Note that the line crossing the points is added as a reference line to detect easily any outliers. To better validate the normality assumption, the Shapiro-Wilk test is performed.

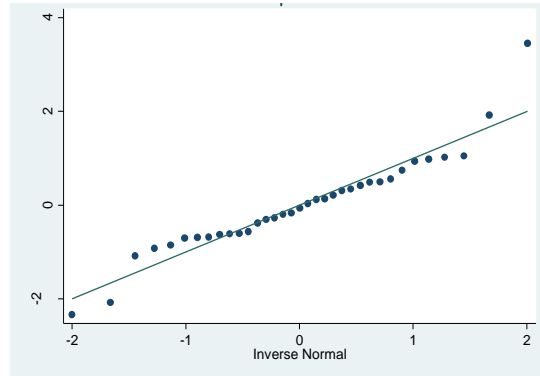


Figure 3.5: Normal probability plot for the thrust force data using MLR

Shapiro test implicates that the null hypothesis is that the data are normally distributed. If the calculated p-value is lower than 0.05, then the null hypothesis that the data are normally distributed is rejected. If the p-value is greater than 0.05, then the null hypothesis is not to be rejected. As indicated in **table 3.6**, the p-value is 0.04411, lower than 0.05. Consequently, the experimental data of the thrust force does not follow a normal distribution.

Table 3.6: Shapiro-Wilk Test for thrust force normality using MLR

Variable	Obs	W	V	z	Prob>z
rthrustforce	35	0.93659	2.263	1.705	0.04411

**Homoscedasticity assumption:** The residual plot below confirms that the variance of the residual is not constant and the values are not fluctuating randomly around zero. The line crossing the points is plotted as a reference to visualize the structure of the values. Most values are concentrated around 0-200 and there are many observations that differ from the main trend. The issue with the variance being not constant may be due to the influence of some observations to the model. However, in several cases, transforming the dependent variable or developing a non-linear model may resolve this issue without excluding any observations.

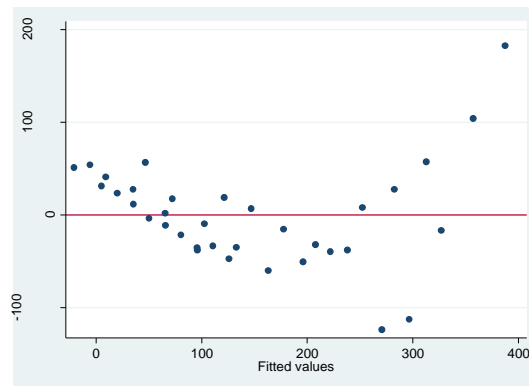


Figure 3.6: Residual plot for the thrust force data using MLR

**Linearity assumption:** The relationship between the independent variable and the dependent variable must be linear. To verify this assumption, the respective scatterplots of the thrust force against the feed rate, and the spindle speed are presented.

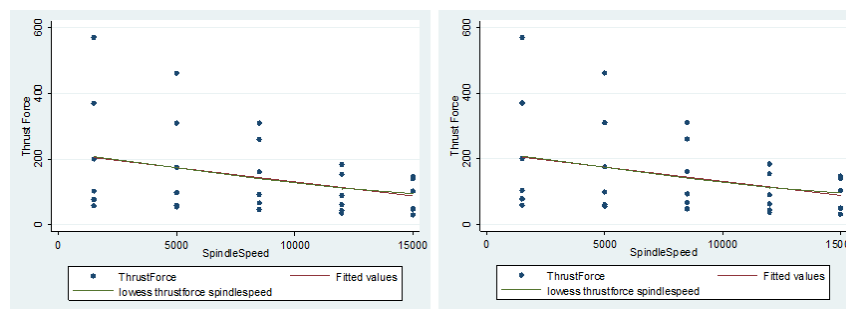


Figure 3.7: Scatterplots of thrust force versus feed rate and spindle speed

The “lowess” outline on both figures is generated by the software as the computations of the locally-weighted polynomial smoother. The name "lowess" is derived from the term "locally weighted scatter plot smooth". Its purpose is to fit a smoother outline to the data under study. If the data follows a linear distribution, this line will overlap the regression line. Otherwise, it will fit another model. **Figure 3.7** indicates that both lowess and the regression outlines are overlapping which confirm that the existent relationship between the thrust force and each independent variable is linear.

As explained previously, to be able to use the MLR model of the thrust force, assumptions must be verified and validated. Below is a summary of the assumptions' validation results for the thrust force multiple linear regression model (**Table 3.7**).

Table 3.7: Illustration of the assumption validation results for thrust force using MLR

Assumptions	Valid for this model
Independency	Yes
Normality	No
Homoscedasticity	No
Linearity	Yes

Based on this series of analysis, this model does not follow the normality and the homoscedasticity assumptions. In this case, the linear regression model does not suit this data and experts cannot rely on this fitting model to predict the trend of the thrust force at any point of this experimental domain. For further verifications and to better understand the relationship between the thrust force and the independent variable, the interaction between the feed rate and the spindle speed is added to the regression model fitting in the next section.

### 3.1.4 Thrust force multiple linear regression analysis with interaction

As described in the previous sections, the feed rate is an important variable to the thrust force regression model. However, the spindle speed was identified to moderately affect the thrust force. In subsequent analysis, an interaction effect between the main variables, feed rate and spindle speed is added to evaluate if this interaction against the thrust force is substantial.

As a standard procedure, the most important tables are pulled from the software to process the multiple regression analysis. **Table 3.8** demonstrates the parameter estimates, ANOVA and coefficient of determination results.

The ANOVA results above indicate that the feed rate with p-value equal to 0, lower than 0.05, and the effect between the feed rate and spindle speed with p-value corresponding to 0, lower than 0.05. As observed, an interaction is revealed important on the thrust force behavior. However, the spindle speed is not as significant as it was in the multiple linear regression analysis without interaction and this was caused by combining the interaction effect into the residual. In fact, the spindle speed has a p-value equal to 0.048, around the critical value of 0.05, which makes this variable non-significant. However, the spindle speed cannot be removed because his effect is reflected in its interaction with the feed rate.

Table 3.8: Parameter estimates and ANOVA results for thrust force using MLR-interaction

Source	SS	df	MS	Number of obs = 35		
Model	520323.15	3	173441.05	F( 3, 31) = 236.87		
Residual	22698.621	31	732.213582	Prob > F = 0.0000		
				R-squared = 0.9582		
				Adj R-squared = 0.9542		
Total	543021.771	34	15971.2286	Root MSE = 27.059		

thrustforce	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.686562	.0332897	20.62	0.000	.6186672	.7544569
spindlespeed	.0029445	.0014322	2.06	0.048	.0000235	.0058654
fds	-.0000372	3.44e-06	-10.82	0.000	-.0000442	-.0000302
_cons	3.921716	13.86343	0.28	0.779	-24.35293	32.19637

Based on **table 3.8**, the following MLR fitted model with interaction is pulled:

$$\mathbf{Tf = 3.922 + 0.687 \cdot F + 0.003 \cdot SS - 0.000037 \cdot F \cdot SS}$$

This model suggests that the thrust force is more sensitive to the feed rate than the spindle speed because the feed rate has a bigger impact and the spindle speed has a minimum influence (via the interaction), the overall trend is to rise. The adjusted  $R^2$  is at 95.42% which is a significant indication that this model may be a good fit. The variability of the multiple first-order regression model developed previously was 78.77% compared to 95.42% with the multiple regression with interaction. Consequently, the variability is improved, thus this model fits better the thrust force experimental data. Below is the plot of the observed against the fitted values of the thrust force using the MLR with interaction. This fitting seems to suit the data. Most of the predicted values are within the 95% interval of confidence. Few observations lie right at the edge but they are not outliers. Refer to **table C.3 in annex C** for the fitted and the observed data.

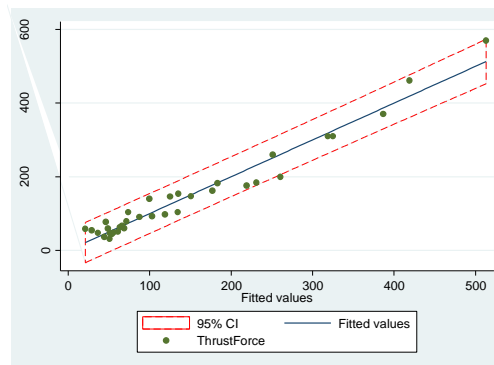


Figure 3.8: Thrust force plot fitted against observed values using MLR-interaction

To verify the adequacy of this model, the same assumptions must be verified except linearity and independency because it was already verified in the previous regression analysis: normality and homoscedasticity.

**Normality assumption:** Figure 3.9 is the normal probability plot of standardized residuals for the thrust force data to validate the normality assumption. It displays a straight line which represents that the random errors are normally distributed and no deviation from the straight line is observed. To better validate the normality assumption, the Shapiro-Wilk test is performed as well. From table 3.9, the calculated p-value corresponds to 0.69273, greater than 0.05. Consequently, the null hypothesis that the standard residuals are normally distributed cannot be rejected. Thus, the experimental data of the thrust force do follow a normal distribution in this case.

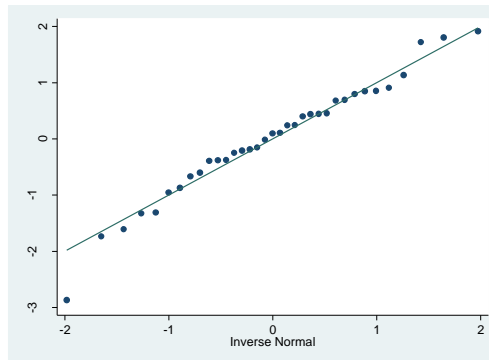


Figure 3.9: Normal probability plot of standardized residuals for the thrust force with MLR-interaction

Table 3.9: Shapiro-Wilk Test for Thrust Force Normality for MLR-interaction

Variable	Obs	W	V	z	Prob>z
rtffds	35	0.97799	0.786	-0.504	0.69273

**Homoscedasticity assumption:** Figure 3.10 is a plot of the predicted values against raw residuals to validate the Homoscedasticity assumption. The reference line at 0 emphasizes that the residuals are split about 50-50 between positive and negative. This plot confirms that this assumption is violated: the variance of the residual is not constant. To make sure that this assumption is violated, the BPCW test is completed and the chi-square value is at 8.76 and its p-value is 0.0031, lower than 0.05. This test indicates that the variance of the residuals is not constant.

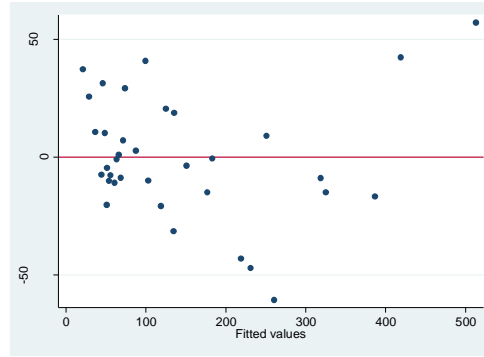


Figure 3.10: Thrust force fitted values against standardized residuals

Below is a summary of the assumptions' validation results. Because the homoscedasticity assumption is violated, this model cannot be accepted even with 96% accuracy. As discussed earlier, if a suitable model is not found with the MLR, a non-linear/transformation model will be attempted.

Table 3.10: Illustration of the assumption validation results

Assumptions	Valid for this model
Independency	Yes
Normality	Yes
Homoscedasticity	No
Linearity	Yes

### 3.1.5 Thrust force data transformation

The software tries different types of transformation and performs a statistical test to evaluate if the data is normally distributed. The transformation with a p-value greater than 0.05, indicates that the hypothesis that the data follows a normal distribution cannot be rejected. One of the important set of statistical tests is the chi-squared test which allows experts to test for deviations of observed frequencies from expected frequencies. In other words, it's called a goodness of fit. In this case, the chi-square is used to test if the transformed thrust force data came from a normal distribution. The null hypothesis in this test is that the data follows a normal distribution and the alternative hypothesis is that the data does not follow a normal distribution. If the p-value is lower than 0.05, the null hypothesis is to be rejected. In **table 3.11**, this method suggests that three transformations are possible with a p-value greater than 0.05:

- Inverse with p-value equals to 0.148,
- Log with p-value equals to 0.318, and

- $1/(\text{square root})$  equals to 0.271.

Table 3.11: Goodness of fit of thrust force transformations

Transformation	formula	chi2(2)	P(chi2)
cubic	$\text{thrust} \sim e^3$	39.60	0.000
square	$\text{thrust} \sim e^2$	30.85	0.000
identity	$\text{thrust} \sim e$	15.93	0.000
square root	$\sqrt{\text{thrust} \sim e}$	7.04	0.030
log	$\log(\text{thrust} \sim e)$	2.29	0.318
$1/(\text{square root})$	$1/\sqrt{\text{thrust} \sim e}$	2.61	0.271
inverse	$1/\text{thrust} \sim e$	3.82	0.148
$1/\text{square}$	$1/(\text{thrust} \sim e^2)$	17.87	0.000
$1/\text{cubic}$	$1/(\text{thrust} \sim e^3)$	30.99	0.000

In this case, the log of dependent variable is the best option whereas its p-value is the highest one to be over 0.05 and the null hypothesis cannot be rejected. In the following analysis, the log transformation is verified to construct the thrust force prediction model.

Table 3.12: Parameter estimates and ANOVA results of log-thrust force

Source	SS	df	MS	Number of obs = 35		
Model	17.9726175	2	8.98630875	F( 2, 32) =	162.16	
Residual	1.77327636	32	.055414886	Prob > F =	0.0000	
Total	19.7458939	34	.580761585	R-squared =	0.9102	
				Adj R-squared =	0.9046	
				Root MSE =	.2354	

ln tf	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.002469	.0001439	17.16	0.000	.0021758	.0027621
spindlespeed	-.0000453	8.27e-06	-5.48	0.000	-.0000622	-.0000285
_cons	4.291173	.0917605	46.76	0.000	4.104263	4.478083

The ANOVA results indicate that the feed rate and the spindle speed p-values are near 0, lower than 0.05 which indicate that these variables are both significant. Based on the table above, the following multiple regression fitted model is pulled:

$$\text{Log (Tf)} = 4.291 + 0.00247 \cdot F - 0.000045 \cdot SS$$

**Table 3.12** indicates as well that the adjusted  $R^2$  is at 90.5%; a very good estimation. The variability of the MLR with interaction developed previously was 95.42% compared to 90.46% with the log transformation. The variability is explained better by the MLR with interaction model; however, the assumptions were violated. In a transformation, the interpretation of the coefficients is slightly different:

1. Intercept  $\exp(\beta_0) = \exp(4.291) = 73.05$  : when the feed rate and the spindle speed are zero, the expected geometric mean of thrust force is 73.05.

2. Feed rate  $\exp(\beta_1)$ : for changes of one unit of feed, the expected geometric mean of thrust force expects to increase 13.4% since  $\exp(0.1259229) = 1.13$ .
3. Spindle Speed  $\exp(\beta_2)$ : for changes of one unit of spindle speed, the expected geometric mean of thrust force expects to have no change, since  $\exp(-0.0000453) = 1.00$ .

**Figure 3.11** is the plot of the log thrust force fitted model using a logarithmic transformation of the dependent. This plot displays a linear fit. Also, the predicted values are slightly lying on the fitted line and within the interval of confidence. This indicates that this model fits well the thrust force data. Refer to **table C.4 in Annex C** for the fitted and observed values of the log thrust force transformation.

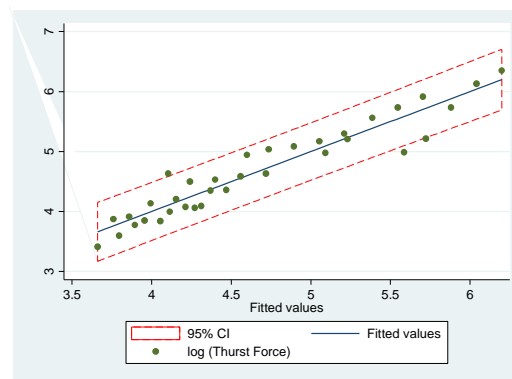


Figure 3.11: Plot of the log thrust force fitted model against the feed rate and spindle speed

To be able to use this model, the following assumptions must be validated: normality, homoscedasticity and linearity.

**Normality Assumption:** To validate the normality assumption, the Shapiro-Wilk test is performed. The p-value from **table 3.13** is equal to 0.08268, greater than 0.05. Consequently, the null hypothesis that the standard residuals are normally distributed cannot be rejected. Thus, this assumption is validated.

Table 3.13: Shapiro-Wilk Test for log thrust force

Variable	Obs	W	V	z	Prob>z
rlntf	35	0.94554	1.944	1.387	0.08268

**Homoscedasticity assumption:** **Figure 3.12** displays no specific pattern of the residuals so this assumption is validated. To validate it mathematically, the BPCW test is completed. The calculated chi-square value is 1.36 and its p-value is 0.2434, greater than 0.05. This test indicates as well that the variance of the residuals is constant.

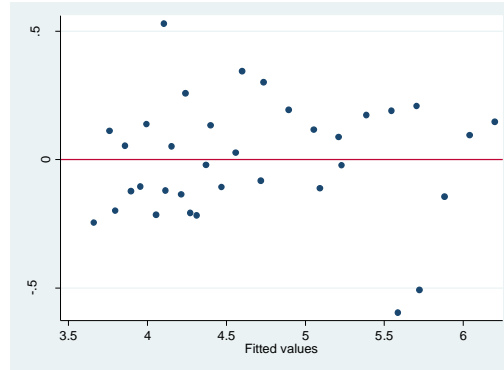


Figure 3.12: Plot of residual vs fitted values for log thrust force transformation

**Linearity Assumption:** As displayed on **figure 3.13**, the relationship is linear. The points on each plot are the observed thrust force value at each corresponding level of feed rate or spindle speed. The line crossing these points is the fitted trend of the overall data to provide us a visual illustration of the distribution.

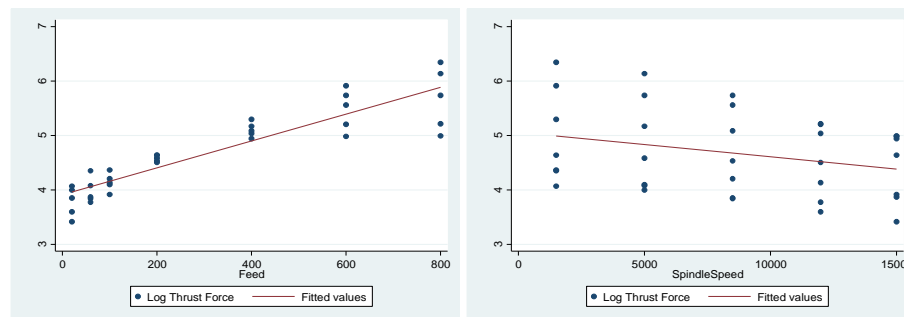


Figure 3.13: Scatterplot of feed rate and spindle speed against log thrust force

**Table 3.14** summarizes the assumptions' validation results. Because all assumptions are validated, this model can be used by the experts with 91% of accuracy.

Table 3.14: Illustration of the assumption validation results for log thrust force model

Assumptions	Valid for this model
Independency	Yes
Normality	Yes
Homoscedasticity	Yes
Linearity	Yes

### 3.1.6 Comparison of the methods used for thrust force

**Table 3.15** explores all the options and indicates the more appropriate model to use for prediction purposes. Based on the best variability and assumptions validation adequacy, researchers can readily apply this model to predict the thrust force trend by:

$$\text{Log (Tf)} = 4.291 + 0.00247 * F - 0.000045 * SS$$

Table 3.15: Comparison of all methods used for the thrust force analysis

Method used	Validity to use	Variability	Significant
MLR	Non-normality and	78.77%	Feed rate and spindle speed
MLR with	Variance not constant	95.42%	Feed rate and interaction
Log transformation	Valid	90.46%	Feed rate and spindle speed

## 3.2 Cutting force analysis

### 3.2.1 Cutting force distribution study over feed rate

For this simple linear regression analysis, the boxplot and the parameter estimates, ANOVA and coefficient of determination results are pulled. **Figure 3.14** indicates that the cutting force is higher as the feed rate increases. Thus, the feed rate is directly proportional to the cutting force. The cutting force reaches its max value when the feed rate is at its max as well. The variability is wider when the feed rate increases. No outliers are detected.

In **table 3.16**, the calculated p-value is near 0, lower than 0.05 which demonstrates that the feed rate has a significant effect on the cutting force. The adjusted  $R^2$  corresponds to 0.7930 which indicates that 79.3% of the variability is explained only by the feed rate; which constitutes a major effect on the cutting force. To perceive the trend of the cutting force against the feed rate, the boxplot below is drawn.

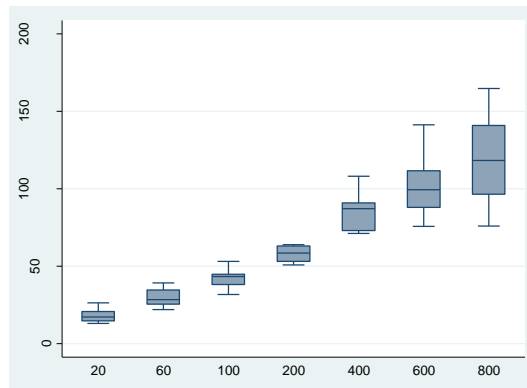


Figure 3.14: Boxplot of feed rate against cutting force

Table 3.16: Parameter estimates and ANOVA results of cutting force against feed rate

Source	SS	df	MS	Number of obs = 35		
Model	42423.1147	1	42423.1147	F( 1, 33) = 131.23		
Residual	10668.0582	33	323.27449	Prob > F = 0.0000		
				R-squared = 0.7991		
				Adj R-squared = 0.7930		
Total	53091.1728	34	1561.50508	Root MSE = 17.98		
cuttingforce	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.1259229	.0109923	11.46	0.000	.1035588	.1482869
_cons	26.12345	4.577721	5.71	0.000	16.81001	35.4369

### 3.2.2 Cutting force distribution study over spindle speed

In this section, a simple linear regression analysis is performed to understand the relationship and interaction between the cutting force against the spindle speed. From **table 3.17**, the calculated p-value corresponds to 0.067 which is greater than 0.05. This indicates that the spindle speed has no major effect on the cutting force. There are other combined variables into the error term that may have influence on the cutting force. The effect of the feed rate is combined into the residual. The adjusted  $R^2$  corresponds to 0.0705 which implies that 7.05% of the variability is explained by the spindle speed effect.

By examining the boxplot, on **figure 3.15**, a slight decrease of the cutting force with the increase of spindle speed is detected. The thrust force is inversely proportional to the spindle speed. The variability of the cutting force on each spindle speed value is very large; cutting force values vary significantly on each spindle speed level.

Table 3.17: Parameter estimates and ANOVA results of cutting force over spindle speed

Source	SS	df	MS	Number of obs = 35		
Model	5192.26432	1	5192.26432	F( 1, 33) = 3.58		
Residual	47898.9085	33	1451.48208	Prob > F = 0.0674		
				R-squared = 0.0978		
				Adj R-squared = 0.0705		
Total	53091.1728	34	1561.50508	Root MSE = 38.098		
cuttingforce	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
spindlespeed	-.002532	.0013387	-1.89	0.067	-.0052556	.0001917
_cons	86.60818	12.95866	6.68	0.000	60.2436	112.9728

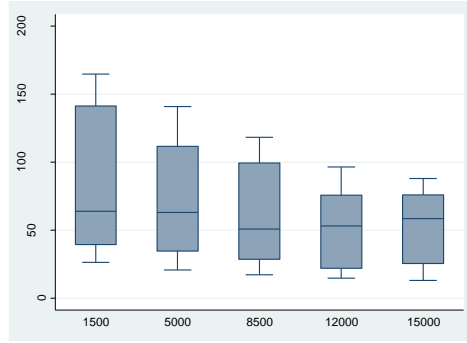


Figure 3.15: Boxplot of cutting force against spindle speed

### 3.2.3 Cutting force multiple linear regression analysis

The parameter estimates, ANOVA and coefficient of determination results are acquired to obtain the fitted model with the multiple linear regression technique. **Table 3.18** indicates that the feed rate and the spindle speed are both significant and affect the cutting force. From a statistical point of view, both p-values are inferior to 0.05. Also, the adjusted  $R^2$  is 89% which indicates that the model fits well the data. The remaining 11% are related to unknown variability.

Table 3.19: Parameter estimates and ANOVA results for cutting force using MLR

Source	SS	df	MS	Number of obs = 35		
Model	47615.379	2	23807.6895	F( 2, 32) = 139.13		
Residual	5475.79384	32	171.118557	Prob > F = 0.0000		
				R-squared = 0.8969		
				Adj R-squared = 0.8904		
Total	53091.1728	34	1561.50508	Root MSE = 13.081		
cuttingforce	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.1259229	.0079975	15.75	0.000	.1096326	.1422132
spindlespeed	-.002532	.0004597	-5.51	0.000	-.0034683	-.0015957
_cons	47.3922	5.099072	9.29	0.000	37.00573	57.77867

The estimated MLR model of the cutting force is given by:

$$Cf = 47.3922 + 0.126 \cdot F - 0.0025 \cdot SS$$

#### Interpretation of the coefficients

1. Constant ( $\beta_0$ ): when feed and spindle are zero, the expected mean of cutting force is 47.3922.
2. Feed ( $\beta_1$ ): for each one-unit change of feed, cutting force will increase by 0.126 when the spindle speed is constant.
3. Spindle Speed ( $\beta_2$ ): for each one-unit change of spindle speed, cutting force will decrease by 0.0025 when the feed rate is constant.

By investigating the regression coefficients, it is noticed that the spindle speed has lower influence on the cutting force compared to the feed rate.

**Figure 3.16** displays the fitted values on the horizontal axis and the observed values on the vertical axis. The utility of this graph is to view whether the model predicts any value poorly. Ideally, the points should lie on a 45-degree straight line. In this case, the points are fluctuating around the fitted line. Hence, the data may fit the model properly. **Table C.5 in annex C** presents the data of this plot.

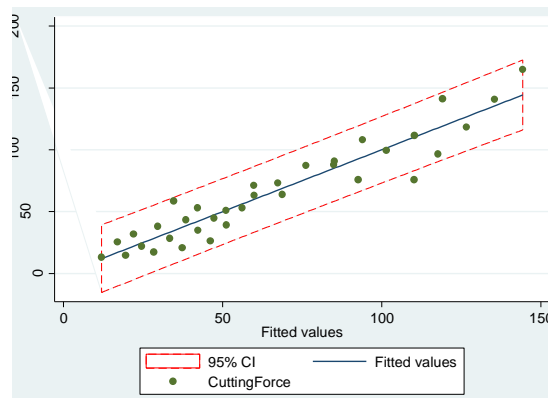


Figure 3.16: Plot of the observed vs predicted values for cutting force using MLR

As explained previously, the assumptions shall be verified to validate the accuracy of the MLR.

**Independency assumption:** **Figure 3.17** displays the graphical representation of the residuals of the fitted cutting force against the run order of the data collection for the cutting force data. As observed, the independence assumption is violated because the residuals have a structured pattern. Some variables that change over order time are essential and haven't been included in the design. The relationship between the case number and the residuals seems to be curvilinear.

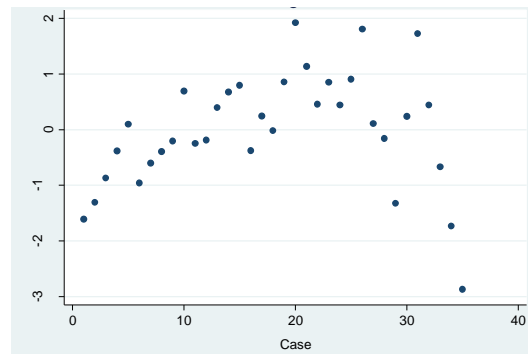


Figure 3.17: Plot of standardized residuals of the cutting force versus case number

**Normality assumption:** Figure 3.18 represents the normality fit of the standardized residuals of the cutting force experimental data. The residual values lie on the straight line, thus the random errors are normally distributed and no outliers are detected. Thus, the cutting force experimental data follows a normal distribution.

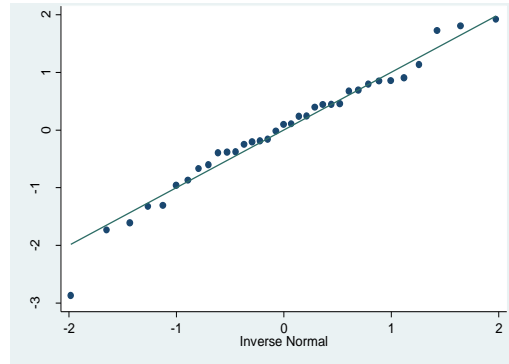


Figure 3.18: Normality plot of the standardized residuals of the cutting force data

To verify mathematically this assumption, the Shapiro-Wilk test for normality is completed in **table 3.19**. The calculated p-value is 0.64561, greater than 0.05. Consequently, the null hypothesis (data follows a normal distribution) cannot be rejected. This test confirms the reasonableness of the conclusion that the error terms are fairly normally distributed.

Table 3.19: Shapiro-Wilk test for cutting force using MLR

Variable	Obs	W	V	z	Prob>z
rcut	35	0.97657	0.836	-0.373	0.64561

**Homoscedasticity assumption:** Figure 3.19 allows a visual assessment between the predicted values of the cutting force and its raw residuals. It illustrates that the variance of the residual is non-constant due to a structured pattern. When observing this figure, some experts will think that the variance is constant. To confirm this result, BPCW test is completed. Thus, the calculated square-chi value is 4.24 and its p-value is at 0.0395, lower than 0.05 which indicates that the variance is not constant.

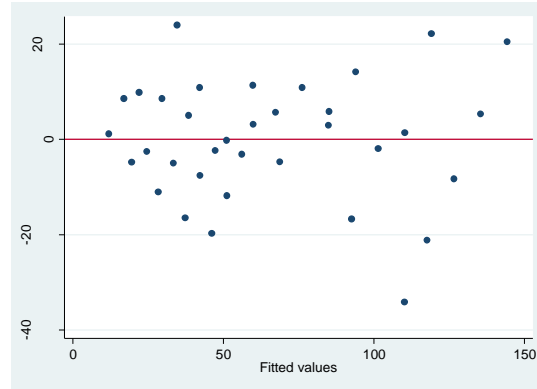


Figure 3.19: Cutting force fitted values versus residuals

**Linearity assumption:** The linearity assumption is assessed by looking at **figure 3.20** representing respectively the scatterplots of the feed rate and the spindle speed against the cutting force.

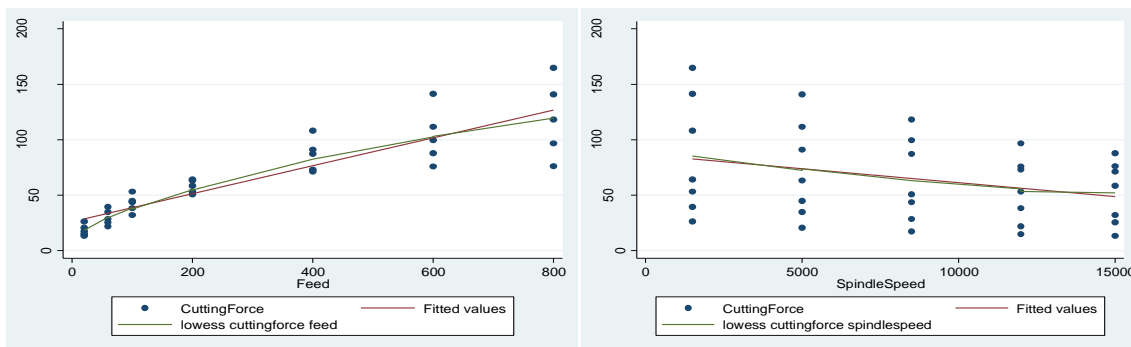


Figure 3.20: Scatterplots feed rate and spindle speed against cutting force

In fact, by observing these plots, they indicate that the relationship between the feed rate and the cutting force, and spindle speed and the cutting force are both linear. The outline of the lowess of the feed rate overlays the fitted values line to indicate a moderately linear fit between the feed rate and cutting force. And, the lowess of the spindle speed overlaps perfectly the fitted values line to imply a linear fit between the spindle speed and cutting force.

**Table 3.20** describes the validation results of the assumptions of the cutting force model using the MLR. The model proposed by the MLR cannot be used because the independency and homoscedasticity assumptions are violated; the fitted values from the linear regression will be biased and estimates are not meaningful. With this model, an increased risk of Type I errors may be perceived. In further sections, another fit of regression analysis will be fitted to find the most

suitable prediction model. Also, the independency assumption may be violated because of the high correlation between the variables.

Table 3.20: Illustration of the assumption validation results for cutting force MLR

Assumptions	Valid for this model
Independency	No
Normality	Yes
Linearity	Yes
Homoscedasticity	No

### 3.2.4 Cutting force multiple linear regression with interaction

In this section, the data are attempted to fit a MLR with interaction. This method evaluates the interaction between the feed rate and the spindle speed against the cutting force in addition to the separate variables.

**Table 3.21** shows that the feed rate and the interaction between the feed rate and spindle speed are significant. It is evident that their p-values are near 0, lower than 0.05. The spindle speed has no impact on the cutting force whereas its p-value equals to 0.43, greater than 0.05. The adjusted  $R^2$  is equal to 94.78% which is considered a very good estimate; since only 5.22% is related to unknown variability. The thrust force is more sensitive to the feed rate than the spindle speed because the feed rate has a bigger impact; the overall trend is to rise.

Table 3.21: Parameter estimates and ANOVA results for cutting force using MLR-interaction

Source	SS	df	MS	Number of obs = 35		
Model	50565.0779	3	16855.026	F( 3, 31) =	206.84	
Residual	2526.09494	31	81.4869335	Prob > F	= 0.0000	
				R-squared	= 0.9524	
				Adj R-squared	= 0.9478	
Total	53091.1728	34	1561.50508	Root MSE	= 9.027	

cuttingforce	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.1839044	.0111054	16.56	0.000	.1612547	.2065541
spindlespeed	-.0003823	.0004778	-0.80	0.430	-.0013568	.0005921
fds	-6.90e-06	1.15e-06	-6.02	0.000	-9.24e-06	-4.56e-06
_cons	29.33509	4.62483	6.34	0.000	19.90269	38.76749

Based on **table 3.21**, the fitted regression model for the cutting force is expressed by:

$$Cf = 29.335 + 0.184 \cdot F - 0.0004 \cdot SS - 0.000007 \cdot F \cdot SS$$

**Figure 3.21** illustrates the estimated regression line of the cutting force using the MLR with interaction. This plot indicates a linear relationship between the dependent and independent variables. No outliers are detected and most predicted values are within the interval of

confidence. This model may be used if the assumptions are validated. Refer to **table C.6 in annex C** for this plot data.

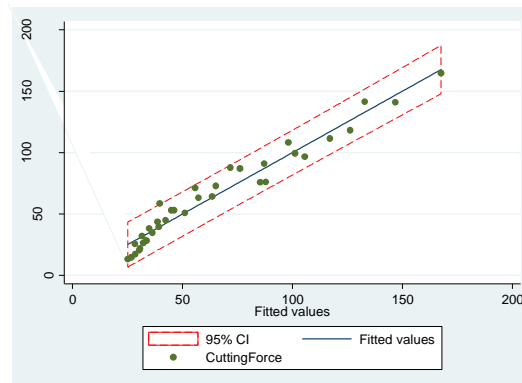


Figure 3.21: Observed vs predicted values of the cutting force using MLR-interaction

The variability of the MLR model developed previously for the cutting force was 89% compared to 94.78% by adding the interaction effect. Accordingly, the variability is enhanced with this type of regression. Before concluding that this model can be used to predict the cutting force, the following assumptions must be validated: normality and homoscedasticity. The independency and linearity assumptions have already been completed in the previous analysis.

**Normality assumption:** Figure 3.22 displays that the distribution follows a normal distribution. The values are distributed around the straight line so we can assume that the random errors are normally distributed and no extreme points are perceived. When the Shapiro-Wilk test (**table 3.23**) is performed, the calculated p-value is 0.35199, greater than 0.05. Consequently, this test confirms the visual examination.

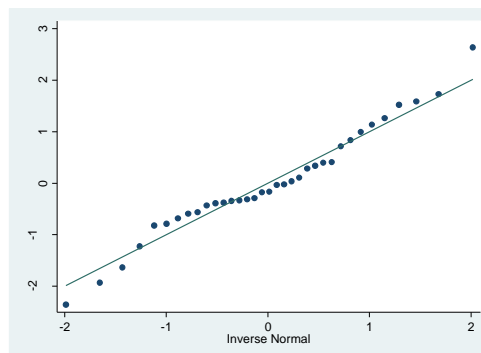


Figure 3.22: Normality plot for cutting force residuals using MLR-interaction

Table 3.22: Shapiro-test for cutting force using MLR-interaction

Variable	Obs	W	V	z	Prob>z
rcffds	35	0.96639	1.200	0.380	0.35199

**Homoscedasticity assumption:** The residual plot below indicates that there is no pattern within the data and the values are fluctuating randomly around zero. Therefore, the variance is constant. To verify that the geometric illustration is interpreted correctly, the BPCW test is completed. The computed square-chi value is 0.22 and the p-value is equal to 0.6413, greater than 0.05. In this case, the variance is constant even if the chart may provide us a different reasoning.

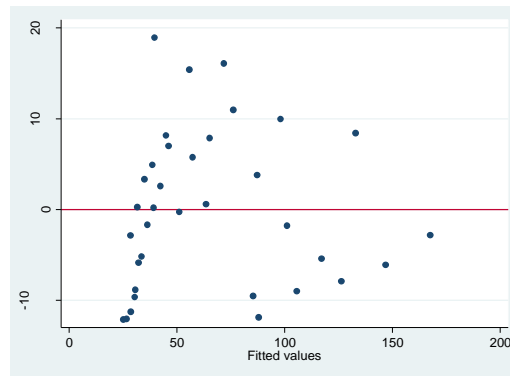


Figure 3.23: Plot of fitted against residual values for cutting force with MLR-interaction

**Table 3.23** describes the validation results of the assumptions for this method. The presence of the interaction term provides a slight better fit of the model from 89% compared to 94.78%. However, it is important to keep in mind that the interpretation of the coefficients is different in the presence of the interaction term. For instance, as the interaction term is statistically significant, it means that the feed rate has an effect on cutting force but the intensity of this effect depends on the spindle speed value.

Table 3.23: Illustration of the assumption validation results for cutting force using MLR-interaction

Assumptions	Valid for this model
Independency	No
Normality	Yes
Linearity	Yes
Homoscedasticity	Yes

### 3.2.5 Comparison of the methods used for cutting force data

To find an accurate model to predict the cutting force with feed rate and spindle speed as inputs, two methods are qualified: multiple regression with and without an interaction effect. To compare the studied models, the AIC is generated for each one.

Table 3.24: Akaike's results for Thrust force model with MLR–no interaction

Model	Obs	ll (null)	ll (model)	df	AIC	BIC
.	35	-177.8402	-138.0859	3	282.1718	286.8378

Table 3.25: Akaike's results for Thrust force model with MLR–with interaction

Model	Obs	ll (null)	ll (model)	df	AIC	BIC
.	35	-177.8402	-124.5468	4	257.0936	263.315

Tables 3.24 and 3.25 indicate that:

$$\text{AIC}_{\text{no interaction}} = 286.84 \text{ compared to } \text{AIC}_{\text{with interaction}} = 263.32$$

Table 3.26 explores all the models and presents the most useful one. The MLR with interaction improves the model by 5.78%. In fact, the AIC of this model is slightly smaller and has a better variability. Experts can apply the following regression model with interaction to predict the cutting force trend:

$$\text{Cf} = 29.335 + 0.184 * \text{F} - 0.0004 * \text{SS} - 0.000007 * \text{F} * \text{SS}$$

Table 3.26: Comparison of main methods used for cutting force analysis

Method used	Assumptions	Variability	Significant variables	AIC
MLR	Variance not constant	89.04%	Feed rate and spindle speed	286.84
MLR with interaction	Valid	94.78%	Feed rate and interaction	263.32

This experiment is achieved by running randomly each observation while drilling the composite. The observations are not dependent on the time of collection, but highly correlated among each other. Therefore, the independency assumption would be violated, this model is still recommended to use.

### 3.3 Torque analysis

#### 3.3.1 Torque distribution study over feed rate

To evaluate the significance of the feed rate on the torque, the parameter estimates, ANOVA and coefficient of determination results are generated. From **table 3.27**, the calculated p-value of the feed rate is near 0, lower than 0.05 which suggests that the feed rate has a significant effect on the torque. The adjusted  $R^2$  corresponds to 0.7930 which is considered a very good estimate. 79.3% of the variability is explained by the feed rate. The rest of the variability (20.7%) is attributed to missing other explanatory variables and unknown parameters.

Table 3.27: Parameter estimates and ANOVA results for the torque against the feed rate

Source	SS	df	MS	Number of obs = 35		
Model	1060585.25	1	1060585.25	F( 1, 33) =	131.23	
Residual	266699.478	33	8081.80237	Prob > F =	0.0000	
				R-squared =	0.7991	
				Adj R-squared =	0.7930	
Total	1327284.73	34	39037.7861	Root MSE =	89.899	

torque	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.6296165	.0549614	11.46	0.000	.5177968	.7414362
_cons	130.6169	22.88852	5.71	0.000	84.04981	177.1839

A box-and-whisker plot of the torque, **figure 3.24**, is drawn below for each group of the feed rate, providing a visual representation of the differences between the ranges under study. The boxplot demonstrates that the torque has a rising relationship with the feed rate. In fact, the torque is higher as the feed rate increases. The feed rate is directly proportional to the torque. Also, the variability of the torque is smaller when the feed rate ranges between 20 and 200. When the feed rate varies between 400 and 800, the torque variance is more stretched which involves more variability between the observations. No outliers are detected.

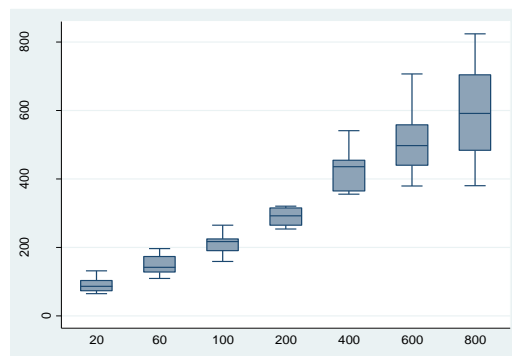


Figure 3.24: Boxplot of the torque against the feed rate

### 3.3.2 Torque distribution study over spindle speed

In this section, the spindle speed is analyzed as the only effect on the torque. To verify the effect of the spindle speed on the torque, the parameter estimates, ANOVA and coefficient of determination results are produced in **table 3.28**. The p-value of the spindle speed corresponds to 0.067, greater than 0.05 which means that the spindle speed is not a significant effect on the torque. As a matter of fact, when the feed rate is removed, its effect is combined to the residual (pooling result) which makes the spindle speed effect inappropriate.

Table 3.28: Parameter estimates and ANOVA results for the torque against spindle speed

Source	SS	df	MS	Number of obs = 35		
Model	129809.995	1	129809.995	F( 1, 33) =	3.58	
Residual	1197474.73	33	36287.1131	Prob > F =	0.0674	
				R-squared =	0.0978	
				Adj R-squared =	0.0705	
Total	1327284.73	34	39037.7861	Root MSE =	190.49	

torque	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
spindlespeed	-.0126601	.0066936	-1.89	0.067	-.0262784	.0009581
_cons	433.0426	64.79333	6.68	0.000	301.2195	564.8656

On **figure 3.25**, the boxplot displays the relationship between the spindle speed and the torque. A slight decrease is observed of the torque value with the increase of the spindle speed. The boxes are all centered which indicates that the distribution follows a normal distribution. Also, no significant shift of the boxes is observed and no outliers are detected.

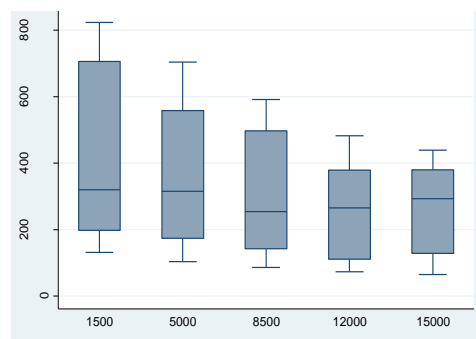


Figure 3.25: Boxplot of the torque against the spindle speed

### 3.3.3 Torque multiple linear regression analysis

In the previous sections, the relationship between each independent variables and the torque was exposed to understand the behavior of the outcome when varying the levels of the feed rate and the spindle speed. In this section, both explanatory variables are put together to present the

multiple linear regression analysis of the torque. The parameter estimates, ANOVA and coefficient of determination tables are generated in **table 3.29** to process the regression analysis. It demonstrates that the feed rate and the spindle speed have significant effects on the torque. In fact, both p-values are close to zero, lower than 0.05. Also, this model has an adjusted  $R^2$  of 89% which reveals a good fit of the model to the data. As observed earlier, the feed rate has the most impact on the torque compared to the spindle speed. The MLR model of the torque is given by:

$$T = 236.962 + 0.63 \cdot F - 0.127 \cdot SS$$

Table 3.29: Parameter estimates and ANOVA tables for the torque using MLR

Source	SS	df	MS	Number of obs = 35		
Model	1190395.24	2	595197.622	F( 2, 32) =	139.14	
Residual	136889.483	32	4277.79635	Prob > F =	0.0000	
				R-squared =	0.8969	
				Adj R-squared =	0.8904	
Total	1327284.73	34	39037.7861	Root MSE =	65.405	

torque	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.6296165	.0399865	15.75	0.000	.5481667	.7110663
spindlespeed	-.0126601	.0022982	-5.51	0.000	-.0173415	-.0079788
_cons	236.962	25.49486	9.29	0.000	185.0307	288.8933

### Interpretation of the coefficients

1. Constant ( $\beta_0$ ): when feed and spindle are zero, the expected mean of torque is 236.9625.
2. Feed rate ( $\beta_1$ ): for each one-unit change of feed rate, the torque will increase by 0.63 when the spindle speed is constant.
3. Spindle speed ( $\beta_2$ ): for each one unit change of spindle speed, the torque will decrease by 0.127 when the feed rate is constant.

**Figure 3.26** illustrates the predicted values on the vertical axis by the observed values on the horizontal axis. The observed relationship is linear and most of the predicted values are within the interval of confidence of 95% which may indicate a good fit to the model. **Table C.7 in annex C** presents the observed and predicted torque data.

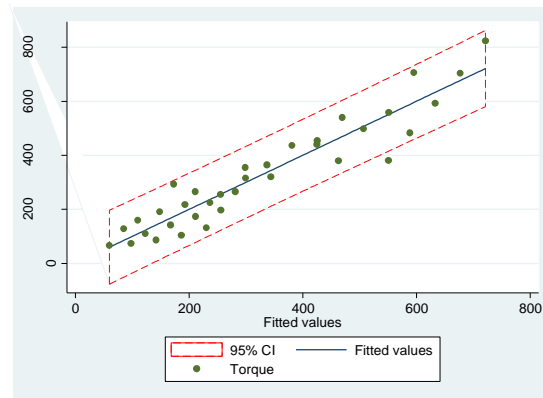


Figure 3.26: Plot of observed versus predicted values for torque using MLR

Even if the model's variability is good, the four assumptions discussed previously shall be validated before confirming the adequacy of this model.

**Independency assumption:** Figure 3.27 is a graphical representation of the standardized residuals of the fitted torque against the run order of the data collection of the torque data. As observed, the independence assumption is violated because the residuals have a curvilinear pattern.

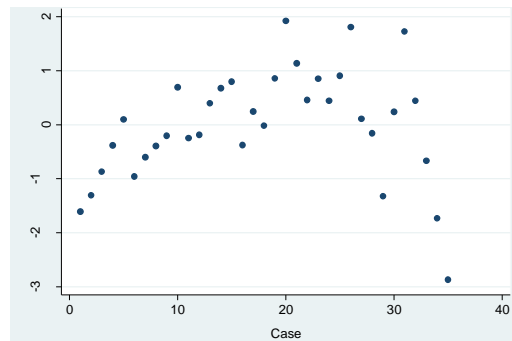


Figure 3.27: Plot of the standardized residuals of the fitted torque against the run order of the data collection

**Normality assumption:** Figure 3.28 shows a straight line. Thus, the random errors are normally distributed and no outliers are identified. By performing the Shapiro-Wilk test for normality checking (table 3.31), the generated p-value is 0.64571, greater than 0.05. This test endorses the same conclusion as the graphical illustration: the normality assumption is validated.

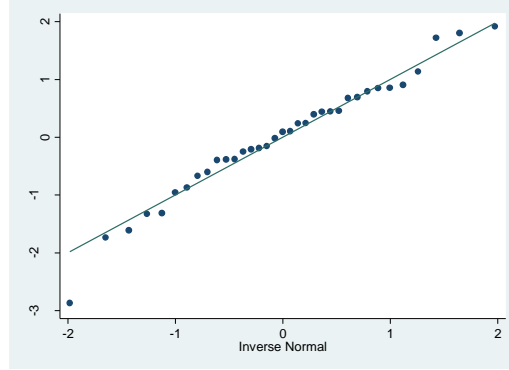


Figure 3.28: Normality plot for the torque experimental data

Table 3.30: Shapiro-Wilk test for normal data for torque using MLR

Variable	Obs	W	V	z	Prob>z
rtor	35	0.97658	0.836	-0.374	0.64571

**Homoscedasticity assumption:** Figure 3.29 presents the standardized residuals of the torque experimental data. It displays that the variance of the residual is not constant with a bent shape pattern. To verify that the geometric illustration is providing accurate information, the BPCW test is completed and the calculated square-chi value is equal to 4.24 and the p-value corresponds to 0.0395, lower than 0.05. In this case, the variance is not constant which confirms the same results pulled from the graph.

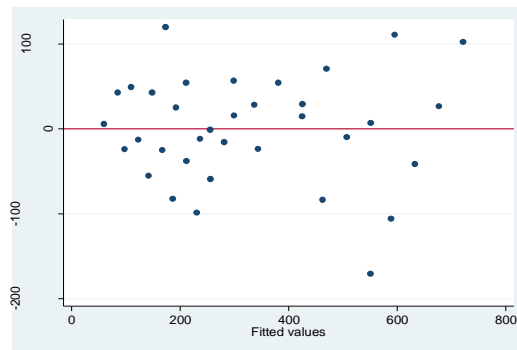


Figure 3.29: Observed vs residual values plot of the torque data using MLR

**Linearity assumption:** To verify the linearity, the relationship of the torque against both independent variables is investigated by plotting the scatterplots below. Figure 3.30 indicates that the relationship between the torque-feed rate, and the torque-spindle speed are both linear. The lowess outline of the feed rate-torque slightly overlaps the fitted regression line. However, the one for the spindle speed-torque is close to a perfect match of a linear fit.

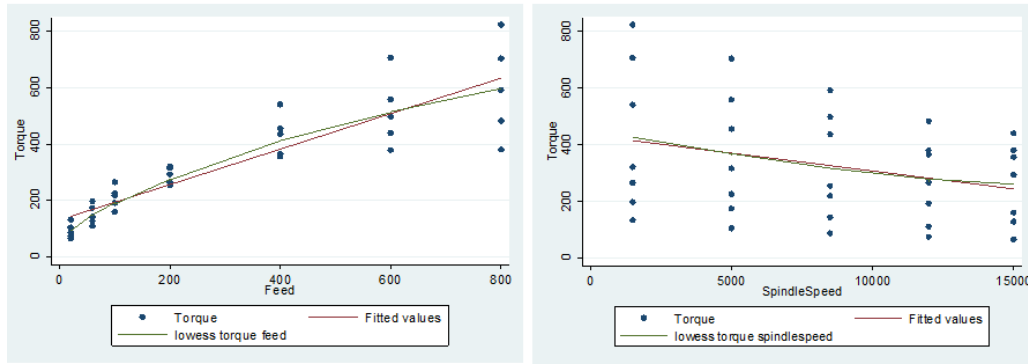


Figure 3.30: Scatterplots feed rate and spindle speed against torque

**Table 3.31** describes the validation results of the assumptions. Because two of the four assumptions are violated, the fitted model above shall not be used for prediction purposes.

Table 3.31: Illustration of the assumption validation results for torque using MLR

Assumptions	Valid for this model
Independency	No
Normality	Yes
Homoscedasticity	No
Linearity	Yes

### 3.3.4 Torque multiple linear regression with interaction

As explored in previous sections, the feed rate has a major effect on the torque unlike the spindle speed. For further investigation, the data is fitted to the MLR technique by adding an interaction effect. This method allows researchers to evaluate the interaction between the feed rate and the spindle speed against the torque in addition to the other effects.

The ANOVA results below (**table 3.32**) show that the feed rate ( $p\text{-value} = 0 < 0.05$ ) and the effect between the feed rate and spindle speed ( $p\text{-value} = 0 < 0.05$ ) are significant. However, the spindle speed ( $p\text{-value} = 0.430 > 0.05$ ) has no important impact on the torque. Also, the adjusted  $R^2$  is very good which corresponds to 94.78%; Very little variability (5.22%) is left unexplained. The presence of the interaction term provides a better fit of the torque model from 89% to 94.78%. The variability is explained better so the second model fits well the thrust force. However, it is important to keep in mind that interpretation of the coefficients is different in the presence of the interaction term. For instance, as the interaction term is statistically significant, the feed rate has an effect on torque but the intensity of this effect depends on the spindle speed.

Table 3.32: Parameter estimates and ANOVA tables for torque using MLR with interaction

Source	SS	df	MS	Number of obs = 35		
Model	1264133.04	3	421377.681	F( 3, 31) =	206.85	
Residual	63151.6843	31	2037.15111	Prob > F =	0.0000	
				R-squared =	0.9524	
				Adj R-squared =	0.9478	
Total	1327284.73	34	39037.7861	Root MSE =	45.135	

	torque	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed		.919515	.0555269	16.56	0.000	.8062673	1.032763
spindlespeed		-.0019122	.0023889	-0.80	0.430	-.0067843	.0029599
fds		-.0000345	5.74e-06	-6.02	0.000	-.0000462	-.0000228
_cons		146.6793	23.12402	6.34	0.000	99.51756	193.8411

From the table above, the following fitted regression model is constructed:

$$T = 146.68 + 0.92 \cdot F - 0.0019 \cdot SS - 0.000035 \cdot F \cdot SS$$

**Figure 3.31** displays the estimated regression line of the torque. The confidence band at 95% is the dash red line. Most of the fitted values are within this interval which indicates a good fit of the estimated model. Also, the predicted values lie approximately on the linear fitted plot. By this visual inspection, the model fits well the data. **Table C.8 in annex C** the data for this plot.

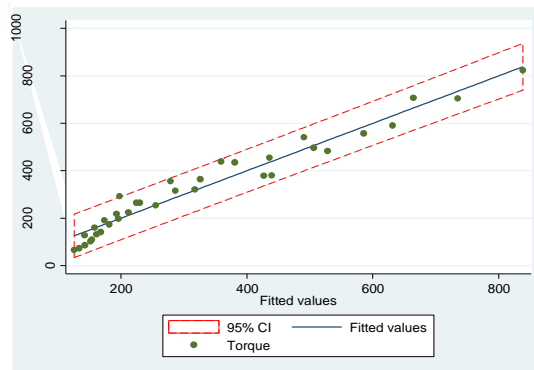


Figure 3.31: Plot of torque observed against predicted values using MLR-interaction

Knowing that the independency and linearity assumptions were already completed in the previous section, the normality and homoscedasticity assumptions will be verified.

**Normality assumption:** **Figure 3.32** shows a straight line which indicates that the random errors are normally distributed and no critical deviation from the straight line is observed. The Shapiro-Wilk test (**table 3.32**) indicates that the p-value corresponds to 0.35176, greater than 0.05. Thus, the null hypothesis that the data follows a normal distribution cannot be rejected. This test confirms the plot's results.

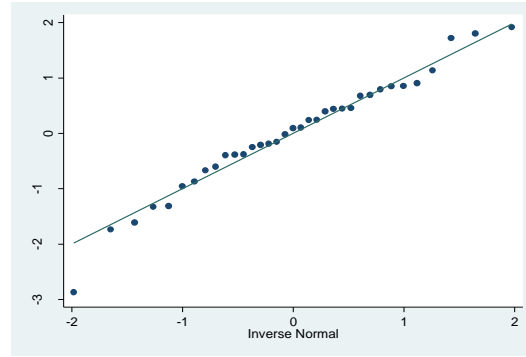


Figure 3.32: Normal probability plot of the torque using the MLR-interaction

Table 3.32: Shapiro-Wilk test for torque using MLR-interaction

Variable	Obs	W	V	z	Prob>z
rtorquefds	35	0.96638	1.200	0.381	0.35176

**Homoscedasticity assumption:** Figure 3.33 is the plot of the predicted values of the torque data against the raw residuals to validate the homoscedasticity assumption. This plot confirms that this assumption is validated. The same observations are displayed in **figure 3.23** due to variables being highly correlated.

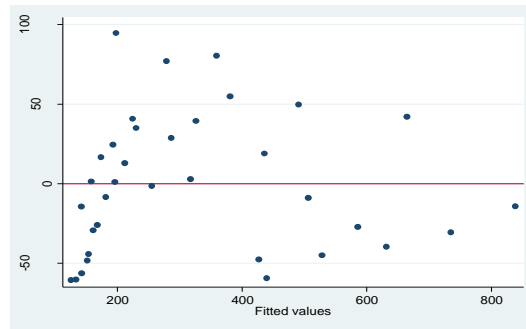


Figure 3.33: Predicted torque values against residuals using MLR-interaction

As discussed in the previous section, the independency assumption is violated due to a strong correlation among the variables which implicates that this model can be used with 94.78% of accuracy. **Table 3.33** describes the validation results of the assumptions.

Table 3.33: Illustration of the assumption validation results for the torque MLR-interaction

Assumptions	Valid for this model
Independency	No
Normality	Yes
Homoscedasticity	Yes
Linearity	Yes

### 3.3.5 Comparison of the methods used for torque

To find an accurate model to predict the torque with feed rate and spindle speed as inputs, MLR method was applied. **Table 3.34** explores all the options and presents which one to use.

Table 3.34: Comparison of all methods used for torque analysis force analysis

Method used	Validity to use	Variability	Significant variables
MLR	Variance not constant	89.04%	Feed rate and spindle speed
MLR with interaction	Valid	94.78%	Feed rate and interaction

Experts shall apply the MLR fitted model with interaction to predict the torque trend by:

$$T = 146.68 + 0.92 \cdot F - 0.0019 \cdot SS - 0.000035 \cdot F \cdot SS$$

As explained in **chapter 1**, the torque and the cutting force are highly correlated thus, they behave similarly. In fact, by examining both fitted models, the torque is observed to be equal to five times the cutting force. Here is a recall of the cutting force model using the MLR with interaction:

$$Cf = 29.335 + 0.184 \cdot F - 0.0004 \cdot SS - 0.000007 \cdot F \cdot SS$$

To confirm this relationship, the following is the mathematical comparison of the values:

- Intercept of the torque = 146.68 vs Intercept of the cutting force = 29.335. By dividing 146.68 by 29.335, the ratio is equal to 5.
- $\beta_{FT} = 0.92$  vs  $\beta_{FCf} = 0.184$ . By dividing 0.92 by 0.184, the ratio is equal to 5.
- $\beta_{SST} = 0.0019$  vs  $\beta_{SSCf} = 0.0004$ . By dividing 0.0019 by 0.0004, the ratio is equal to 5.
- $\beta_{IT} = 0.000035$  vs  $\beta_{ICf} = 0.000007$ . By dividing 0.000035 by 0.000007, the ratio is equal to 5.

## 3.4 Delamination at entry analysis

### 3.4.1 Delamination at entry distribution study over feed rate

In this section, the feed rate is considered the only effect on the delamination at entry. **Figure 3.34** displays the association between the feed rate and the delamination at entry to be linear. The outline shape is investigated in details in an anterior section. In fact, the delamination at entry is higher as the feed rate increases. The feed rate is directly proportional to the delamination at entry. The delamination at entry is directly proportional to the feed rate; it varies depending if the feed rate increases or decreases. At feed rate 20, 60 and 100, flat boxes are identified because the delamination at entry has the same value (it's equal to 1). The first outlier is distinguished at feed

rate equals to 20 whereas delamination at entry is at 1.081. Six other outliers are detected on the plot at different feed rate values when feed rate ranges from 200 to 800. The variance at each range of feed rate is considered small.

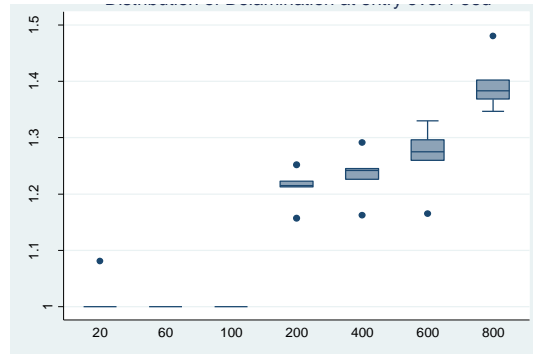


Figure 3.34: Boxplot of delamination at entry against feed rate

To evaluate the effect of the feed rate on the delamination at entry, the parameter estimates and ANOVA table is generated. **Table 3.35** indicates that the computed p-value of the feed rate is near 0, lower than 0.05 which designates that the feed rate has a significant effect on the delamination at entry. The adjusted R<sup>2</sup> corresponds to 0.8215. This indication is very good considering that only the feed rate is involved in the delamination at entry. Only 17.85% of the variability is not explained.

Table 3.35: Parameter estimates and ANOVA results for the delamination at entry against feed rate

Source	SS	df	MS	Number of obs = 35		
Model	.643078995	1	.643078995	F( 1, 33) = 157.46		
Residual	.134770913	33	.004083967	Prob > F = 0.0000		
Total	.777849908	34	.022877938	R-squared = 0.8267		
				Adj R-squared = 0.8215		
				Root MSE = .06391		
delentry	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0004903	.0000391	12.55	0.000	.0004108	.0005698
_cons	1.007659	.0162706	61.93	0.000	.9745558	1.040762

### 3.4.2 Delamination at entry distribution study over spindle speed

In this section, the spindle speed is the only effect analyzed on the delamination at entry. **Figure 3.35** displays the relationship between the spindle speed and delamination at entry. It indicates a slight decrease of the delamination at entry with the increase of spindle speed. Also, no outliers

are detected. However, the variance at each spindle speed range is high which suggests that the data is spread out and not concentrated around similar values.

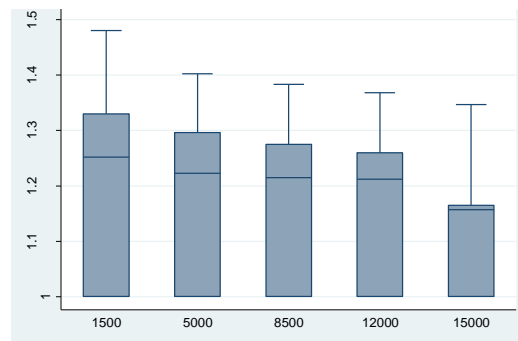


Figure 3.35: Boxplot of delamination at entry against spindle speed

To detect the effect of the spindle speed on the delamination at entry, the parameter estimates and ANOVA results are generated in **table 3.36**. The calculated p-value of the spindle speed is 0.449, greater than 0.05. Thus, the spindle speed is not a major factor and has no impact on the delamination at entry. The adjusted  $R^2$  which intends to approximate the actual percentage of explained variance is slightly negative and very close to zero; -0.0123. It's an indication that the regressor in this model, spindle speed, is useless. More terms should be added to the model for better interpretation.

Table 3.36: Parameter estimates and ANOVA results for the delamination at entry against spindle speed

Source	SS	df	MS	Number of obs = 35		
Model	.013620254	1	.013620254	F( 1, 33) =	0.59	
Residual	.764229654	33	.023158474	Prob > F =	0.4486	
				R-squared =	0.0175	
				Adj R-squared =	-0.0123	
Total	.777849908	34	.022877938	Root MSE =	.15218	
delentry	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
spindlespeed	-4.10e-06	5.35e-06	-0.77	0.449	-.000015	6.78e-06
_cons	1.19479	.0517618	23.08	0.000	1.08948	1.3001

### 3.4.3 Delamination at entry multiple linear regression analysis

This section presents the multiple linear regression analysis of the delamination at entry which is dependent on the feed rate and spindle speed. As a start, the parameter estimates and ANOVA tables are generated to detect the variables that have substantial effect on the delamination at

entry. **Table 3.37** indicates that the feed rate is the significant variable on the delamination at entry with a p-value equal to 0 lower than 0.05, in despite to the spindle speed whereas its p-value is at 0.067, greater than 0.05. Accordingly, the delamination at entry is affected significantly by the feed rate. However, the spindle speed variation has no significant effect on the delamination at entry. The adjusted  $R^2$  for delamination at entry using MLR is 83.45%. Based on this indicator only, the model fits well the data because a good amount of the variability is explained.

Table 3.37: Parameter estimates and ANOVA results for delamination at entry using MLR

Source	SS	df	MS	Number of obs = 35		
Model	.656699248	2	.328349624	F( 2, 32) =	86.73	
Residual	.12115066	32	.003785958	Prob > F =	0.0000	
				R-squared =	0.8442	
				Adj R-squared =	0.8345	
Total	.777849908	34	.022877938	Root MSE =	.06153	

delentry	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0004903	.0000376	13.03	0.000	.0004136	.0005669
spindlespeed	-4.10e-06	2.16e-06	-1.90	0.067	-8.50e-06	3.03e-07
_cons	1.042106	.0239845	43.45	0.000	.9932513	1.090961

From **table 3.37**, the MLR model of the delamination at entry is given by:

$$\text{Den} = 1.042 + 0.00049 \cdot F - 4.1 \times 10^{-6} \cdot SS$$

### **Interpretation of the coefficients**

1. Constant ( $\beta_0$ ): when feed rate and spindle are zero, the expected mean of delamination at entry is 1.042.
2. Feed rate ( $\beta_1$ ): for each one unit change of feed rate, delamination at entry will increase by 0.00049 when the spindle speed is constant.
3. Spindle Speed ( $\beta_2$ ): for each one unit change of spindle speed, delamination at entry will decrease by  $4.10 \times 10^{-6}$  when the feed rate is constant.

**Figure 3.36** shows the estimated plot of the delamination at entry using the multiple regression technique. The relationship between the independent variables and the dependent is linear. The predicted values are within the 95% interval of confidence and no outliers are identified. **Table C.9 in annex C** presents the observed and predicted values of this delamination at entry model.

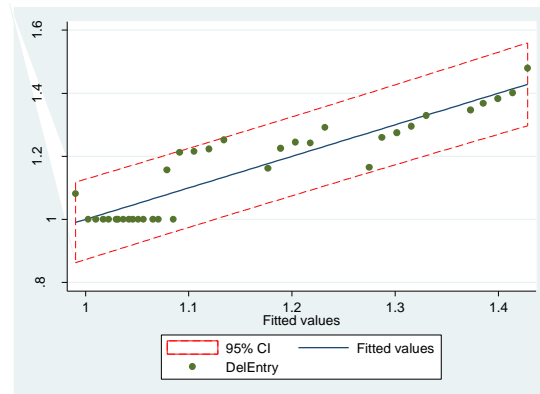


Figure 3.36: Delamination at entry observed values against predicted using MLR

To be able to use this model, the same four assumptions shall be validated.

**Independency assumption:** Figure 3.37 is a graphical representation of the standardized residuals of the fitted delamination at entry against the run order of the data collection of the delamination at entry data. As observed, the independence assumption is valid because the residuals do not follow a structured pattern.

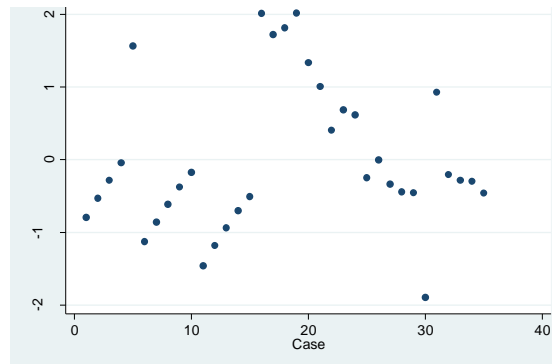


Figure 3.37: Plot of the standardized residuals of the fitted delamination at entry against the run order of the data collection

**Normality assumption:** Figure 3.38 represents the normality fit of the standardized residuals of the delamination at entry. The curved line indicates a non-normal distribution. It has a sharp lower and upper bound that differs from a bell shaped curve. The Shapiro-Wilk test is completed **table 3.38**. The calculated p-value is 0.02320, lower than 0.05. Therefore, the null hypothesis that the data does follow a normal distribution is rejected. This test confirms the graphical results.

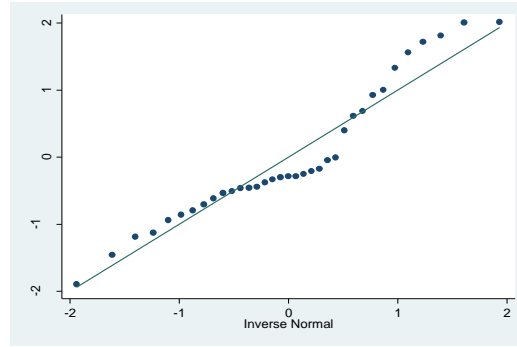


Figure 3.38: Normality plot for the delamination at entry

Table 3.38: Shapiro-Wilk test for normal data for delamination at entry

Variable	Obs	W	V	z	Prob>z
rdelentry	35	0.92725	2.597	1.992	0.02320

**Homoscedasticity assumption:** Figure 3.39 displays that the variance of the residual is constant with an undefined pattern. The BPCW test is completed to confirm this result. The calculated square-chi value is equal to 1.18 and the p-value is 0.2768 is greater than 0.05. In this case, the variance is constant and the assumption is validated.

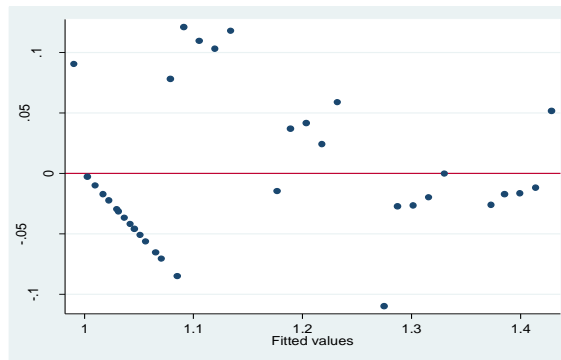


Figure 3.39: Predicted vs standardized residual values plot of the delamination at entry

**Linearity assumption:** To validate the linearity assumption, the scatterplot of each independent variable against the delamination at entry is presented. Figure 3.40 confirms that the relationship is linear between each independent variable and the delamination at entry. As a result, the linearity assumption is validated.

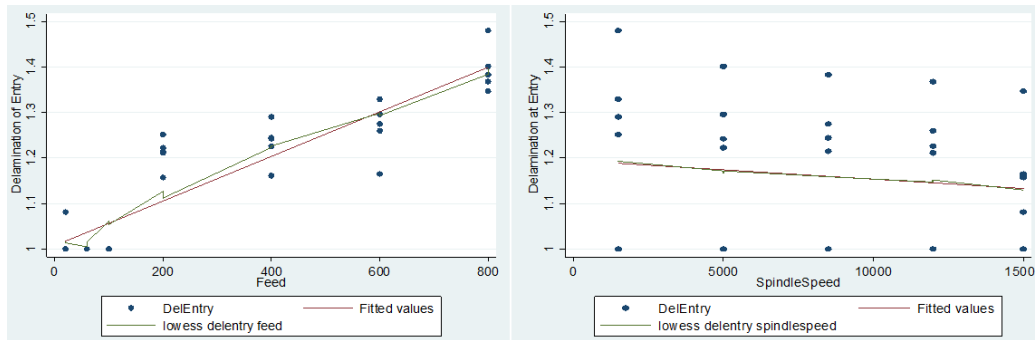


Figure 3.40: Scatterplots feed rate and spindle speed vs delamination at entry

**Table 3.39** describes the validation results of the assumptions.

Table 3.39: Illustration of the assumption validation results

Assumptions	Valid for this model
Independency	Yes
Normality	No
Homoscedasticity	Yes
Linearity	Yes

The normality assumption is violated subsequently; the fitted model generated may be used for prediction purposes. However, it may not be the ideal model; the next section will explore the interaction effect.

### 3.4.4 Delamination at entry multiple linear regression with interaction

In research, studying multiple effects rather than only the isolated effects of the single variables may contribute considerably to the process performance. As a consequence, in this part of the study, the regression analysis of the delamination at entry is performed by adding the interaction effect of the feed rate and the spindle speed then, a prediction equation is constructed. The parameter estimates ANOVA and coefficient of determination results are computed in **table 3.40**. The ANOVA outcomes indicates that the feed rate (p-value = 0), and the interaction of the feed rate and spindle speed (p-value = 0.033) are significant; their p-values are both lower than 0.05. The p-value of the spindle speed corresponds to 0.741, greater than 0.05, which designates that this variable has minimum effect on the delamination at entry. However, the presence of a significant interaction indicates that the effect of the feed rate on the delamination at entry depends on the values of the spindle speed. The adjusted  $R^2$  corresponds to 85.28% which is considered to be a good estimate. Adding an interaction term to this regression model can greatly

expand the understanding of the relationships among the variables in the model. In fact, this model provided 1.83% more explained variability than the model without the interaction.

Table 3.40: Parameter estimates and ANOVA results for delamination at entry using MLR-interaction

Source	SS	df	MS	Number of obs = 35		
Model	.673454409	3	.224484803	F( 3, 31) = 66.66		
Residual	.104395499	31	.003367597	Prob > F = 0.0000		
				R-squared = 0.8658		
				Adj R-squared = 0.8528		
				Root MSE = .05803		
delentry	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0006285	.0000714	8.80	0.000	.0004829	.0007741
spindlespeed	1.02e-06	3.07e-06	0.33	0.741	-5.24e-06	7.29e-06
fds	-1.65e-08	7.38e-09	-2.23	0.033	-3.15e-08	-1.41e-09
_cons	.9990699	.0297312	33.60	0.000	.9384328	1.059707

From **table 3.40**, the estimated regression model with interaction for the delamination at entry is:

$$\text{Den} = 1 + 0.00063 * F + 1.02 \times 10^{-6} * SS - 1.7 \times 10^{-8} * F * SS$$

**Figure 3.41** shows the estimated regression line of the delamination at entry. This model fits well the data and the relationship is linear. **Table C.10 in annex C** presents the data of this plot.

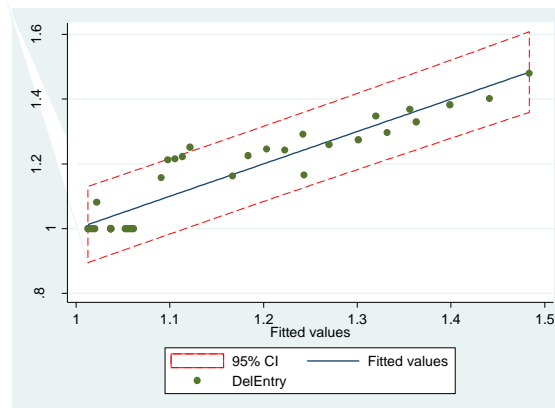


Figure 3.41: Delamination at entry observed against predicted values using MLR-interaction

To be able to use this fitted model, the following assumptions shall be validated before using the fitted model: normality and homoscedasticity.

**Normality assumption:** **Figure 3.42** indicates that the data does not follow a normal distribution. The lower and upper tails are not following the straight line and many outliers are identified. The Shapiro-Wilk test is performed to check the adequacy of the results observed by

the graphical illustration. The computed p-value is 0.00432 (from **table 3.41**), lower than 0.05. Thus, the null hypothesis is rejected, and the data does not follow a normal distribution.

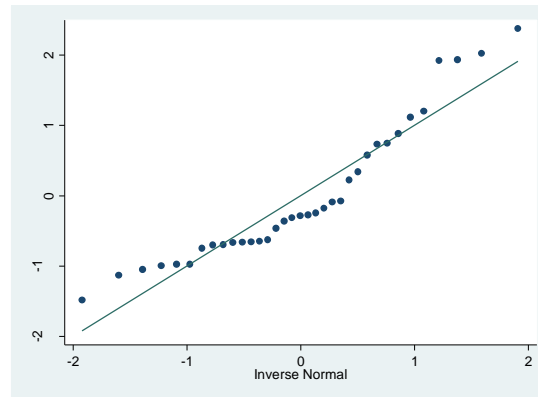


Figure 3.42: Normality plot for delamination at entry using MLR-interaction

Table 3.41: Shapiro-Wilk test for normal data for delamination at entry

Variable	Obs	W	V	z	Prob>z
rdelentryfds	35	0.90143	3.518	2.626	0.00432

**Homoscedasticity assumption:** This assumption aims to verify that the dependent variable exhibits similar amounts of variance across the range of fitted values. This assumption is checked by visual examination of a plot of the standardized residuals against the fitted values. **Figure 3.43** displays no specific pattern and the values are fluctuating randomly around zero. To confirm this result, the BPCW test is completed. The calculated square-chi value is equal to 1.43 and the p-value is at 0.2324, greater than 0.05. In this case, the null hypothesis is accepted and the variance is homogeneous.

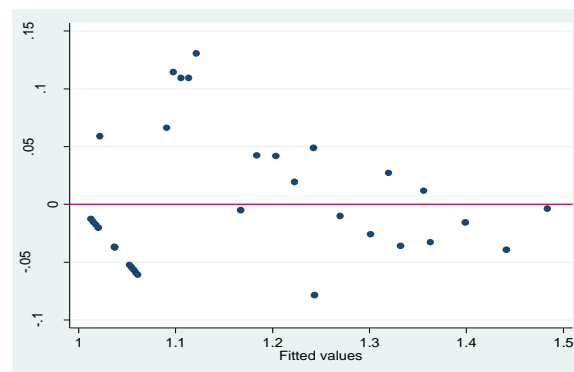


Figure 3.43: Residual plot versus fitted data of delamination at entry using MLR-interaction

**Table 3.42** describes the validation results of the assumptions. This model may be used. However, a transformation and a non-linear regression fitting will be attempted in the next section to check if the normality issue can be fixed.

Table 3.42: Illustration of the assumption validation results of delamination at entry using MLR-interaction

Assumptions	Valid for this model
Independency	Yes
Normality	No
Homoscedasticity	Yes
Linearity	Yes

### 3.4.5 Delamination at entry nonlinear regression

The main objective of this section is to predict delamination at entry during drilling of composites. In the previous sections, it hasn't been possible to outline a satisfactory estimated multiple regression model. In fact, in case of the assumptions are violated, the results of the analysis are incorrect and misleading which makes the model unusable. As introduced at the beginning of this study, if the multiple regression methodology can't be used, the nonlinear regression technique will be developed to fit a model. Researchers are looking for a model that is not linear in its parameters. Finding the function that provides the best fit of the data under investigation and the type of nonlinear regression that coexists between the predictor variable and the response variable is a challenge. It requires a good knowledge of the process, a validation of the nonlinear regression assumptions and trial and error analyses. The software helps computing different type of transformation in order to find the most suitable one: exponential, logarithmic, trigonometric, sigmoidal curves, etc. The transformation with a p-value greater than 0.05, indicates that the hypothesis that the data follows a normal distribution cannot be rejected. From **table 3.43**, this method suggests two transformations with a p-value greater than 0.05:

- Cubic with p-value equals to 0.227; and
- Square with p-value equals to 0.123.

In this case, the cubic model is the best option, whereas its p-value is the highest one to be over 0.05. In the following analysis, the cubic transformation is verified to determine the delamination at entry fitted model.

Table 3.43: Goodness of fit of the delamination at entry transformations

Transformation	formula	chi2(2)	P(chi2)
cubic	delentry^3	2.97	0.227
square	delentry^2	4.20	0.123
identity	delentry	7.18	0.028
square root	sqrt(delentry)	9.36	0.009
log	log(delentry)	11.97	0.003
1/(square root)	1/sqrt(delentry)	14.97	0.001
inverse	1/delentry	18.33	0.000
1/square	1/(delentry^2)	25.85	0.000
1/cubic	1/(delentry^3)	33.88	0.000

The ANOVA table below indicates that both feed rate and spindle speed are significant variables on the delamination at entry with respective p-values 0 and 0.017, lower than 0.05.

Table 3.44: Parameter estimates and ANOVA results of cubic transformation of the delamination at entry

Source	SS	df	MS	Number of obs = 35		
Model	12.0834386	2	6.04171929	F( 2, 32) = 102.66		
Residual	1.88320376	32	.058850117	Prob > F = 0.0000		
				R-squared = 0.8652		
				Adj R-squared = 0.8567		
Total	13.9666423	34	.410783598	Root MSE = .24259		
delentry3	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0020924	.0001483	14.11	0.000	.0017903	.0023945
spindlespeed	-.0000214	8.52e-06	-2.51	0.017	-.0000387	-4.01e-06
_cons	1.16837	.0945619	12.36	0.000	.9757539	1.360987

From **table 3.44**, the following transformation fitted model is determined:

$$Dee^3 = 1.01681 + 0.0021 * F - 0.0000214 * SS$$

**Table 3.44** indicates that the adjusted  $R^2$  is at 85.67%; a good estimation of the variability. The adjusted  $R^2$  of the MLR model with interaction developed previously was 85.28% compared to 85.67% with the cubic transformation. In fact, a little more variability is explained by this model. To select the model to use, the AIC of all the developed models are computed below.

Table 3.45: AIC of the transformation of the delamination at entry

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	35	-33.58601	1.47869	3	3.042621	7.708665

Table 3.46: AIC of the MLR model of the delamination at entry

Model	Obs	ll (null)	ll (model)	df	AIC	BIC
.	35	16.95212	49.49335	3	-92.9867	-88.32065

Table 3.47: AIC of the MLR model with interaction of the delamination at entry

Model	Obs	ll (null)	ll (model)	df	AIC	BIC
.	35	16.95212	52.09819	4	-96.19639	-89.975

Table 3.48: Critical values of the models developed for delamination at entry

Method	Adjusted R <sup>2</sup>	AIC
MLR	83.45%	-92.9867
MLR with interaction	85.28%	-96.19639
Non-linear regression	85.67%	3.042621

When comparing models among each other, several parameters can be used: the adjusted R<sup>2</sup> coefficient (the higher), the MSE value (the lowest) and the AIC (the lowest). Based on the AIC value from the tables above, the cubic transformation model should not be retained, as it did not improve the multiple linear regression models. The AIC of the multiple linear regression models with and without interaction are very similar. The adjusted R<sup>2</sup> of the MLR with interaction (85.28%) is slightly better than the MLR without interaction (83.45%). In this case, the suggested model to use is the MLR with interaction:

$$\text{Den} = 1 + 0.00063 \cdot F + 1.02 \times 10^{-6} \cdot SS - 1.7 \times 10^{-8} \cdot F \cdot SS$$

Normality is not needed to fit the regression line. Therefore, even if this model doesn't follow a normal distribution, it still can be used for prediction purposes because the model is strongly linear and the variance behavior is constant. The prediction accuracy whereas measured by mean square error, is little affected by the normality assumption. Below is a table comparing all the methods used to analyze the delamination at entry.

Table 3.49: Comparison of all methods used for delamination at entry analysis

Method used	Validity to use	Variability	Significant variables
MLR	Non-normality	83.45%	Feed rate
MLR with interaction	Non-normality	85.28%	Feed rate and interaction
Cubic Transformation	AIC too high	85.67%	Feed rate and spindle speed

### 3.5 Delamination at exit analysis

#### 3.5.1 Delamination at exit distribution study over feed rate

In this section, the feed rate is analyzed as the only effect on the delamination at exit. **Table 3.50** indicates that the p-value of the feed rate is near 0, lower than 0.05. Consequently, the feed rate has a significant impact on the delamination at entry behavior. Only 55.17% of the variability is explained by the feed rate effect.

Table 3.50: Parameter estimates and ANOVA results for the delamination at exit against feed rate

Source	SS	df	MS	Number of obs = 35		
Model	.077817093	1	.077817093	F( 1, 33) =	42.84	
Residual	.059937636	33	.001816292	Prob > F =	0.0000	
				R-squared =	0.5649	
				Adj R-squared =	0.5517	
Total	.137754729	34	.00405161	Root MSE =	.04262	

delexit	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0001705	.0000261	6.55	0.000	.0001175	.0002236
_cons	1.00043	.0108507	92.20	0.000	.9783542	1.022506

To understand the relationship between the feed rate and delamination at exit, the boxplot below is drawn.

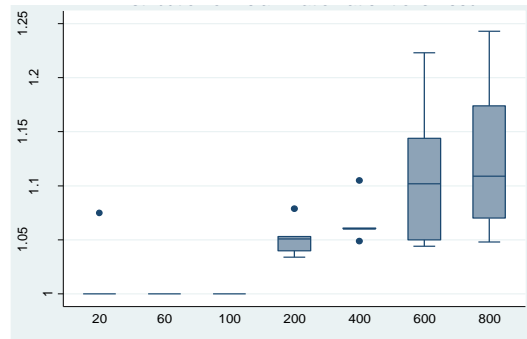


Figure 3.44: Boxplot of delamination at exit against feed rate

**Figure 3.44** displays that the delamination at exit, similar to delamination at entry, tends to rise when the feed rate increases. This figure indicates that:

- The flat boxes at feed rate 20, 60 and 100 indicate that all values of delamination at exit are equal to 1,
- The observation at feed rate equal to 20 whereas the delamination at exit is at 1.075 is away from the data but is not identified as an outlier. In fact, the delamination at this low

feed is high due to potential rubbing between the tool and the work piece and thermal softening at high speed and low feed [32], and

- There are three more outliers found at feed rate 200 and 400. The variability at feed rate 600 and 800 is very high compared to the other boxes.

Thus, a substantial amount of outliers is explored: 11.4% (4/35) of the total number of observations. The existence of extreme points in this data may affect the model variability and make the model unusable. To evaluate the effect of the feed rate on the delamination at exit, the parameter estimates, ANOVA and coefficient of determination results are generated.

### 3.5.2 Delamination at exit distribution study over spindle speed

In this section, the spindle speed is analyzed as the only effect on the delamination at exit. To detect the effect of the spindle speed on the delamination at exit, the parameter estimates, ANOVA and coefficient of determination results are computed in **table 3.51**. The p-value of the spindle speed equals 0.055, greater than 0.05, which means that the spindle speed is not a major factor and has no impact on the delamination at exit.

Table 3.51: Parameter estimates and ANOVA results for the delamination at exit against spindle speed

Source	SS	df	MS	Number of obs = 35		
Model	.014765352	1	.014765352	F( 1, 33) =	3.96	
Residual	.122989377	33	.003726951	Prob > F =	0.0549	
				R-squared =	0.1072	
				Adj R-squared =	0.0801	
Total	.137754729	34	.00405161	Root MSE =	.06105	
delexit	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
spindlespeed	-4.27e-06	2.15e-06	-1.99	0.055	-8.63e-06	9.46e-08
_cons	1.089409	.020765	52.46	0.000	1.047162	1.131656

**Figure 3.45** displays the relationship between the spindle speed and delamination at exit. It indicates a minor decrease of the delamination at entry with the increase of spindle speed. Therefore, the delamination at entry is inversely proportional to the spindle speed. No outliers are detected. However, the variance at each spindle speed range is high which suggests that the data is spread out and not concentrated at similar values. However the higher the spindle speed is, the smallest the variability is.

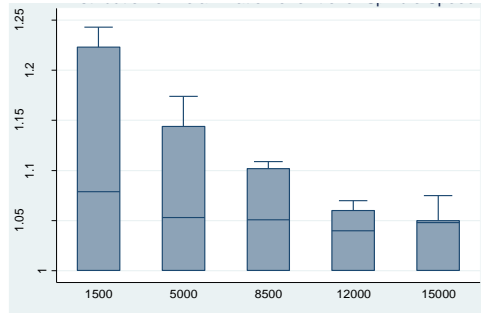


Figure 3.45: Boxplot of delamination at exit against spindle speed

### 3.5.3 Delamination at exit multiple regression analysis

This section presents the multiple regression analysis for the delamination at exit which is dependent on the feed rate and spindle speed. As a start, the parameter estimates, ANOVA and coefficient of determination results are computed in **table 3.52** to detect the factors that have substantial impacts on the delamination at exit. It indicates that the feed rate is a significant variable on the delamination at exit with a p-value equal to 0 which is lower than 0.05. The spindle speed has a p-value of 0.003, lower than 0.05. Accordingly, the delamination at entry is affected significantly by the feed rate and the spindle speed. The adjusted  $R^2$  of the MLR for the delamination at exit is 0.6516. Based on this indicator only, the model may fit the data. However, this value is considered to be reasonably low; 34.84% of the variability is not explained.

Table 3.52: Parameter estimates and ANOVA results for delamination at exit using MLR

Source	SS	df	MS	Number of obs = 35		
Model	.092582445	2	.046291223	F( 2, 32) = 32.79		
Residual	.045172284	32	.001411634	Prob > F = 0.0000		
Total	.137754729	34	.00405161	R-squared = 0.6721		
				Adj R-squared = 0.6516		
				Root MSE = .03757		
delexit	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0001705	.000023	7.42	0.000	.0001238	.0002173
spindlespeed	-4.27e-06	1.32e-06	-3.23	0.003	-6.96e-06	-1.58e-06
_cons	1.036296	.0146455	70.76	0.000	1.006464	1.066128

From **table 3.52**, the estimated regression model of the Delamination at exit is given by:

$$\text{Dex} = 1.036 + 0.00017 \cdot F - 4.27 \times 10^{-6} \cdot \text{SS}$$

#### Interpretation of the coefficients

1. Constant ( $\beta_0$ ): when feed and spindle are zero, the expected mean of delamination at exit is 1.036.

2. Feed ( $\beta_1$ ): for each one unit change of feed, delamination at exit will increase by 0.00017 when the spindle speed is held constant.
3. Spindle Speed ( $\beta_2$ ): for each one unit change of spindle speed, delamination at exit will decrease by  $4.27 \times 10^{-6}$  when the feed rate is held constant.

**Figure 3.46** shows the estimated plot of the delamination at exit using the MLR technique. The grey area represents the 95% interval of confidence. At least three points are outside the interval of confidence which indicates that this model may not be a good fit. **Table C.11 in annex C** displays the values of this plot.

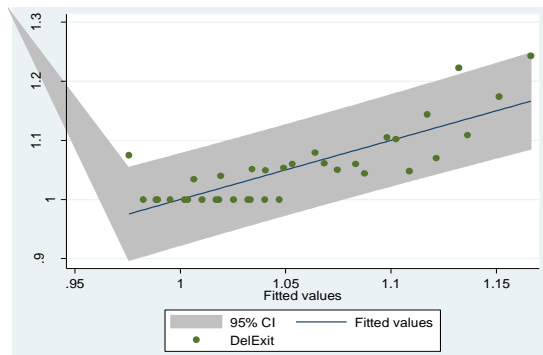


Figure 3.46: Delamination at exit estimated plot using MLR

To be able to use this model, the four assumptions shall be verified.

**Independency assumption:** As observed on **figure 3.47**, the independence assumption is validated because the residuals do not follow a structured pattern.

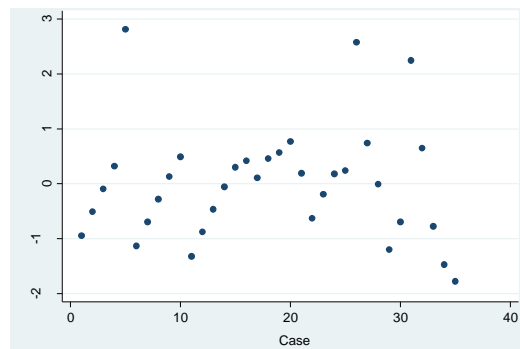


Figure 3.47: Plot of the standardized residuals of the fitted delamination at exit against the run order of the data collection

**Normality assumption:** The plot in **figure 3.48** suggests a slight deviation of the straight line at the upper tail due to outliers and influential observations. Shapiro-Wilk test is performed and it

indicates that the calculated p-value is 0.01409 (**table 3.53**), lower than 0.05. Therefore, the null hypothesis that the data does follow a normal distribution is rejected which endorses the graphical illustration result.

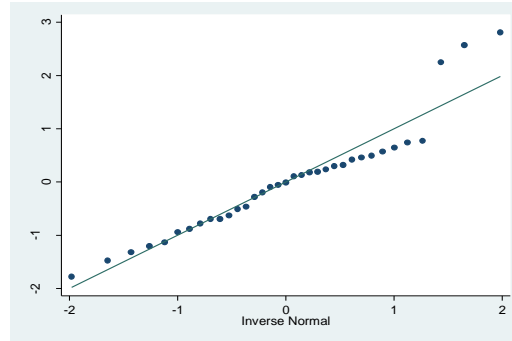


Figure 3.48: Normality plot for the delamination at exit with MLR

Table 3.53: Shapiro-Wilk test for delamination at exit using MLR

Variable	Obs	W	V	z	Prob>z
rdelexit	35	0.91982	2.862	2.195	0.01409

**Homoscedasticity assumption:** Figure 3.49 displays that the variance of the residual is constant; no specific pattern is identified. To verify this interpretation, the BPCW test is completed. The computed square-chi value is equal to 2.98 and the p-value equal to 0.0844, greater than 0.05. In this case, the null hypothesis is accepted and the variance is constant.

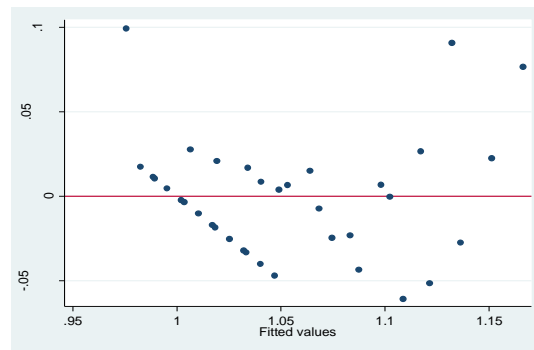


Figure 3.49: Predicted vs standardized residual values plot of the delamination at exit

**Linearity assumption:** To validate the linearity assumption, the scatterplots in **figure 3.50** of each independent variable against the delamination at entry is presented. The scatterplots below indicate that the relationship is linear between the feed rate and the delamination at exit and is

moderately linear between the spindle speed and the delamination at exit; the shape is slightly bell-shaped. Both lowess outlines overlap a linear fit.

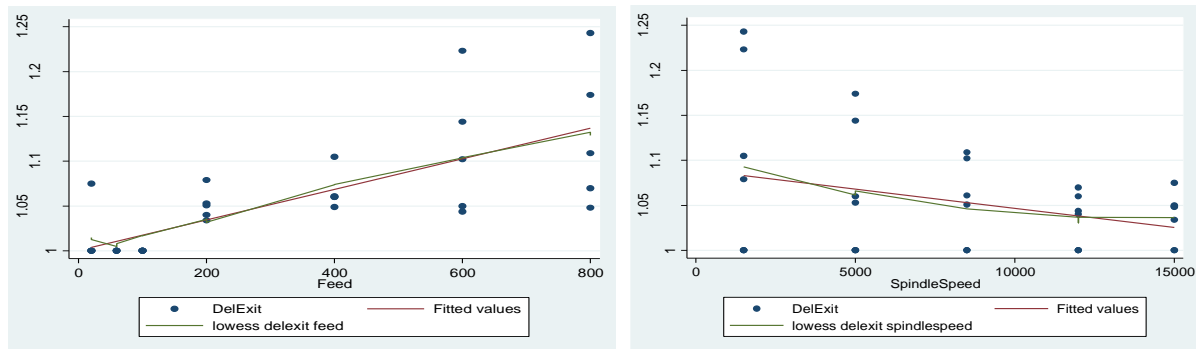


Figure 3.50: Scatterplot feed rate and spindle speed vs delamination at exit

**Table 3.54** describes the validation results of the assumptions using the multiple regression technique.

Table 3.54: Illustration of the assumption validation results of the delamination at exit using MLR

Assumptions	Valid for this model
Independency	Yes
Normality	No
Homoscedasticity	Yes
Linearity	Yes

The normality assumption is violated. Therefore, the fitted model generated may be used for prediction purposes but the MLR with interaction is attempted for better variability estimation.

### 3.5.4 Delamination at exit multiple linear regression analysis with interaction

This section presents the multiple regression analysis for the delamination at exit with the interaction between the feed rate and the spindle speed. As a start, the parameter estimates, ANOVA and coefficient of determination results are generated in **table 3.55**. It indicates that the feed rate, spindle speed and their interaction effect are significant on the delamination at exit. Their p-values are respectively 0, 0.016 and 0; lower than 0.05. Accordingly, the delamination at entry is affected significantly by the feed rate, the spindle speed and their interaction. The adjusted  $R^2$  of the MLR with interaction model corresponds to 89.44%. A considerable amount of variability is explained by this model. The adjusted  $R^2$  of the previous model was 65.16%. The current model is offering 24.28% better explained variability by adding the interaction effect.

Table 3.55: Parameter estimates and ANOVA results for delamination at exit using MLR-interaction

Source	SS	df	MS	Number of obs = 35		
Model	.124486724	3	.041495575	F( 3, 31) =	96.95	
Residual	.013268005	31	.000428	Prob > F =	0.0000	
				R-squared =	0.9037	
				Adj R-squared =	0.8944	
Total	.137754729	34	.00405161	Root MSE =	.02069	

delexit	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0003612	.0000255	14.19	0.000	.0003093	.0004131
spindlespeed	2.80e-06	1.09e-06	2.56	0.016	5.67e-07	5.03e-06
fds	-2.27e-08	2.63e-09	-8.63	0.000	-2.81e-08	-1.73e-08
_cons	.9769103	.0105992	92.17	0.000	.955293	.9985275

From **table 3.55**, the estimated regression model with interaction of the delamination at exit is given by:

$$\text{Dex} = 0.98 + 0.00036 * F + 2.8 \times 10^{-6} * SS - 2.3 \times 10^{-8} * F * SS$$

**Figure 3.51** shows the estimated plot of the delamination at exit using the MLR with interaction. **Table C.12 in annex C** displays the data of this plot. To be able to use this model, the following assumptions shall be validated: normality and homoscedasticity.

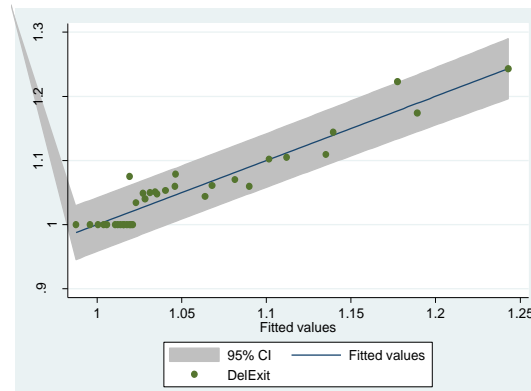


Figure 3.51: Delamination at exit observed values against predicted using MLR-interaction

**Normality assumption:** **Figure 3.52** represents the normality fit of the standardized residuals of the delamination at exit experimental data. This plot suggests a slight deviation of the straight line at the lower and upper tail due to outliers and influential observations. To confirm these results, the Shapiro-Wilk test is computed. The calculated p-value is 0.02383 (**table 3.56**), lower than 0.05. Hence, the null hypothesis that the data does follow a normal distribution is rejected.

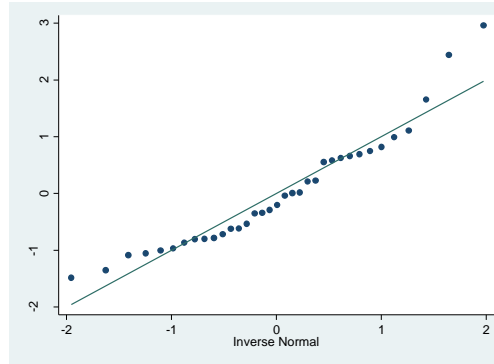


Figure 3.52: Normality plot for the delamination at exit with MLR-interaction

Table 3.56: Shapiro-Wilk test for normal data for delamination at exit MLR-interaction

Variable	Obs	W	V	z	Prob>z
rdelexitfds	35	0.92765	2.582	1.980	0.02383

**Homoscedasticity assumption:** Figure 3.53 presents the standardized residuals of the fitted delamination at exit. It displays that the variance of the residual is constant because no pattern is identified. To verify that the plot is interpreted correctly, the BPCW test is completed. The computed square-chi value is equal to 0.51 and the p-value is at 0.4739, greater than 0.05. In this case, the variance is declared constant.

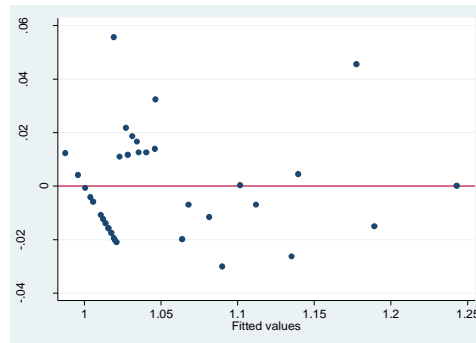


Figure 3.53: Predicted vs standardized residual values plot of the delamination at exit using MLR-interaction

**Table 3.57** describes the validation results of the assumptions. One assumption is violated. In fact, the lack of normality for prediction purposes in regression (to get unbiased estimates of the regression coefficients) may not be as important as the homoscedasticity and linearity [33]; whereas the normality of residuals is mandatory to validate testing hypothesis. For further investigation, a transformation of the dependent variable is attempted. In **table 3.59**, no

transformation is suggested because no p-value is found to be greater than 0.05. It may be due to the presence of outliers.

Table 3.57: Illustration of the assumption validation results for delamination at exit MLR-interaction

Assumptions	Valid for this model
Independency	Yes
Normality	No
Homoscedasticity	Yes
Linearity	Yes

Table 3.58: Goodness of fit of the delamination at exit transformations

Transformation	formula	chi2 (2)	P(chi2)
cubic	delexit^3	15.58	0.000
square	delexit^2	13.77	0.001
identity	delexit	11.99	0.002
square root	sqrt(delexit)	11.11	0.004
log	log(delexit)	10.25	0.006
1/(square root)	1/sqrt(delexit)	9.41	0.009
inverse	1/delexit	8.60	0.014
1/square	1/(delexit^2)	7.08	0.029
1/cubic	1/(delexit^3)	5.74	0.057

Table 3.59: AIC of the MLR model of the delamination at exit

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	35	47.24565	66.75799	3	-127.516	-122.8499

Table 3.60: AIC of the MLR model with interaction of the delamination at exit

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	35	47.24565	88.19774	4	-168.3955	-162.1741

Based on the AIC value from the tables above, the MLR model with interaction is suggested:

$$\text{Dex} = 0.98 + 0.00036 \cdot F + 2.8 \times 10^{-6} \cdot \text{SS} - 2.3 \times 10^{-8} \cdot F \cdot \text{SS}$$

Table 3.61 presents all the methods developed to analyze the delamination at exit fitting model.

Table 3.61: Comparison of all methods used for delamination at exit analysis

Method used	Validity to use	Variability	Significant variables
MLR	Non-normality	65.16%	Feed rate and spindle speed
MLR with interaction	Non-normality	89.44%	Feed rate and interaction

### 3.6 Surface roughness analysis

#### 3.6.1 Surface roughness distribution study over feed rate

In this section, the feed rate is analyzed as the only effect on the surface roughness. **Figure 3.54** displays the relationship between the feed rate and the surface roughness. This boxplot indicates a very large variability when the feed rate is at 20. At the other feed rate levels, the variability is acceptable and relatively small. It also displays a moderately linear fit and identifies an outlier when the feed rate is at 600. To evaluate the effect of the feed rate on the surface roughness, the parameter estimates, ANOVA and coefficient of determination results are generated in **table 3.62**.

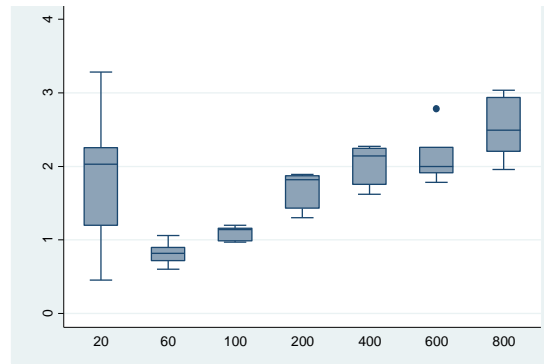


Figure 3.54: Boxplot of surface roughness against feed rate

Table 3.62: Parameter estimates and ANOVA results of the surface roughness over feed rate

Source	SS	df	MS	Number of obs = 35		
Model	7.0859538	1	7.0859538	F( 1, 33) = 22.24		
Residual	10.5123218	33	.318555206	Prob > F = 0.0000		
				R-squared = 0.4027		
				Adj R-squared = 0.3845		
Total	17.5982756	34	.517596341	Root MSE = .56441		
holesurfro~h	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0016274	.0003451	4.72	0.000	.0009254	.0023295
_cons	1.221086	.1436997	8.50	0.000	.9287266	1.513445

The computed p-value of the feed rate is near 0, lower than 0.05 which indicates that the feed rate has a significant effect on the surface roughness. Only 38.45% of the variability is explained by the feed rate.

### 3.6.2 Surface roughness distribution study over spindle speed

In this section, the spindle speed is analyzed as the only effect on the surface roughness. **Figure 3.55** displays the relationship between them. It demonstrates that the variability at all spindle speed level is very large; the data is very spread and not concentrated around a specific value.

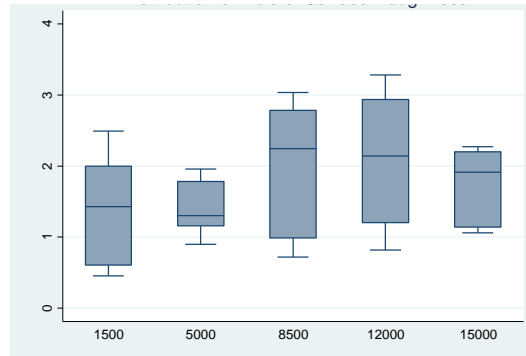


Figure 3.55: Boxplot of surface roughness against spindle speed

To detect the effect of the spindle speed on the surface roughness, the parameter estimates, ANOVA and coefficient of determination results are computed in **table 3.63**. The calculated p-value of the spindle speed equals 0.080, greater than 0.05, which indicates that the spindle speed is not a major factor and has no impact on the surface roughness. Only 6.27 % of the variability is explained by the spindle speed.

Table 3.63: Parameter estimates and ANOVA results of the surface roughness over spindle speed

Source	SS	df	MS	Number of obs = 35		
Model	1.58774913	1	1.58774913	F( 1, 33) =	3.27	
Residual	16.0105265	33	.485167469	Prob > F =	0.0796	
				R-squared =	0.0902	
				Adj R-squared =	0.0627	
Total	17.5982756	34	.517596341	Root MSE =	.69654	
holesurfro~h	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
spindlespeed	.0000443	.0000245	1.81	0.080	-5.52e-06	.0000941
_cons	1.35599	.236919	5.72	0.000	.8739746	1.838005

### 3.6.3 Surface roughness multiple linear regression analysis

This section presents the MLR analysis for the surface roughness which is dependent on the feed rate and spindle speed. The parameter estimates, ANOVA and coefficient of determination results are computed in **table 3.64**, to detect the variables that have substantial effect on the output and to

evaluate the known variability. This table indicates that the feed rate and spindle speed are both statistically significant with respectively a p-value equal to 0 and 0.023; both lower than 0.05. Accordingly, the surface roughness is affected significantly by the feed rate and spindle speed. The adjusted  $R^2$  of the MLR for the surface roughness is 46.12%. This indicator specifies that the model does not fit the data properly. In fact, this value is considered to be low; 53.88% variability is not explained by the feed rate and spindle speed.

Table 3.64: Parameter estimates and ANOVA results for surface roughness using MLR

Source	SS	df	MS	Number of obs = 35		
Model	8.67370293	2	4.33685146	F( 2, 32) =	15.55	
Residual	8.92457268	32	.278892896	Prob > F =	0.0000	
				R-squared =	0.4929	
				Adj R-squared =	0.4612	
Total	17.5982756	34	.517596341	Root MSE =	.5281	

holesurfro-h	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0016274	.0003229	5.04	0.000	.0009698	.0022851
spindlespeed	.0000443	.0000186	2.39	0.023	6.48e-06	.0000821
_cons	.8491616	.205855	4.13	0.000	.4298488	1.268474

From **Table 3.64**, the estimated MLR model of the surface roughness is given by:

$$SR = 0.85 + 0.0016 * F + 0.000044 * SS$$

### **Interpretation of the coefficients**

1. Constant ( $\beta_1$ ): when feed rate and spindle are equal to zero, the expected mean of surface roughness is 0.85.
2. Feed rate ( $\beta_2$ ): for each one unit change of feed, the surface roughness will increase by 0.0016 when the spindle speed is constant.
3. Spindle Speed ( $\beta_3$ ): for each one unit change of Spindle Speed, the surface roughness will increase by 0.000044 when the feed rate is constant.

**Figure 3.56** shows the estimated plot of the surface roughness using the multiple regression technique. **Table C.13 in annex C** shows the observed and predicted surface roughness values.

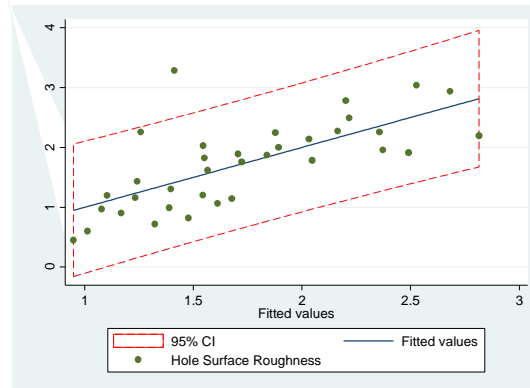


Figure 3.56: Surface roughness observed values against predicted using MLR

To be able to use this model, the following assumptions must be validated: independency, linearity, normality and homoscedasticity.

**Independency assumption:** Figure 3.57 is a graphical representation of the standardized residuals of the fitted surface roughness against the run order of the data collection of the surface roughness data. As observed, the independence assumption is valid because the residuals do not follow a structured pattern.

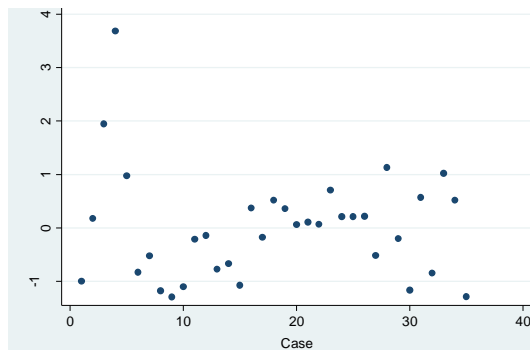


Figure 3.57: Plot of the standardized residuals of the fitted surface roughness against the run order of the data collection

**Normality assumption:** Figure 3.58 displays the normality fit of the standardized residuals of the surface roughness experimental data. This plot suggests a slight deviation of the straight line at the lower and upper tails due to outliers' observations. To verify this geometric observation, the Shapiro-Wilk test is performed. Table 3.65 indicates that the p-value is 0.00197, lower than 0.05. Thus, the null hypothesis that the data does follow a normal distribution is rejected. This test confirms the results seen on the plot.

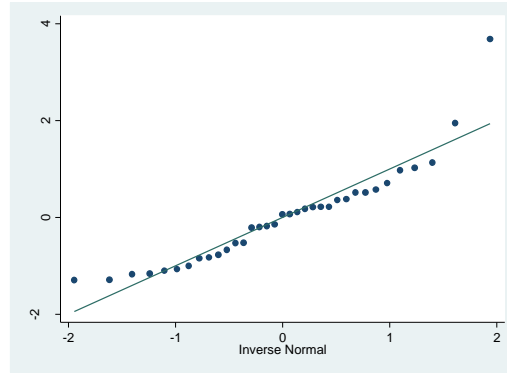


Figure 3.58: Normality plot for the surface roughness experimental data

Table 3.65: Shapiro-Wilk test for normal data for surface roughness

Variable	Obs	W	V	z	Prob>z
rhsr	35	0.88851	3.979	2.883	0.00197

**Homoscedasticity assumption:** Figure 3.59 illustrates the standardized residuals of the surface roughness data against the fitted ones. It displays that the variance of the residual is constant and no pattern is identified.

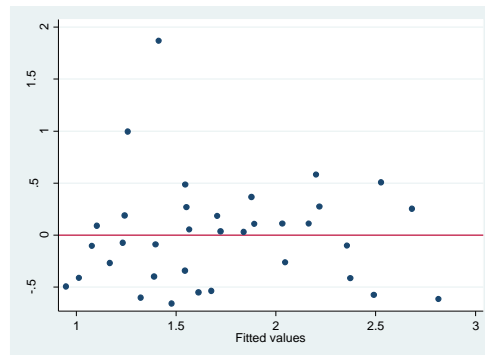


Figure 3.59: Predicted vs standardized residual values plot of the delamination at exit

To verify that the plot is providing the right conclusion, the BPCW test is completed. The calculated square-chi value is equal to 1.05 and the p-value to 0.3053, greater than 0.05. Consequently, this test demonstrates that the variance is constant.

**Linearity assumption:** To validate the linearity assumption, the scatterplot of each independent variable against the surface roughness is presented in **figure 3.60**. It confirms that the relationship is linear for both the feed rate and the spindle speed against the surface roughness.

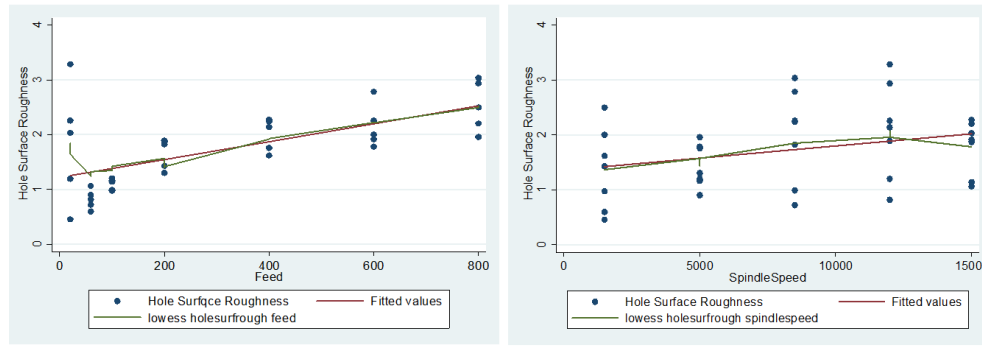


Figure 3.60: Scatterplots feed rate and spindle speed against surface roughness

**Table 3.66** describes the validation results of the assumptions. The normality assumption is violated. Hence, the fitted model generated may be used for prediction purposes. However, the MLR with interaction will be developed in search of better variability estimation.

Table 3.66: Illustration of the assumption validation results for the surface roughness using MLR

Assumptions	Valid for this model
Independency	Yes
Normality	No
Homoscedasticity	Yes
Linearity	Yes

### 3.6.4 Surface roughness multiple linear regression analysis with interaction

This part of the study presents the multiple regression analysis for the surface roughness by adding the interaction effect between the feed rate and spindle speed.

As a start, the parameter estimates, ANOVA and coefficient of determination results are generated in **table 3.67** to detect the variables that have important effects on the surface roughness. It indicates that the both the feed rate and the spindle speed are respectively significant with a p-value equals to 0.001 and 0.016. However, this table reveals that the interaction between the feed rate and spindle speed is not statistically significant; its p-value corresponds to 0.210, greater than 0.05.

The adjusted  $R^2$  of the MLR with interaction is 0.4717. Only 47.17 % of the variability is explained by the feed rate, spindle speed and their interaction. In fact, adding the interaction fact improved the known variability by 1.057% compared to the previous model (46.12%).

Table 3.67: Parameter estimates and ANOVA results for surface roughness using MLR-interaction

Source	SS	df	MS	Number of obs = 35		
Model	9.12147138	3	3.04049046	F( 3, 31) = 11.12		
Residual	8.47680423	31	.273445298	Prob > F = 0.0000		
Total	17.5982756	34	.517596341	R-squared = 0.5183		
				Adj R-squared = 0.4717		
				Root MSE = .52292		
holesurfro~h	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0023418	.0006433	3.64	0.001	.0010297	.0036539
spindlespeed	.0000708	.0000277	2.56	0.016	.0000143	.0001272
fds	-8.50e-08	6.65e-08	-1.28	0.210	-2.21e-07	5.05e-08
_cons	.6266841	.2679088	2.34	0.026	.0802805	1.173088

From **table 3.67**, the estimated regression model with interaction of the surface roughness is given by:

$$SR = 0.63 + 0.00234 * F + 0.000071 * SS - 8.5 \times 10^{-8} * F * SS$$

**Figure 3.61** shows the estimated plot of the surface roughness using the MLR with interaction. **Table C.14 in annex C** displays the data of this estimated plot.

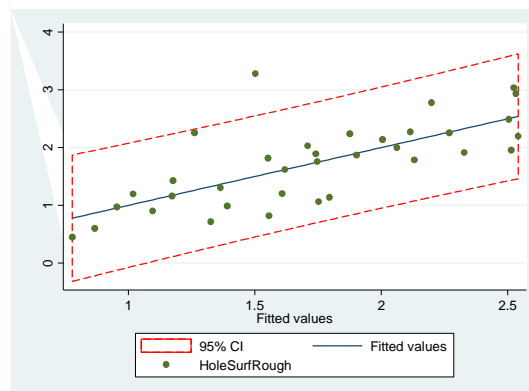


Figure 3.61: Surface roughness observed values against predicted using MLR-interaction

To be able to use this model, the following assumptions shall be validated: normality and homoscedasticity.

**Normality assumption:** **Figure 3.62** displays the normality plot of the standardized residuals of the delamination at exit experimental data. This plot suggests a slight deviation of the straight line at the lower and upper tail due to outlying and influential observations. The Shapiro-Wilk test is performed to confirm this observation. The computed p-value is 0.00630 (**table 3.68**), lower than 0.05. Hence, the null hypothesis that the data does follow a normal distribution is rejected.

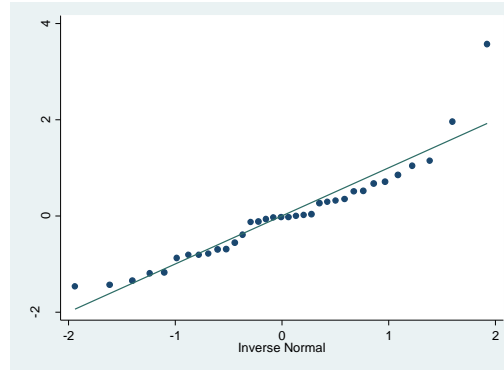


Figure 3.62: Normality plot of the surface roughness data using MLR-interaction

Table 3.68: Shapiro-Wilk test for surface roughness using MLR-interaction

Variable	Obs	W	V	z	Prob>z
rholesurfr~s	35	0.90742	3.304	2.495	0.00630

**Homoscedasticity assumption:** Figure 3.63 presents the standardized residuals of the surface roughness. It displays that the variance of the residual is constant. To confirm this interpretation, the BPCW test is completed. The computed square-chi value corresponds to 0.43 and the p-value is equal to 0.5105, greater than 0.05. In this case, the variance is homogenous.

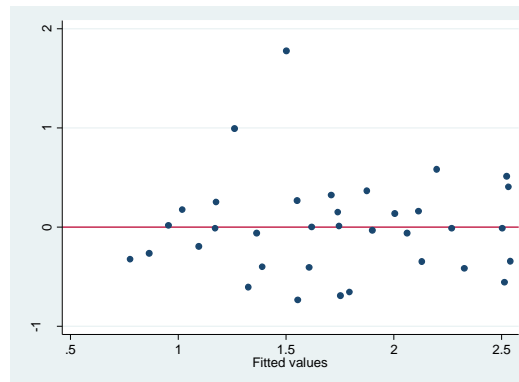


Figure 3.63: Predicted vs standardized residual values plot of the surface roughness using MLR-interaction

**Table 3.69** describes the validation results of the assumptions. The normality assumption is violated. Therefore, the fitted model generated may be used for prediction purposes. However, the correlation between the surface roughness and the independent variables (feed rate and spindle speed) is poor. In consequence, the inclusion of the interaction term did not improve the model drastically. The non-normality of the residuals is probably due to the presence of outliers as shown in the normality plot.

Table 3.69: Illustration of the assumption validation results for surface roughness with MLR-interaction

Assumptions	Valid for this model
Independency	Yes
Normality	No
Homoscedasticity	Yes
Linearity	Yes

To find a suitable model, the influence of the outliers must be investigated and a transformation of the surface roughness variable might be successful. In previous sections, a suitable prediction model couldn't be determined using the multiple regression method. The outlier observation is displayed in **table 3.70**.

Table 3.70: Outlier in surface roughness transformation

rhsr	feed	spindl~d
.3734655	200	1500
.5179452	800	12000
.5180496	200	8500
.5727679	800	1500
.705222	400	8500
.9740008	20	15000
1.02524	800	8500
1.132659	600	8500
1.947428	20	8500
3.682326	20	12000

To evaluate the influence of this outlier, the DFBETA of the feed rate and the spindle speed are evaluated and displayed in **table 3.71**. Refer to **annex B note 5** for the DFBETA definition. In this case, the cut-off value for DFBETAs is  $2/\sqrt{35}=0.338$ . The results show that observation at feed rate (20) and spindle speed (12000), strongly affect the coefficients of regression. Including this observation, the regression coefficient of the feed rate and the spindle speed will respectively increase by about 0.88 times and 0.63 times the standard error than the case with the observation excluded. Thus, one alternative is possible is to re-fit a model without this observation.

Table 3.71: DFBETA of feed rate and spindle speed in surface roughness transformation

feed	spindl~d	_dfbeta_7	_dfbet~8
20	1500	.1896229	.258041
20	5000	-.0326362	-.021884
20	8500	-.3752383	.0074004
20	12000	-.8850979	.6284093
20	15000	-.1842228	.2397928

The next analysis is the regression results after excluding the observation at feed (20) and spindle speed (12000). **Table 3.72** indicates that both the feed rate and the spindle speed are significant variables on the surface roughness data; their p-values are near 0, lower than 0.05. Also, the adjusted  $R^2$  is at 0.6377 which indicates that 63.77% of the variability is explained by the feed rate and spindle speed.

Table 3.72: Parameter estimates and ANOVA results for the surface roughness with n=34

Source	SS	df	MS	Number of obs = 34		
Model	9.96914096	2	4.98457048	F( 2, 31) = 30.05		
Residual	5.1429173	31	.165900558	Prob > F = 0.0000		
Total	15.1120583	33	.457941159	R-squared = 0.6597		
				Adj R-squared = 0.6377		
				Root MSE = .40731		

holesurfro~h	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0018478	.0002533	7.30	0.000	.0013313	.0023644
spindlespeed	.0000353	.0000144	2.44	0.020	5.84e-06	.0000647
_cons	.7982598	.1591269	5.02	0.000	.4737184	1.122801

The fitted regression model is:

$$SR = 0.80 + 0.00185 * F + 0.000035 * SS$$

### Interpretation of the coefficients

1. Constant ( $\beta_0$ ): when feed and spindle are zero, the expected mean of surface roughness is 0.8.
2. Feed rate ( $\beta_1$ ): for each one-unit change of feed, surface roughness will increase by 0.00185 when the spindle speed is constant.
3. Spindle Speed ( $\beta_2$ ): for each one-unit change of spindle speed, surface roughness will increase by 0.000035 when the feed rate is constant.

**Figure 3.64** is the plot of this fitted model. The fitted line represents a linear fit. Some outliers are identified on this plot. **Table C.15 in annex C** presents the data for this plot.

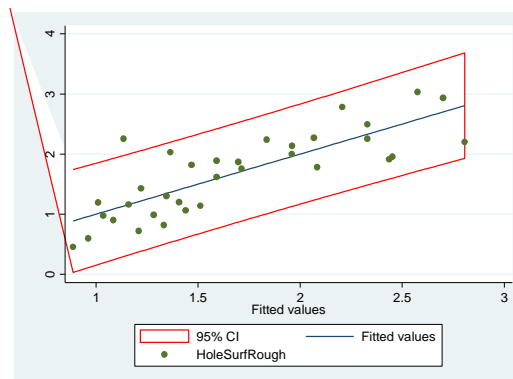


Figure 3.64: Surface roughness observed values against predicted using MLR with n=34

To be able to use this model, the following assumptions are verified: normality of residuals and constant variance of residuals.

**Normality assumption:** Figure 3.65 illustrates the normality plot of the standardized residuals of the surface roughness with the excluded observation. It suggests that the data is spread over the straight line which indicates that the data follow a normal distribution.

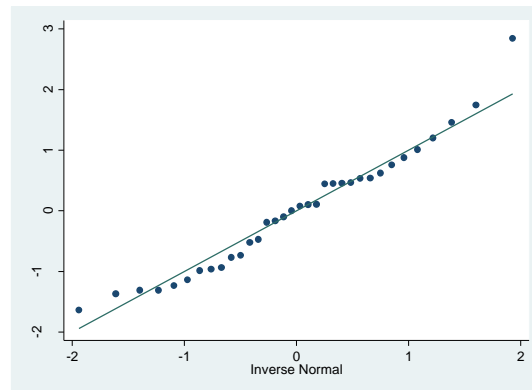


Figure 3.65: Plot of normality of the surface roughness with n=34

The Shapiro-Wilk test is performed to confirm the visual examination. **Table 3.73** presents these results. The p-value corresponds to 0.283, greater than 0.05. Subsequently, the experimental data of the diameter error at exit does follow a normal distribution.

Table 3.73: Shapiro-Wilk Test for surface roughness with n=34

Variable	Obs	W	V	z	Prob>z
rhsr2	34	0.96229	1.317	0.573	0.28329

**Homoscedasticity assumption:** To verify the variance status, the residual plot is drawn. The residuals on **figure 3.67** are structure less which indicates that the variance is constant. After performing the BPCW test, the computed square-chi value is 0.04 and the p-value is equal to 0.8342, greater than 0.05. In this case, the null hypothesis is not rejected and the variance is recognized to be constant.

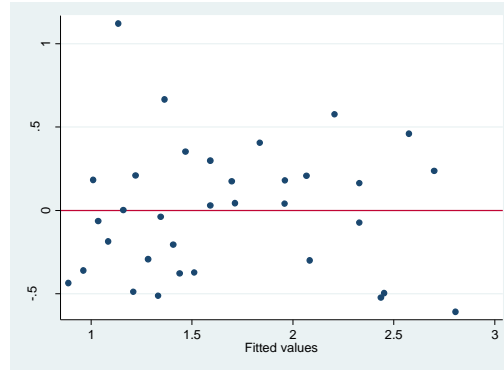


Figure 3.66: Residual plot of the surface roughness experimental data with  $n=34$

To compare the studied models, the Akaike's information for each one of them is explored. The “best” model is the one with the lowest AIC. The models to compare are:

1. Multiple regression model with interaction with  $n=35$ , and
2. Multiple regression model with  $n=34$  (observation excluded at feed rate equal to 20 and spindle speed equal to 12000).

Table 3.74: Akaike's results for surface roughness model with MLR-interaction

Model	Obs	ll (null)	ll (model)	df	AIC	BIC
.	35	-37.63077	-25.74841	3	57.49681	62.16286

Table 3.75: Akaike's results for surface roughness model with MLR  $n=34$

Model	Obs	ll (null)	ll (model)	df	AIC	BIC
.	34	-34.45916	-16.13533	3	38.27066	42.84974

**Tables 3.74** and **3.75** indicate that:

$$\text{AIC}_{n=35} = 57.49681 \text{ compared to } \text{AIC}_{n=34} = 38.27066$$

From a mathematical point of view, the second model is better than the first one for the following reasons:

1. The adjusted  $R^2$  coefficient obtained after exclusion of an observation compared to the original model (0.4717 vs 0.6377) is higher.
2. The AIC for the new model is -21.03 compared to 96.28. The smallest the AIC is, the better the model is.
3. The residual plot in the second model turned out to be normally distributed.

The outlier identified by the software doesn't constitute a "real" outlier to the drilling process. A mechanical failure (at the surface roughness) is expected to happen at low feed rate and high values of spindle speed. Therefore, the model with an excluded observation cannot be selected for prediction purpose. The MLR technique cannot fit models that include outliers that are identified mathematically. Below is a table presenting all the models produced for the surface roughness. No model is recommended.

Table 3.77: Comparison of all methods used for surface roughness

Method used	Validity to use	Variability	Significant variables
MLR	Non-normality	46.12%	Feed rate and spindle speed
MLR with interaction	Non-normality	47.17%	Feed rate and spindle speed
MLR with Observation excluded	Not valid	63.77%	Feed rate and spindle speed

### 3.7 Diameter error at exit analysis

#### 3.7.1 Diameter error at exit distribution study over feed rate

In this section, the feed rate is analyzed as the only input to the diameter error at exit. The boxplot below provides information on the skewness of the distribution, the central location, the variability and the presence of outliers. **Figure 3.67** displays the relationship between the feed rate and the diameter error at exit. The correlation between these two variables seems to be very poor and there is no movement. The graphical illustration indicates a very low variability of the data at all feed rate levels. From levels 20 to 600, the boxes are flat which indicates that all observations are clustered around a single value. However, a small variability is noticed when the feed rate is at 800 and three outliers are observed when the feed rate corresponds respectively to 20, 200 and 600. This representation suggests a lack of influence of the feed rate on the diameter error at exit.

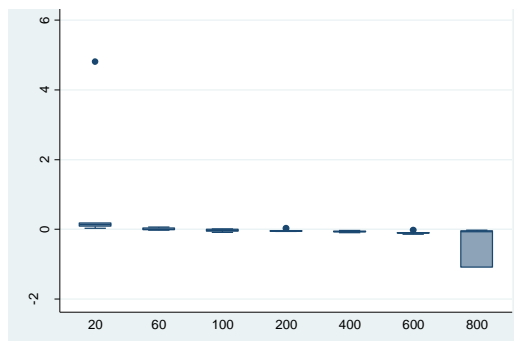


Figure 3.67: Boxplot of diameter error at exit against feed rate

To evaluate the effect of the feed rate on the delamination at exit, the parameter estimates, ANOVA and coefficient of determination results are computed. **Table 3.77** indicates that the p-value of the feed rate is 0.048 which suggests that the feed rate has a limited effect on the diameter error at exit. The known variability explained by the feed rate is at 8.69%. This result confirms the relationship observed on **figure 3.68**.

Table 3.77: Parameter estimates and ANOVA results for the diameter error at exit against feed rate

Source	SS	df	MS	Number of obs = 35		
Model	2.89771492	1	2.89771492	F( 1, 33) = 4.23		
Residual	22.5832562	33	.684341097	Prob > F = 0.0476		
Total	25.4809711	34	.749440327	R-squared = 0.1137		
				Adj R-squared = 0.0869		
				Root MSE = .82725		

holediaerr~t	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	-.0010407	.0005058	-2.06	0.048	-.0020697	-.0000117
_cons	.3783936	.2106202	1.80	0.082	-.0501163	.8069036

### 3.7.2 Diameter error at exit distribution study over spindle speed

In this section, the spindle speed is analyzed to be the only effect on the diameter error at exit. In **figure 3.68**, the boxplot displays the relationship between the spindle speed and diameter error at exit. This plot suggests a lack of influence of the spindle speed on the diameter error at exit. Also, six outliers are identified at the following spindle speed levels: 1500, 5000, 8500 and 15000. The variability is very small and observations are all gathered at the same value.

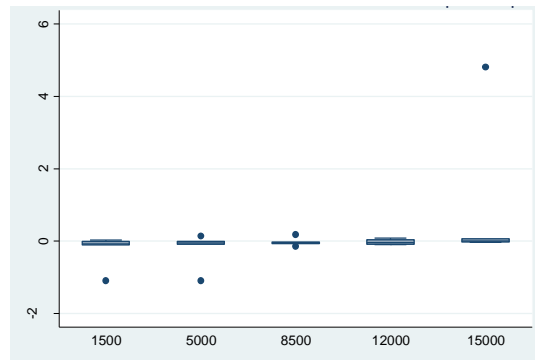


Figure 3.68: Boxplot of diameter error at exit against spindle speed

To detect the effect of the spindle speed on the diameter at exit, the parameter estimates, ANOVA and the coefficient of determination results (**Table 3.78**) are investigated. The computed p-value

of the spindle speed equals 0.067, greater than 0.05. Thus, the spindle speed is not a major factor and has minor impact on the diameter error at exit. Only 7.1% of the variability is explained by the spindle speed.

Table 3.78: Parameter estimates and ANOVA results for the diameter error at exit against spindle speed

Source	SS	df	MS	Number of obs = 35		
Model	2.50130909	1	2.50130909	F( 1, 33) =	3.59	
Residual	22.979662	33	.696353395	Prob > F =	0.0668	
				R-squared =	0.0982	
				Adj R-squared =	0.0708	
Total	25.4809711	34	.749440327	Root MSE =	.83448	
holediaerr~t	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
spindlespeed	.0000556	.0000293	1.90	0.067	-4.08e-06	.0001152
_cons	-.4125319	.283837	-1.45	0.156	-.9900026	.1649387

### 3.7.3 Diameter error at exit multiple linear regression analysis

In the previous sections, the feed rate and spindle speed were determined to be individually irrelevant on the diameter error at exit behavior. Therefore in this section, the feed rate and the spindle speed are put together as inputs to construct a regression model with the diameter error at exit as outcome. As a start, the parameter estimates, ANOVA and coefficient of determination results are calculated in **table 3.79** to detect the variables that have important effect on the diameter error at exit. This table shows that the feed rate is a significant variable on the diameter error at exit with a p-value equal to 0.039, lower than 0.05. However, the spindle speed has a p-value equal to 0.054, over 0.05, which indicates that the spindle speed variation has no trivial effect on the diameter error at exit. Accordingly, the diameter error at exit is affected more by the feed rate than the spindle speed. The adjusted  $R^2$  is at 0.1626. Only 16.26% of the variability is explained by the feed rate and the spindle speed. This indicator is extremely low to justify the appropriateness of the estimated model.

Table 3.79: Parameter estimates and ANOVA results for diameter error at exit using MLR

Source	SS	df	MS	Number of obs = 35		
Model	5.39902401	2	2.699512	F( 2, 32) =	4.30	
Residual	20.0819471	32	.627560847	Prob > F =	0.0222	
				R-squared =	0.2119	
				Adj R-squared =	0.1626	
Total	25.4809711	34	.749440327	Root MSE =	.79219	

holediaerr~t	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	-.0010407	.0004843	-2.15	0.039	-.0020272	-.0000542
spindlespeed	.0000556	.0000278	2.00	0.054	-1.13e-06	.0001123
_cons	-.088424	.3087952	-0.29	0.776	-.7174193	.5405712

From **table 3.80**, the estimated regression model of the diameter error at exit is given by:

$$\text{Drx} = -0.0884 - 0.001041 \cdot F + 0.0000556 \cdot SS$$

### **Interpretation of the coefficients**

1. Constant ( $\beta_0$ ): when the feed rate and spindle are equal zero, the expected mean of diameter error at exit is -0.0884.
2. Feed rate ( $\beta_1$ ): for each one unit change of the feed rate, the diameter error at exit will decrease by 0.001041 when the spindle speed is constant.
3. Spindle Speed ( $\beta_2$ ): For each one unit change of the spindle speed, the diameter error at exit will increase by 0.00006 when the feed rate is constant.

This model cannot be used because the justified variability is very low. To analyze what the next step would be to determine an accurate regression model, the following assumption is investigated: homoscedasticity.

**Homoscedasticity assumption:** Figure 3.69 presents the plot of the standardized residuals of the diameter error at exit experimental data against the fitted values. By visually checking this representation, it indicates that the residuals are not evenly scattered around the line and the variance of the residual is not constant. To verify this interpretation, the BPCW test is completed. The calculated square-chi value is 22.27 and the p-value is equal to 0, lower than 0.05. Thus, the null hypothesis is rejected and the variance is acknowledged as not constant. In this case, a transformation of the diameter error at exit may be good alternative to find a suitable model. In the next section, both procedures will be investigated.

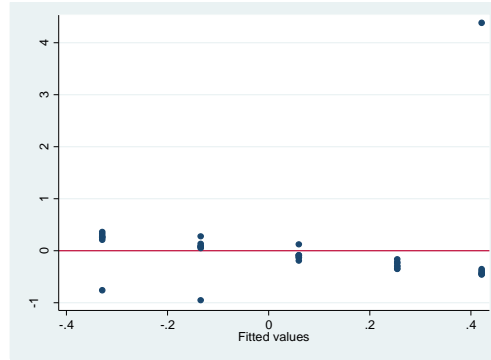


Figure 3.69: Predicted vs standardized residual values plot of the diameter error at exit

### 3.7.4 Diameter error at exit transformation

In previous sections, a suitable prediction model couldn't be determined using the MLR model. In this part of the study, a transformation is applied to the diameter error at exit to attempt finding the best fit model. **Table 3.80** displays the testing results of multiple types of transformations of the diameter error at exit.

Table 3.80: Testing results of the transformations of the diameter error at exit

Transformation	formula	chi2 (2)	P(chi2)
cubic	holedi~t^3	53.91	0.000
square	holedi~t^2	53.65	0.000
identity	holedi~t	48.29	0.000
square root	sqrt(holedi~t)	.	.
log	log(holedi~t)	.	.
1/(square root)	1/sqrt(holedi~t)	.	.
inverse	1/holedi~t	12.67	0.002
1/square	1/(holedi~t^2)	34.06	0.000
1/cubic	1/(holedi~t^3)	16.66	0.000

To accept any type of transformation, the p-value should be over 0.05 which means that the hypothesis if the data follows a normal distribution is accepted. As observed in **table 3.80**, none of the transformation is recommended to fit the diameter error at exit data. To find a suitable model, the transformation of the independent variables is attempted. As a first step, the feed rate will be transformed. **Table 3.81** indicates that the log transformation of the feed rate may be a good option; the p-value corresponds to 0.072, over 0.05.

Table 3.81: Test of the transformation of the feed rate

Transformation	formula	chi2 (2)	P (chi2)
cubic	feed^3	10.51	0.005
square	feed^2	6.87	0.032
identity	feed	6.70	0.035
square root	sqrt(feed)	11.32	0.003
log	log(feed)	5.27	0.072
1/(square root)	1/sqrt(feed)	7.01	0.030
inverse	1/feed	12.83	0.002
1/square	1/(feed^2)	16.64	0.000
1/cubic	1/(feed^3)	17.10	0.000

To analyze furthermore this model, the parameter estimates, ANOVA and coefficient of determination results are generated in **table 3.82**. The feed rate is the only significant effect on the diameter error at exit whereas the p-value is equal to 0.008. The p-value of the spindle speed is at the conventional reference point (0.05) which indicates that this variable has no effect on the outcome. The adjusted  $R^2$  corresponds to 0.234 which still a very low indication of the known variability and the model may not be satisfactory.

Table 3.82: Parameter estimates and ANOVA results for diameter error at exit

Source	SS	df	MS	Number of obs = 35		
Model	7.10970195	2	3.55485097	F( 2, 32) =	6.19	
Residual	18.3712692	32	.574102161	Prob > F =	0.0053	
				R-squared =	0.2790	
				Adj R-squared =	0.2340	
Total	25.4809711	34	.749440327	Root MSE =	.7577	
holediaerr~t	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnfeed	-.2936028	.1036287	-2.83	0.008	-.5046875	-.082518
spindlespeed	.0000556	.0000266	2.09	0.045	1.34e-06	.0001098
_cons	1.100217	.592878	1.86	0.073	-.1074355	2.30787

The prediction model under investigation is:

$$Drx = 1.1 - 0.2936028 * \log(F) + 0.000056 * SS$$

After the log transformation of the independent variable, feed rate, the following assumption is verified: homoscedasticity.

**Homoscedasticity assumption:** To verify the variance status, the residual plot is drawn. Refer to **figure 3.70**. It is clear on the plot below that the variance is not constant and a pattern is observed within the data. The Log transformation of the independent variable, feed rate, did not improve the model fitting. Also, an outlier is identified to be far from the overall data trend. For instance,

this observation arises ( $Drx=20.89178$ ) when the feed rate is at 20 and the spindle speed is at 15000.

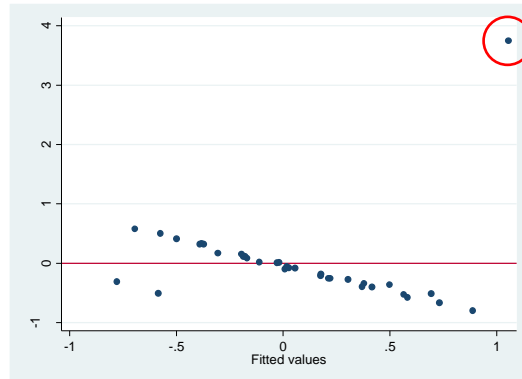


Figure 3.70: Residual plot for the diameter error at exit experimental data

This mathematical outlier is causing the regression method to be ineffective and inconclusive. It is important to keep in mind that a poor correlation exists between the diameter at exit and the independent variables. Below is a table resuming all the models developed for the diameter error at exit fitting. Unfortunately, no model is recommended.

Table 3.83: Comparison of all methods used for the diameter error at exit

Method used	Validity to use	Variability	Significant
MLR	Non-normality and variance not constant	16.26%	Feed rate
Transformation of the feed rate	Non-normality and variance not constant	23.4%	Feed rate

### 3.8 Diameter error at entry analysis

#### 3.8.1 Diameter error at entry distribution study over feed rate

In this section, the feed rate is analyzed as the only effect on diameter error at entry. To understand the relationship between the feed rate and the diameter error at entry, the boxplot below is presented. **Figure 3.71** indicates multiple outliers at the following feed rate levels: 20, 60, 100, 200, 400 and 600. The overall trend of the diameter error at entry has a tendency to slightly decrease with the rise of the feed rate level. The outlier at feed rate 20 is very far from the overall data distribution. This influential point may have a major effect on the multiple regression analysis and the fitting model. The variability of the diameter error at entry within each sample of

the feed rate is very small except at level 800, the variability within this sample is larger than the others.

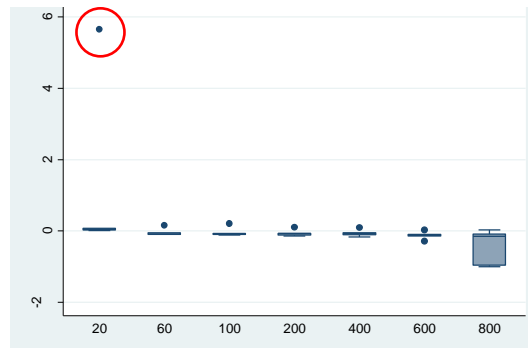


Figure 3.71: Diameter error at entry boxplot against the feed rate

**Table 3.84** indicates that the feed rate is not an important variable on the diameter error at entry whereas its p-value is equal to 0.083, greater than 0.05. Only 6% of the variability is explained by the feed rate.

Table 3.84: Parameter estimates and ANOVA results for diameter error at entry against feed rate

Source	SS	df	MS	Number of obs = 35		
Model	3.01219599	1	3.01219599	F( 1, 33) = 3.19		
Residual	31.1400376	33	.943637505	Prob > F = 0.0832		
				R-squared = 0.0882		
				Adj R-squared = 0.0606		
Total	34.1522336	34	1.00447746	Root MSE = .97141		
holediaerr-y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	-.0010611	.0005939	-1.79	0.083	-.0022694	.0001472
_cons	.3897053	.247324	1.58	0.125	-.1134791	.8928898

### 3.8.2 Diameter error at entry distribution study over spindle speed

In this section, the spindle speed over the diameter error at entry is analyzed to understand their correlation. In **figure 3.72**, a slight increase of the diameter error at entry is observed with the increase of the spindle speed. Accordingly, the diameter error at entry is inversely proportional to the spindle speed. Also, it is noticed that the variability of the diameter error at entry within each sample of the spindle speed is very small. This is an indication that the values are concentrated around the same value. Also, five outliers are identified on the boxplot below at the following spindle speed levels: 1500, 5000, 8500 and 12000. Another extreme observation is observed, at the top right on the plot, to be far from the overall distribution at spindle speed level 15000.

These outliers and influential points will have an imperative impact on the analysis of regression and will make it unusable.

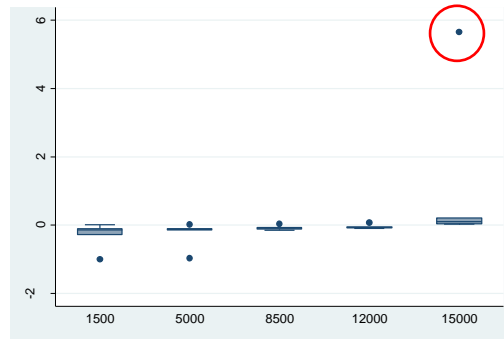


Figure 3.72: Diameter error at entry boxplot with the spindle speed

**Table 3.85** indicates that the spindle speed is an important variable on the diameter error at entry whereas its p-value is equal to 0.041, lower than 0.05. Only 9.37% of the variability is explained by the spindle speed effect.

Table 3.85: Parameter estimates and ANOVA results for diameter error at entry against spindle speed

Source	SS	df	MS	Number of obs = 35		
Model	4.11151157	1	4.11151157	F( 1, 33) =	4.52	
Residual	30.0407221	33	.910324911	Prob > F =	0.0411	
				R-squared =	0.1204	
				Adj R-squared =	0.0937	
Total	34.1522336	34	1.00447746	Root MSE =	.95411	
holediaerr~y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
spindlespeed	.0000713	.0000335	2.13	0.041	3.04e-06	.0001395
_cons	-.5392431	.3245281	-1.66	0.106	-1.199501	.1210143

### 3.8.3 Hole diameter error at entry multiple linear regression analysis

After looking over the relationship of the diameter error at entry against the feed rate and the spindle speed individually, this section puts together those independent variables to observe their effect on the diameter error at entry. As a first step, the MLR analysis is attempted by varying the feed rate and spindle speed on the diameter error at entry data. The regression analysis starts by examining the most important independent variables to the output. As shown in **table 3.86**, the p-value of the feed rate is equal to 0.068, greater than 0.05, which indicates that this independent variable is not significant on the diameter error at entry. However, the p-value of the spindle

speed corresponds to 0.035, lower to 0.05, which suggests that this variable is slightly significant on the outcome. The adjusted  $R^2$  corresponds to 0.1591.

Thus, this parameter is very low and indicates that only 15.91% of the variability of the diameter error at entry is explained by the feed rate and the spindle speed. In general, this indication makes the model inaccurate to be used for any prediction; it cannot determine whether the coefficient estimates and predictions are biased, the residual plot must be assessed.

Table 3.86: ANOVA and parameter estimates tables for the diameter error at entry using MLR

Source	SS	df	MS	Number of obs = 35		
Model	7.12370756	2	3.56185378	F( 2, 32) =	4.22	
Residual	27.0285261	32	.84464144	Prob > F =	0.0237	
				R-squared =	0.2086	
				Adj R-squared =	0.1591	
Total	34.1522336	34	1.00447746	Root MSE =	.91904	

holediaerr~y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	-.0010611	.0005619	-1.89	0.068	-.0022056	.0000834
spindlespeed	.0000713	.0000323	2.21	0.035	5.47e-06	.000137
_cons	-.2087949	.3582439	-0.58	0.564	-.9385139	.520924

From **table 3.86**, the fitted regression model of the diameter error at entry is:

$$\text{Dre} = -0.209 - 0.00106 \cdot F + 0.000071 \cdot SS$$

### Interpretation of the coefficients

1. Constant ( $\beta_0$ ): when feed and spindle are zero, the expected mean of diameter error at entry is -0.209.
2. Feed rate ( $\beta_1$ ): for each one-unit change of feed rate, the diameter error at entry will decrease by 0.00106 when the spindle speed is constant.
3. Spindle speed ( $\beta_2$ ): for each one-unit change of spindle speed, the diameter error at entry will increase by 0.000071 when the feed rate is constant.

On **figure 3.73**, a clear pattern is observed: a straight line with outliers rather than the randomness that is expected. This indicates a bad fit. To find a suitable fitted model, a transformation will be attempted.

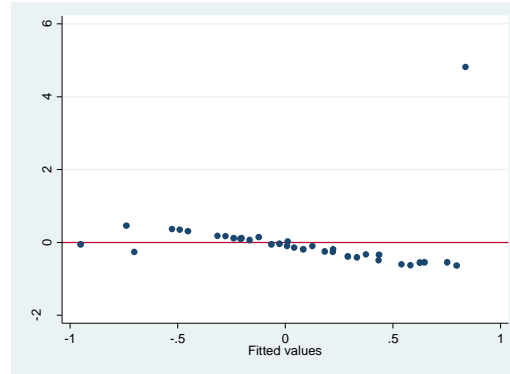


Figure 3.73: Residual plot for diameter error at entry using MLR

### 3.8.4 Diameter error at entry transformation

In this part of the study, a transformation is researched for the diameter error at entry to find a suitable prediction model. **Table 3.87** presents the results of all these transformations. As demonstrated by these results, none of the transformation is recommended to fit the diameter error at entry data. In fact, no p-value is found to be over 0.05; this data distribution does not follow a normal distribution and no transformation is possible.

Table 3.87: Results for the transformation of the diameter error at entry

Transformation	formula	chi2 (2)	P (chi2)
cubic	holedi~y^3	53.93	0.000
square	holedi~y^2	53.83	0.000
identity	holedi~y	50.44	0.000
square root	sqrt(holedi~y)	.	.
log	log(holedi~y)	.	.
1/(square root)	1/sqrt(holedi~y)	.	.
inverse	1/holedi~y	15.76	0.000
1/square	1/(holedi~y^2)	37.40	0.000
1/cubic	1/(holedi~y^3)	42.92	0.000

When checking the observations, an outlier is found to be at feed rate 20 and spindle speed 15000 (same outlier as in previous sections) which can be affecting the data distribution comportment and may be restraining the researches from finding a suitable model at this point. The inclusion of an observation could either contribute to an increase or a decrease in a regression coefficient. In this case, the cut-off value for DFBETAs is  $2/\sqrt{35} = 0.338$ . The DFBETA of the feed rate is 5.7 (in absolute value) and the DFBETA of the spindle speed is 7.39. That indicates that including this observation in the fitting, the regression coefficient of the feed rate and the spindle speed will respectively increase by about 6 times and 7.4 times the standard error than the case with the observation excluded. This outlier strongly affects the coefficients of regression. Again,

this outlier cannot be excluded to complete the regression modelling because this observation is part of the normal behavior of the drilling process.

Table 3.88: BFBETAs for diameter error at entry

feed	spindl~d	_dfbeta_3	_dfbeta_4
20	1500	-.0301517	-.0410308
20	5000	.0219094	.0146913
20	8500	.0674835	-.0013309
20	12000	.11494	-.0816061
20	15000	-5.682657	7.396806

**Table 3.89** explores the only analysis method performed to determine the best fitting model for the diameter error at entry. An interaction effect hasn't been explored for this output because it's clear from the previous analysis that the independent variables have no impact on the dependent.

Table 3.89: Comparison of all methods used for the diameter error at entry analysis

Method used	Validity to use	Variability	Significant
MLR	Non-normality and variance not constant	15.91%	Feed rate

### 3.9 Circularity at exit analysis

#### 3.9.1 Circularity at exit distribution study over feed rate

In this section, an ANOVA is performed to understand the relationship between circularity at exit and feed rate. A boxplot is drawn below to discern the trend of the circularity at exit against the feed rate and to indicate the degree of dispersion in the data and identify any outliers.

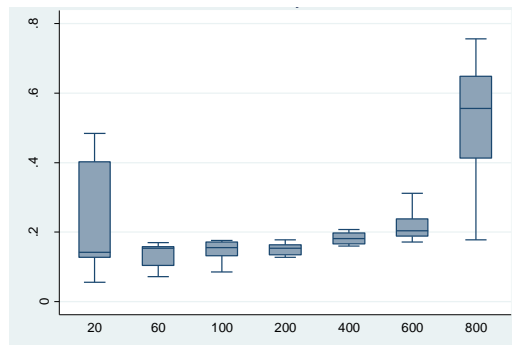


Figure 3.74: Boxplot of the circularity at exit over feed rate

As displayed on **figure 3.74**, the circularity at exit outline resembles to a bell shaped curve. At feed rate levels 20 and 800, the observations cover a wide range which anticipates a large spread

of the data. At feed rate levels 60, 100, 200, 400 and 600, the observations are clustered around a single value to suggest a smaller spread of the data. No outliers are distinguished. In **table 3.90**, the calculated p-value of the feed rate is near 0, lower than 0.05 which demonstrates that the feed rate has a significant effect on the circularity at exit.

Table 3.90: Parameter estimates and ANOVA results for circularity at exit against feed rate

Source	SS	df	MS	Number of obs = 35		
Model	.298065912	1	.298065912	F( 1, 33) =	16.58	
Residual	.593260745	33	.017977598	Prob > F =	0.0003	
				R-squared =	0.3344	
				Adj R-squared =	0.3142	
Total	.891326657	34	.02621549	Root MSE =	.13408	

holecirexit	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0003338	.000082	4.07	0.000	.000167	.0005006
_cons	.1225944	.0341373	3.59	0.001	.0531415	.1920473

### 3.9.2 Circularity at exit distribution study over spindle speed

In this section, an ANOVA analysis is performed to understand the relationship between the value of the circularity at exit and the explanatory variable, the spindle speed. The boxplot below illustrates the relationship and liaison between the circularity at entry against the spindle speed and to indicate the degree of variability in the data. As displayed on **figure 3.75**, the dispersion is large at spindle speed levels 1500, 12000 and 15000. However, a smaller spread is displayed at levels 5000 and 8500. The relationship between both variables seems to be linear which will be confirmed in the next section. Also, this boxplot singles out two outliers at spindle speed levels 5000 and 8500. These observations may affect the fitted model and result in less accurate results.

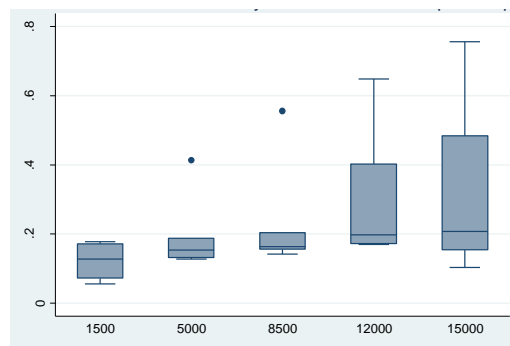


Figure 3.75: Boxplot of the circularity at exit over spindle speed

In **table 3.91**, the calculated p-value is 0.01, lower than 0.05 which indicates that the spindle speed has a significant effect on the circularity at exit.

Table 3.91: Parameter estimates and ANOVA results for circularity at exit against spindle speed

Source	SS	df	MS	Number of obs = 35		
Model	.1635304	1	.1635304	F( 1, 33) =	7.41	
Residual	.727796258	33	.022054432	Prob > F	=	0.0103
				R-squared	=	0.1835
				Adj R-squared	=	0.1587
Total	.891326657	34	.02621549	Root MSE	=	.14851
holecirexit	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
spindlespeed	.0000142	5.22e-06	2.72	0.010	3.59e-06	.0000248
_cons	.1071818	.0505129	2.12	0.041	.0044125	.209951

### 3.9.3 Circularity at exit multiple linear regression

In previous sections, the relationship of the circularity at exit against the feed rate and the spindle speed was analyzed individually, and both variables were found to have significant effect on the outcome. This segment presents the MLR analysis of the circularity at exit. As a first step, the parameter estimates, ANOVA and coefficient results are generated and studied in **table 3.92**. It indicates that the p-values of the feed rate (0.00) and spindle speed (0.001) are both lower than 0.05, which specifies that both independent variables are significant on the dependent. The adjusted  $R^2$  corresponds to 0.4877 which explains half of the variance in the process.

Table 3.92: Parameter estimates and ANOVA results for the diameter error at exit MLR

Source	SS	df	MS	Number of obs = 35		
Model	.461596312	2	.230798156	F( 2, 32) =	17.19	
Residual	.429730346	32	.013429073	Prob > F	=	0.0000
				R-squared	=	0.5179
				Adj R-squared	=	0.4877
Total	.891326657	34	.02621549	Root MSE	=	.11588
holecirexit	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0003338	.0000708	4.71	0.000	.0001895	.0004781
spindlespeed	.0000142	4.07e-06	3.49	0.001	5.92e-06	.0000225
_cons	.0032333	.0451716	0.07	0.943	-.0887783	.0952449

From **table 3.92**, the fitted regression model generated is:

$$Cex = 0.0032 + 0.000334 * F + 0.000014 * SS$$

### **Interpretation of the coefficients**

1. Constant ( $\beta_0$ ): when feed rate and spindle speed values are zero, the expected mean of circularity at exit corresponds to 0.0032.
2. Feed rate ( $\beta_1$ ): for each changes of one unit of feed, circularity at exit will increase by 0.000334 when the spindle speed is constant.
3. Spindle speed ( $\beta_2$ ): for each changes of one unit of spindle speed, circularity at exit will decrease by 0.000014 when the feed rate is constant.

**Figure 3.76** is the plot of this fitted model and **table C.16 in annex C** presents the circularity at exit fitted and observed values.

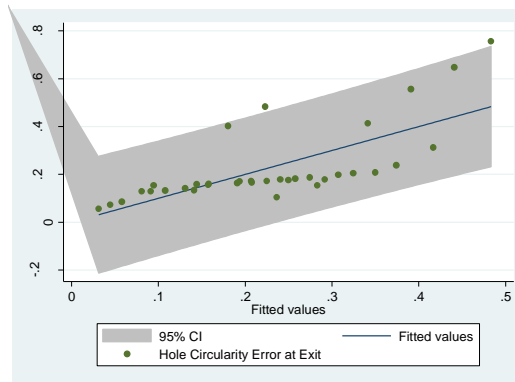


Figure 3.76: Plot of the circularity at exit fitted vs observed values using MLR

To be able to use this model, the regression assumptions shall be verified: independency, normality, homoscedasticity and linearity.

**Independency assumption:** The plot below represents the standardized residuals against the cases. The independency assumption is violated due to an observed pattern in the data.

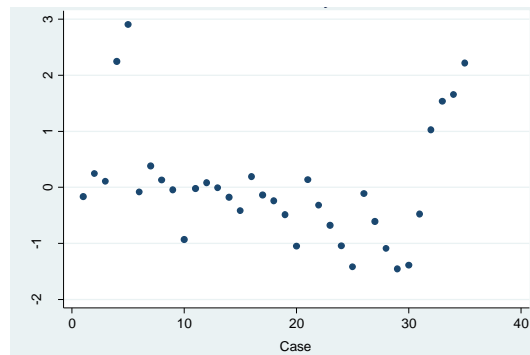


Figure 3.77: Plot of the standardized residuals of the circularity at exit data against the number of cases

**Normality assumption:** The residual plot in **figure 3.78** suggests a slight deviation of the normal distribution at the lower and upper tail, which could be due to the influence of the outliers. The Shapiro-Wilk test (in **table 3.93**) confirms that the standardized residuals are not normally distributed as p-value corresponds to 0.00312, lower than 0.05 and the null hypothesis is rejected.

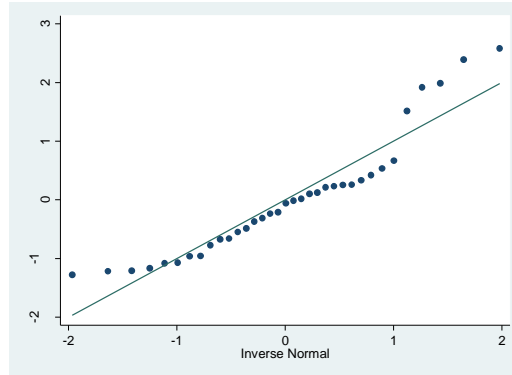


Figure 3.78: Normal plot of the standardized residuals of the circularity at exit data

Table 3.93: Shapiro-Wilk test for normal data of the circularity at exit data using MLR

Variable	Obs	W	V	z	Prob>z
rcirexit	35	0.89612	3.708	2.735	0.00312

**Homoscedasticity assumption:** **Figure 3.79** displays that the variance is not constant. When observing this plot, some experts may think that the variable is constant. Therefore, to confirm this observation, the BPCW test is completed. In this case, it's expected that the p-value will be lower than 0.05 to accept the null hypothesis. After performing this test, the calculated square-chi value corresponds to 9.14 and the p-value is at 0.0025, lower than 0.05 which indicates that  $H_0$  cannot be rejected. Hereafter, the variance is not constant in this case.

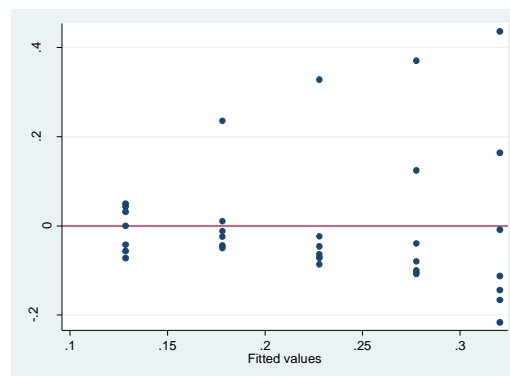


Figure 3.79: Residuals plot for circularity at exit

**Linearity assumption:** Figure 3.80 indicates that the relationship between the circularity at exit and the feed rate is not linear. However, the association between the circularity at exit and the spindle speed is linear. Thus, this assumption is not valid unless both scatterplots reflect a linear fitting.

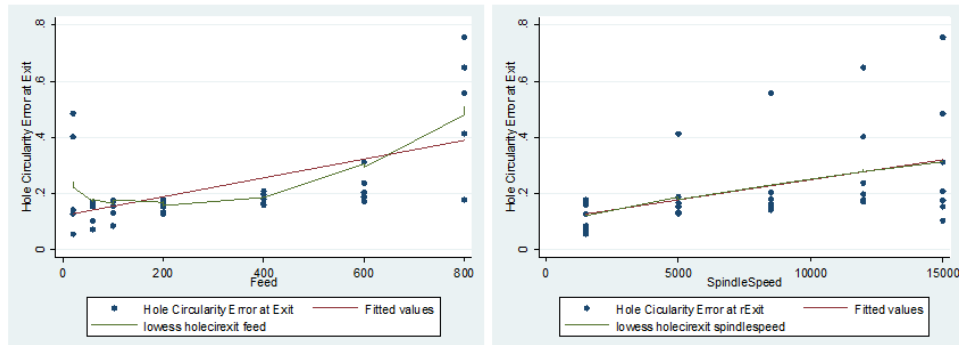


Figure 3.80: Scatterplots feed rate and spindle speed against circularity at exit

**Table 3.94** presents a brief description of each assumption's validation. Hence, this model cannot be used because all the assumptions are violated. This issue may be caused by some outliers. Prior to building another regression model, the observations with extreme values of the independent variables is identified and their influence investigated.

Table 3.94: Illustration of the assumption validation results for circularity at exit using MLR

Assumptions	Valid for this model
Independency	No
Normality	No
Linearity	No
Homoscedasticity	No

The leverage against the normalized residual squared plot is drawn (**figure 3.81**).

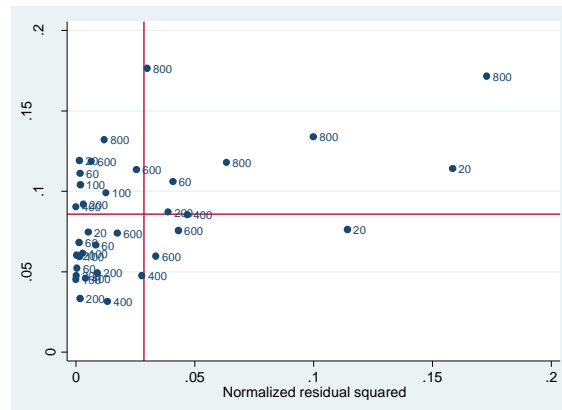


Figure 3.81: Plot of the leverage against normalized residual squared of the circularity at exit

The leverage measures how far an independent variable deviates from its mean. High leverage points could have a prodigious effect on the estimate of the coefficients of regression. The inclusion of these influential points can substantially change the estimate of the regression coefficients. **Figure 3.81** reveals high leverage points on the upper left corner of the plot and high points in the absolute of residuals at the lower right corner. Also, there are points high in leverage and in the absolute of residuals are situated on the upper right portion.

In this case, the cut-off value for DFBETAs is  $2/\sqrt{35} = 0.338$ . To identify which observations are associated with the extreme point on the plot, the DFBETA of the feed rate and spindle speed are generated and all points with dfbeta\_5 and dfbeta\_6 over 0.34 are the researched extreme values. **Table 3.95** presents the calculated DFBETA of the feed rate (dfbeta\_5) and spindle speed (df\_beta6). It indicates as well that the observation at feed rate [800] and spindle speed [15000] would increase the coefficient for feed rate and spindle speed by respectively 0.94 and 0.73 standard errors. Thus, one alternative would be to refit a model without this observation.

Table 3.95: Values of DFBETA of feed rate and spindle speed for circularity at exit

	feed	spindl~d	_dfbeta_5	_dfbeta_6
1.	20	1500	-.0425977	-.0579675
2.	20	5000	-.0771417	-.0517269
3.	20	8500	-.0182153	.0003592
4.	20	12000	-.3877818	.2753206
5.	20	15000	-.4919004	.6402801
6.	60	1500	-.0403156	-.06359
7.	60	5000	-.0838542	-.0651734
8.	60	8500	-.0192315	.0004396
9.	60	12000	.0333368	-.0274342
10.	60	15000	.1979255	-.2986153
11.	100	1500	-.0344371	-.0645941
12.	100	5000	-.0284356	-.026282
13.	100	8500	.0016023	-.0000436
14.	100	12000	.0411455	-.0402663
15.	100	15000	.0905163	-.1624007
16.	200	1500	-.0234244	-.0833686
17.	200	5000	.0042752	.0074975
18.	200	8500	.0160362	-.0008272
19.	200	12000	.038221	-.0709722
20.	200	15000	.0836339	-.2847151
21.	400	1500	.000981	-.0043925
22.	400	5000	-.0201866	.044538
23.	400	8500	-.036115	-.0023436
24.	400	12000	-.0535439	-.1250834
25.	400	15000	-.0732426	-.3136859
26.	600	1500	-.0901102	.123837
27.	600	5000	-.1413853	.0957438
28.	600	8500	-.1952087	-.003888
29.	600	12000	-.2258269	-.161922
30.	600	15000	-.179476	-.2359272
31.	800	1500	-.3564699	.2893509
32.	800	5000	.2110224	-.0844035
33.	800	8500	.4923854	.0057924
34.	800	12000	.6452944	.2732835
35.	800	15000	.9384779	.7286533

As a first step of the re-fitting model without the observation at feed rate [800] and spindle speed [15000] is to generate the parameter estimates, ANOVA and coefficient of determination results in **table 3.96**.

Table 3.96: Parameter estimates and ANOVA results for hole circularity at exit with n=34

Source	SS	df	MS	Number of obs = 34		
Model	.262726881	2	.13136344	F( 2, 31) =	11.98	
Residual	.340030081	31	.010968712	Prob > F	=	0.0001
				R-squared	=	0.4359
				Adj R-squared	=	0.3995
Total	.602756962	33	.018265362	Root MSE	=	.10473

holecirexit	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0002737	.0000674	4.06	0.000	.0001362	.0004111
spindlespeed	.0000115	3.80e-06	3.04	0.005	3.78e-06	.0000193
_cons	.0350705	.0423153	0.83	0.414	-.0512321	.121373

**Table 3.96** indicates that the feed rate and spindle speed are both significant variables on the circularity at exit with respective p-values of 0 and 0.005; both lower than 0.05. Also, the adjusted  $R^2$  corresponds to 0.3995 revealing that 39.95% of the variability is explained by the feed rate and spindle speed.

Based on this table, the fitted regression model is:

$$C_{ex} = 0.0351 + 0.000274 \cdot F + 0.0000115 \cdot SS$$

### Interpretation of the coefficients

1. Constant ( $\beta_0$ ): when feed and spindle are zero, the expected mean of circularity at exit is 0.0351.
2. Feed rate ( $\beta_1$ ): for each one-unit change of feed, circularity at exit will increase by 0.000274 when the spindle speed is constant.
3. Spindle Speed ( $\beta_2$ ): for each one-unit change of Spindle Speed, circularity at exit will increase by 0.0000115 when the feed rate is constant.

**Figure 3.82** is the plot of this fitted model and **table C.17 in annex C** presents the circularity at exit data values after excluding the observation at feed rate [800] and spindle speed [15000].

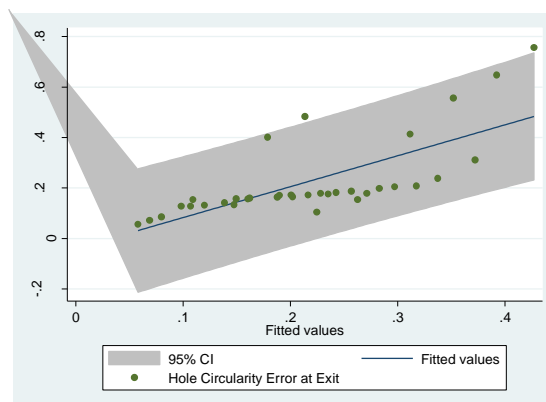


Figure 3.82: Plot of the circularity at exit fitted vs observed values n=34

To compare the studied models, the Akaike's information is calculated for each one.

Table 3.97: Akaike's results for circularity at exit model with MLR

Model	Obs	ll (null)	ll (model)	df	AIC	BIC
.	35	14.56902	27.3362	3	-48.67239	-44.00635

Table 3.98: Akaike's results for circularity at exit model with multiple regression–n=34

Model	Obs	ll (null)	ll (model)	df	AIC	BIC
.	34	15.01388	29.41249	3	-52.82499	-48.24591

**Tables 3.97 and 3.98** indicate that:

$$\mathbf{AIC}_{n=35} = \mathbf{-48.67} \text{ compared to } \mathbf{AIC}_{n=34} = \mathbf{-52.82}$$

In this case, the first model is better than the second one for the following reasons:

1. The adjusted  $R^2$  coefficient obtained after exclusion of an observation compared to the original model is lower: 0.3995 vs 0.4877. No improvement seen after excluding the influential point.
2. As indicated in the ANOVA results for both models:
  - i. The SS residuals of the 2<sup>nd</sup> model is smaller (0.34) compared to the SS residual of the 1<sup>st</sup> model (0.43). However, percentage-wise, SS residuals of the 2<sup>nd</sup> model constitute 56.67% from the total SS versus 48.31% for the 1st model.
  - ii. The root MSE is reduced by the new model (0.105 vs 0.116). The smaller this value, the closer the fit is to the data.
3. The AIC for the new model is -52.82 compared to -48.67. The smallest the AIC is, the better the model is.

The model with the excluded observation did not provide better results. A transformed or non-linear model won't provide better results because of the presence of an outlier. The table below presents a summary of all the analysis performed on the circularity at exit. However, none of the models are recommended to be used to their low significance.

Table 3.99: Comparison of all methods used for the circularity at exit analysis

Method used	Validity to use	Variability	Significant variables
MLR	Non-normality, non-linearity and variance not constant	48.77%	Feed rate and spindle speed
MLR with excluded observation	Non-normality and variance not constant	39.95%	Feed rate and interaction

### 3.10 Circularity at entry analysis

#### 3.10.1 Circularity at entry distribution study over feed rate

In this section, an ANOVA is accomplished to understand the relationship between the circularity at entry and the feed rate. A boxplot is drawn below to discern the trend of the circularity at entry against the feed rate and to indicate the degree of dispersion in the data and identify any outliers. As displayed on **figure 3.83**, the circularity at entry outline is close to a linear. At feed rate levels 20 and 800, outliers are identified at the upper side which may cause issues in the fitted regression model. The dispersion of the data within each level is large which indicates that the data doesn't vary around the same value. In **table 3.100**, the calculated p-value is near 0.097, greater than 0.05 which demonstrates that the feed rate has no significant effect on the circularity at entry. Also, only 5.36% of the variability is explained by the feed rate.

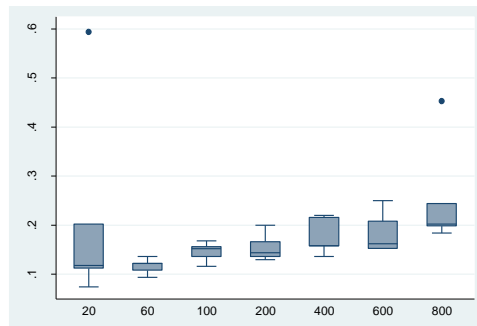


Figure 3.83: Boxplot of circularity at entry over the feed rate

Table 3.100: Parameter estimates and ANOVA results for circularity at entry over feed rate

Source	SS	df	MS	Number of obs = 35		
Model	.026022786	1	.026022786	F( 1, 33) =	2.93	
Residual	.293481604	33	.008893382	Prob > F =	0.0965	
				R-squared =	0.0814	
				Adj R-squared =	0.0536	
Total	.31950439	34	.009397188	Root MSE =	.0943	

holecirentury	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0000986	.0000577	1.71	0.097	-.0000187	.0002159
_cons	.1486858	.0240103	6.19	0.000	.0998366	.1975351

### 3.10.2 Circularity at entry distribution study over spindle speed

After verifying the relationship between the circularity at entry and feed rate, this section will attempt to verify the association between the circularity at entry and the spindle speed. The plot below is a visualization to view how the data is distributed. As displayed on **figure 3.84**, the circularity at entry against the spindle speed outline is slightly curved. Also, the boxplot detects outliers and extreme points at spindle speed levels 5000 and 15000 on the upper side. From the length of each box, the variability is determined to be large which indicates that the data is spread within each range. By observing the horizontal line inside each box which represents the median, it suggests that the distribution is skewed because the median is mostly not centered. In **table 3.101**, the calculated p-value is 0.764, greater than 0.05 which indicates that the feed rate has no significant effect on the circularity at entry. The adjusted  $R^2$  is roughly equal to 0 which is in indication of lack of explanation of the known variability. The spindle speed doesn't contribute at all to the circularity at entry performance.

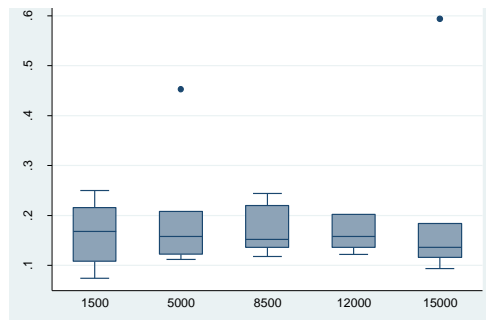


Figure 3.84: Boxplot of circularity at entry over the spindle speed

Table 3.101: Parameter estimates and ANOVA results for circularity at entry over spindle speed

Source	SS	df	MS	Number of obs = 35		
Model	.000882456	1	.000882456	F( 1, 33) =	0.09	
Residual	.318621934	33	.00965521	Prob > F =	0.7643	
				R-squared =	0.0028	
				Adj R-squared =	-0.0275	
Total	.31950439	34	.009397188	Root MSE =	.09826	
holecirentry	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
spindlespeed	1.04e-06	3.45e-06	0.30	0.764	-5.98e-06	8.07e-06
_cons	.1706318	.0334222	5.11	0.000	.1026338	.2386298

### 3.10.3 Circularity at entry multiple linear regression

In previous sections, the relationship of the circularity at entry against the feed rate and the spindle speed were examined. As a first step in the analysis of regression, the parameter estimates, ANOVA and coefficient of determination results are generated and studied. **Table 3.102** indicates that the p-value of the feed rate (0.101) and spindle speed (0.758) are both greater than 0.05, which specifies that both independent variables are not significant on the hole circularity at entry data. The adjusted  $R^2$  corresponds to 0.027; this indication is very low and means that only 2.7% of the variability is explained by the feed rate and spindle speed.

Table 3.102: Parameter estimates and ANOVA results for circularity at entry using MLR

Source	SS	df	MS	Number of obs = 35		
Model	.026905241	2	.013452621	F( 2, 32) =	1.47	
Residual	.292599148	32	.009143723	Prob > F =	0.2448	
				R-squared =	0.0842	
				Adj R-squared =	0.0270	
Total	.31950439	34	.009397188	Root MSE =	.09562	
holecirentry	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
feed	.0000986	.0000585	1.69	0.101	-.0000205	.0002177
spindlespeed	1.04e-06	3.36e-06	0.31	0.758	-5.80e-06	7.89e-06
_cons	.1399176	.0372738	3.75	0.001	.0639933	.215842

From **table 3.102**, the fitted regression model is:

$$\text{Cen} = 0.14 + 0.00010 \cdot F + 1 \times 10^{-6} \cdot SS$$

#### Interpretation of the coefficients

1. Constant ( $\beta_0$ ): when feed and spindle are zero, the expected mean of the circularity at entry is 0.14
2. Feed rate ( $\beta_1$ ): For each changes of one unit of feed, the circularity at entry will increase by 0.00010 when the spindle speed is constant.

3. Spindle Speed ( $\beta_2$ ): for each changes of one unit of spindle speed, the circularity at entry will decrease by 0.000001 when the deed rate is held constant.

**Figure 3.85** is the plot of this fitted model and **table C.18 in annex C** presents the circularity at entry fitted and observed values.

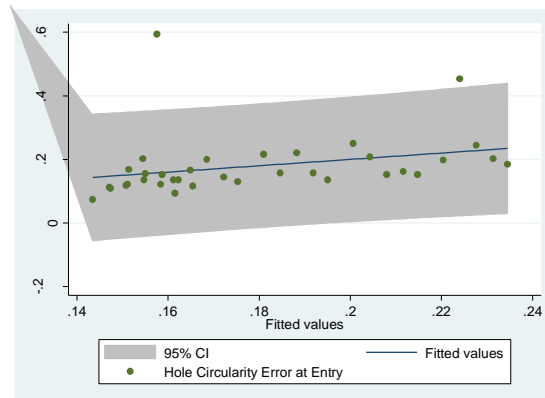


Figure 3.85: Plot of the circularity at entry fitted vs observed values using MLR

As observed previously, the plot displays the occurrence of outliers. This model is not useful because of lack of known variability and presence of outliers. Despite the uselessness of this model, the assumptions will be verified to better understand the lack of fitness of this case.

**Independency assumption:** As observed on **figure 3.86**, this assumption is violated because the residual follow a repetitive structured pattern. Also, outliers are observed.

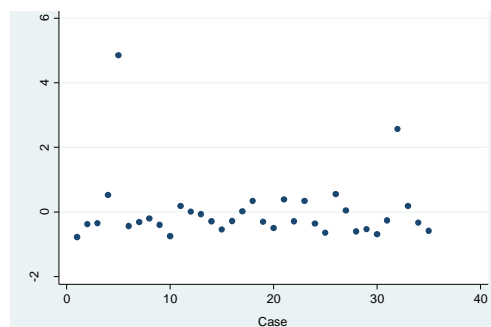


Figure 3.86: Plot of standardized residuals of the circularity at entry against case number

**Normality assumption:** The residual plot, in **figure 3.87**, suggests a deviation of the normal distribution at the lower and upper tail; and outliers are detected as well at the upper side. It is an indication that the standardized residuals are not normally distributed. The Shapiro-Wilk test, from **table 3.103**, confirms the results observed on the illustration below with a p-value equal to 0 lower than 0.05; the null hypothesis of the normality test is rejected.

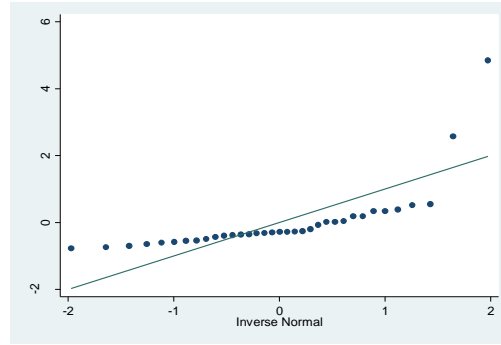


Figure 3.87: Plot of standardized residuals of the circularity at entry

Table 3.103: Shapiro-Wilk test for the circularity at entry

Variable	Obs	W	V	z	Prob>z
rcirentry	35	0.57167	15.288	5.693	0.00000

**Linearity assumption:** Figure 3.88 suggests that the relationship between the feed rate and the circularity at entry is slightly linear. However, the outline in the scatterplot of the spindle speed and the circularity at entry is not perfectly linear. This assumption is violated.

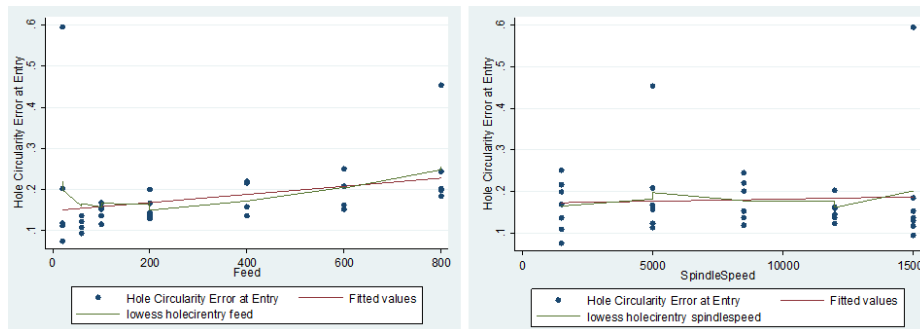


Figure 3.88: Scatterplots feed rate and spindle speed against circularity at entry

**Homoscedasticity assumption:** The residuals variance of the circularity at entry in figure 3.89 suggests that the residuals behavior is not constant. BPCW test is carried out to verify the results of the illustration above. The calculated square-chi value is 4.90 and p-value is at 0.0268, lower than 0.05 which indicates that  $H_0$  is rejected and the variance is confirmed to be not constant.

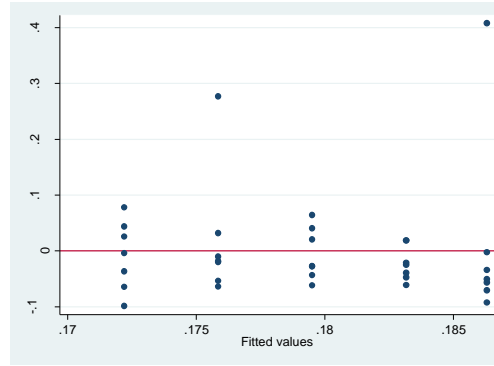


Figure 3.89: Residuals plot for circularity error at entry

**Table 3.104** presents a brief description of each assumption's validation.

Table 3.104: Illustration of the assumption validation results for circularity at exit

Assumptions	Valid for this model
Independency	No
Normality	No
Linearity	No
Homoscedasticity	No

Hence, this model cannot be used because all the assumptions are violated. This issue may be caused by the outliers identified in the analysis. Because the independent variables, the feed rate and spindle speed do not affect the circularity at entry and because the variance of the residuals is not constant, the exclusion of the outliers will not significantly change the regression coefficients.

The table below presents a summary of all the analysis performed on the circularity at entry.

Table 3.105: Comparison of all methods used for the circularity at entry analysis

Method used	Validity to use	Variability	Significant variables
MLR	Non-normality, non-linearity and variance not constant	2.7%	None

In previous sections, the analysis was done using the feed rate and the spindle speed as the independent input variables to study their effect combined together on the uncontrollable variables and the desired quality characteristics. As explained previously, the cutting force and the torque are highly correlated which indicates that when they are put together in the same analytical model as inputs, it may generate inaccurate and unreliable results. To get conclusive results and avoid variables cancelling each other, researches decided to study the effect of the thrust force and the cutting force on the delamination at entry and exit using the simple and multiple regression applications. The boxplots won't be presented because the variables under

study are continuous and can't be grouped together. Boxplots are made to compare groups. In previous sections, the distribution of the uncontrollable variables is pulled across the different group of values of the feed rate and the spindle speed. In the following sections, the two comparable variables (thrust force and cutting force) are continuous.

### 3.11 Delamination at exit analysis versus thrust force and cutting force

#### 3.11.1 Delamination at exit distribution study over thrust force

In this section, the thrust force is analyzed as the only effect on the delamination at exit. As a start, the parameter estimates, ANOVA and coefficient of determination results are generated in **table 3.106**.

Table 3.106: Parameter estimates and ANOVA results of delamination at exit against thrust force

Source	SS	df	MS	Number of obs = 35		
Model	.11936803	1	.11936803	F( 1, 33) =	214.24	
Residual	.018386699	33	.000557173	Prob > F =	0.0000	
				R-squared =	0.8665	
				Adj R-squared =	0.8625	
Total	.137754729	34	.00405161	Root MSE =	.0236	
delexit	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
thrustforce	.0004689	.000032	14.64	0.000	.0004037	.000534
_cons	.9855196	.0061251	160.90	0.000	.9730579	.9979813

This table reveals that the calculated p-value is near 0, lower than 0.05 which indicates that the thrust force has a significant effect on the delamination at exit. The adjusted  $R^2$  corresponds to 0.8625 which implies that 86.25% of the variability is explained by the thrust force. Thus, the thrust force has an important influence on the delamination at exit comportment. The simple linear model is represented by:

$$\text{Dex} = 0.9855196 + 0.0004689 \cdot \text{TF}$$

**Figure 3.90** is the plot of this simple linear regression of the thrust force against the delamination at exit. It indicates that the observed values lie closely to the fitted line except two observations. **Table C.19 in annex C** presents the data for this plot.

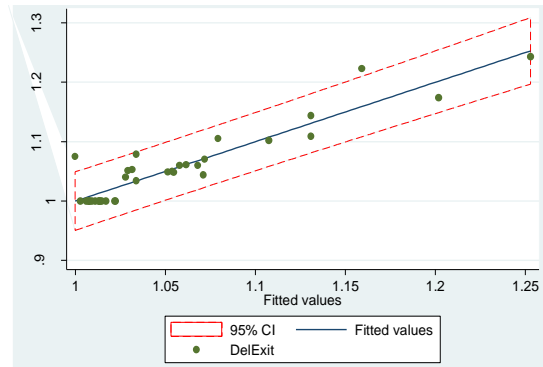


Figure 3.90: Plot of the thrust force against the delamination at exit

Before moving forward with the cutting force analysis, it's important to investigate if this simple regression model can represent the delamination at exit. Thus, the three assumptions described previously are verified: normality, homoscedasticity and linearity.

**Normality assumption:** The normal probability plot below suggests a deviation of the straight line at the lower and upper tail which could be potentially be due to the influence of the outliers noticed earlier. Shapiro test (**table 3.107**) indicates that the p-value is 0.00004, lower than 0.05. Consequently, this assumption is violated.

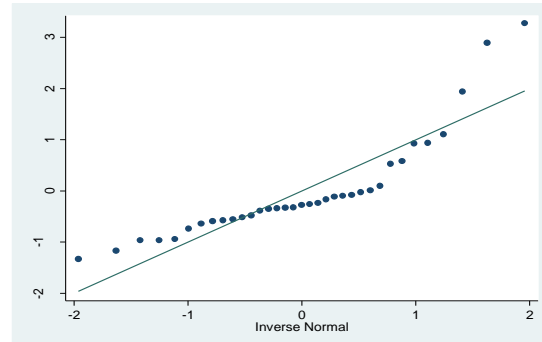


Figure 3.91: Normal probability plot for the delamination at exit simple regression

Table 3.107: Shapiro-Wilk Test for delamination at exit normality delamination at exit against thrust force

Variable	Obs	W	V	z	Prob>z
rstanddele~f	35	0.81495	6.605	3.941	0.00004

**Homoscedasticity assumption:** The residual plot below demonstrates that the variance of the residual is constant and the values are fluctuating randomly around zero. The BPCW test is performed and the calculated chi-square corresponds to 0.54 and the p-value of the chi-square is

0.4613; higher than 0.05 which indicates that the null hypothesis that the variance of residuals is constant cannot be rejected.

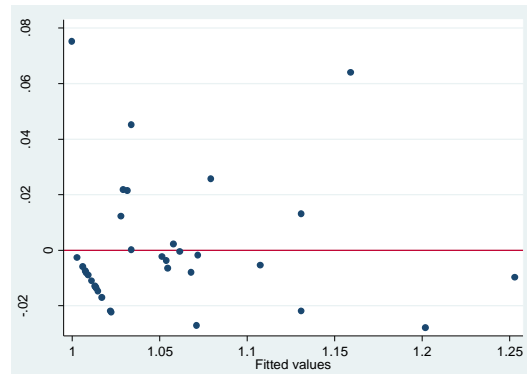


Figure 3.92: Residual plot for the delamination at exit data using simple regression

**Linearity assumption:** Figure 3.93 indicates that the lowess outline closely overlaps the regression outline which confirms that the existent relationship between the thrust force and the delamination at exit is linear.

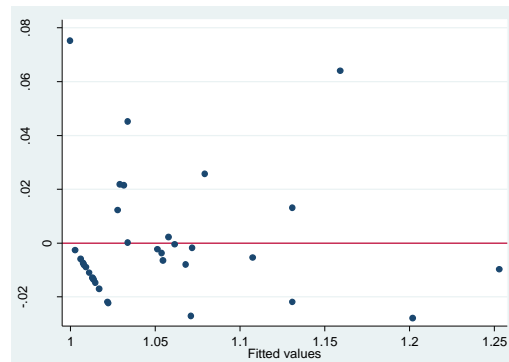


Figure 3.93: Scatterplot of thrust force versus feed rate

Below is a summary of the assumptions' validation results for the delamination at exit simple linear regression model (**Table 3.108**) while the only input is the thrust force.

Table 3.108: Illustration of the assumption validation results for delamination at exit-simple regression

Assumptions	Valid for this model
Normality	No
Homoscedasticity	Yes
Linearity	Yes

Based on this series of analysis, this model does not follow the normality assumption. As discussed previously, the normality is not a mandatory requirement for prediction purposes. In this case, the linear regression model does suit this data and experts can rely on this fitting model to predict the trend of the thrust force at any point of this experimental domain of the delamination at exit.

### 3.11.2 Delamination at exit distribution study over cutting force

In this section, the cutting force is analyzed over the delamination at exit to understand their relationship and dependency. As shown in **table 3.109**, the p-value of the cutting force is lower than 0 which is below the conventional reference point, which indicates that the cutting force does have an important effect on the delamination at exit. The adjusted  $R^2$  equals to 0.8042 which indicates that 80.42% of the variability of the delamination at exit is explained by the cutting force.

Table 3.109: Parameter estimates and ANOVA results of cutting force against delamination at exit

Source	SS	df	MS	Number of obs = 35		
Model	.111581588	1	.111581588	F( 1, 33) =	140.69	
Residual	.026173141	33	.000793125	Prob > F	=	0.0000
				R-squared	=	0.8100
				Adj R-squared	=	0.8042
Total	.137754729	34	.00405161	Root MSE	=	.02816
delexit	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
cuttingforce	.0014497	.0001222	11.86	0.000	.0012011	.0016984
_cons	.9588188	.0092972	103.13	0.000	.9399034	.9777341

The simple linear model is represented by:

$$\text{Dex} = 0.9588188 + 0.0014497 \cdot \text{CF}$$

**Figure 3.94** is the plot of the simple linear regression of the cutting force against the delamination at exit. The overall trend tends to rise. The plot indicates that the scatter of points lie closely to the regression line and are within the 95% confidence of interval of the predicted values. However, one outlier is identified to be indicating a possible lack fit of the model. Further investigations are needed into the assumptions to check the validity of this model. **Table C.20 in annex C** presents the data of this plot.

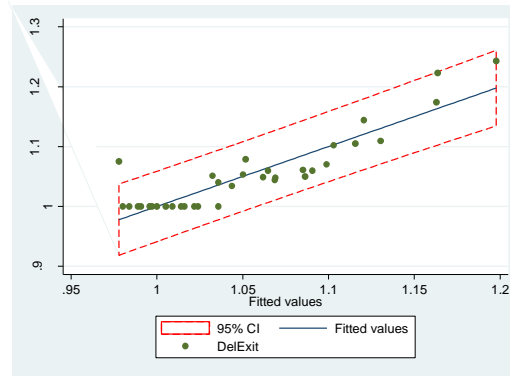


Figure 3.94: Plot of the cutting force against the delamination at exit

Before presenting any conclusions, the appropriateness of the regression model for the delamination at exit data is examined.

**Normality assumption:** The normal probability plot below shows a deviation of the straight line at the lower and upper tail. The Shapiro-Wilk test is performed and the results are presented in **table 3.110**. It indicates that the calculated p-value is 0.00137, lower than 0.05. Thus, the thrust force does not follow a normal distribution.

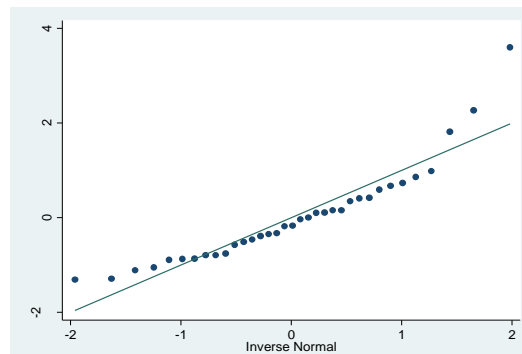


Figure 3.95: Normal probability plot for the cutting force data against the cutting force

Table 3.110: Shapiro-Wilk Test for cutting force normality using simple regression

Variable	Obs	W	V	z	Prob>z
rdelexitcf~d	35	0.88238	4.198	2.995	0.00137

**Homoscedasticity assumption:** The residual plot below confirms that the variance of the residual is constant. The calculated chi-squared corresponds to 0.03 and the p-value of the chi-squared is 0.8572, higher than 0.05 which indicates that the null hypothesis that the variance of residuals is constant cannot be rejected.

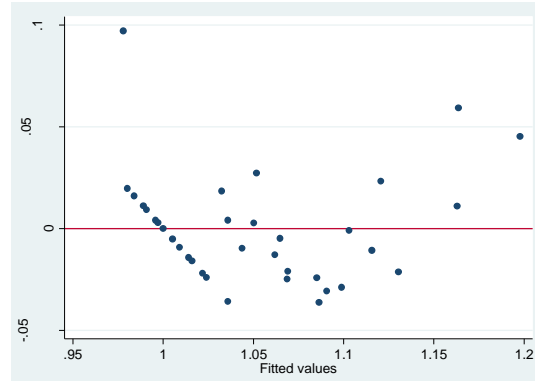


Figure 3.96: Residual plot for the cutting force versus delamination at exit data using simple linear regression

**Linearity assumption:** The relationship between the cutting force and the delamination at exit must be linear. Thus, to verify this assumption, the scatterplot of the cutting force against the delamination at exit is illustrated in **figure 3.97**. The lowess fitted curve suggests a bell-shaped model: the quadratic.

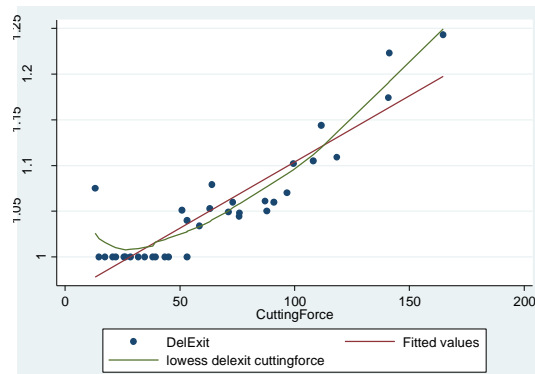


Figure 3.97: Scatterplot of cutting force against delamination at exit

Below is a summary of the assumptions' validation results for the delamination at exit simple regression model (**Table 3.111**).

Table 3.111: Illustration of the assumption validation results for delamination at exit versus cutting force using simple regression

Assumptions	Valid for this model
Normality	No
Homoscedasticity	Yes
Linearity	No

Based on this series of analysis, this model does not follow the normality and the linearity assumptions. In this case, the linear regression model does not suit this data. Because the linearity assumption is not validated, the option to fit this model is to develop a nonlinear model. Note that data transformation is used to refit the data only when the variance is not constant. Therefore, as illustrated on **figure 3.97**, the most apparent model to fit properly this data is the quadratic fit. The following analysis outlines the nonlinear analysis to get a mathematical model.

**Table 3.112** indicates as well that the adjusted  $R^2$  is at 88.85%; a very good estimation. The variability of the simple regression model developed previously was 80.42% compared to 88.85% with the quadratic model.

Table 3.112: Parameter estimates and ANOVA results of delamination at exit against cutting force using nonlinear regression

Source	SS	df	MS	Number of obs = 35		
Model	.123300239	2	.06165012	F( 2, 32) =	136.48	
Residual	.01445449	32	.000451703	Prob > F =	0.0000	
				R-squared =	0.8951	
				Adj R-squared =	0.8885	
Total	.137754729	34	.00405161	Root MSE =	.02125	
delexit	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
cuttingforce	-.0002544	.000347	-0.73	0.469	-.0009613	.0004525
cf2	.0000108	2.12e-06	5.09	0.000	6.48e-06	.0000151
_cons	1.007646	.0118796	84.82	0.000	.9834481	1.031844

Based on the table above, the following nonlinear regression fitted model is pulled:

$$\text{Dex} = 1.007646 - 0.0002544 \cdot \text{CF} + 0.0000108 \cdot \text{CF}^2$$

Below is the plot of the quadratic fitted model of the delamination at exit against the cutting force. The predicted values are slightly lying on the fitted line and within the interval of confidence. This indicates that this model fits well the thrust force data. **Table C.21 in annex C** presents the data of this plot.

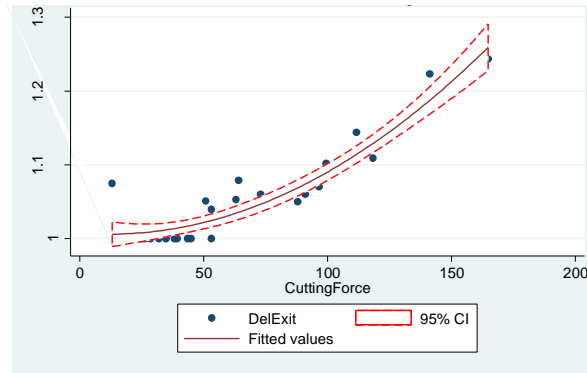


Figure 3.98: Plot of the quadratic delamination at exit fitted model

### 3.11.3 Delamination at exit distribution study over thrust force and cutting force

In this section, the thrust force and cutting force are combined together against the delamination at exit to understand their effect's impact. As indicated in **table 3.113**, the calculated p-value of the cutting force is 0.308 higher than 0.05 which indicates that the cutting force does not have an important effect on the delamination at exit. The p-value of the thrust force is 0 lower than 0.05 which indicates that the thrust force have an important effect on the delamination at exit.

Table 3.113: Parameter estimates and ANOVA results of thrust force and cutting force against delamination at exit

Source	SS	df	MS	Number of obs = 35		
Model	.11996525	2	.059982625	F( 2, 32) = 107.90		
Residual	.017789479	32	.000555921	Prob > F = 0.0000		
Total	.137754729	34	.00405161	R-squared = 0.8709		
				Adj R-squared = 0.8628		
				Root MSE = .02358		
delexit	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
cuttingforce	.0003197	.0003085	1.04	0.308	-.0003086	.000948
thrustforce	.0003745	.0000964	3.88	0.000	.0001781	.000571
_cons	.9783124	.009262	105.63	0.000	.9594463	.9971785

The adjusted  $R^2$  is equal to 0.8628 which indicates that 86.28% of the variability of the delamination at exit is explained by the cutting force and the thrust force. The MLR model is represented by:

$$\text{Dex} = 0.9783124 + 0.0003197 \cdot \text{CF} + 0.003745 \cdot \text{TF}$$

**Figure 3.99** is the plot of the multiple linear regression model of delamination at exit. The overall trend tends to rise. The plot indicates that the scatter of points lie closely to the regression line and are within the 95% confidence of interval of the predicted values. However, two outliers are

identified to likely indicate a lack fit of the model. Further investigations are needed into the assumptions to check the model's validity. **Table C.22 in annex C** presents the data of this plot.

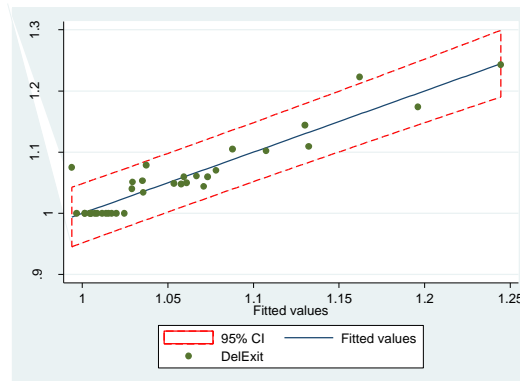


Figure 3.100: Plot of the delamination at exit using MLR

**Normality assumption:** The normal probability plot below shows a deviation of the straight line at the lower and upper tail. Also, as presented in the Shapiro-Wilk test (in **table 3.114**) the calculated p-value is 0, lower than 0.05. Consequently, the experimental data of the thrust force do not follow a normal distribution.

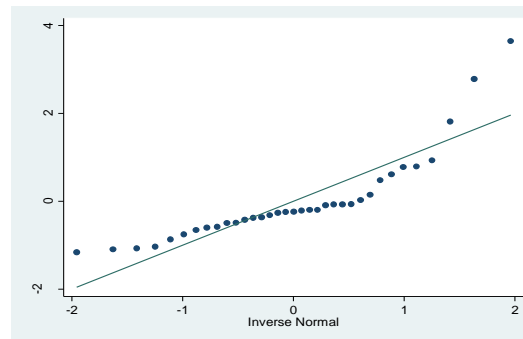


Figure 3.101: Normal probability plot for the delamination at exit using MLR

Table 3.114: Shapiro-Wilk Test for delamination at exit normality using MLR

Variable	Obs	W	V	z	Prob>z
rdelexit_c~d	35	0.79332	7.377	4.171	0.00002

**Homoscedasticity assumption:** The residual plot below confirms that the variance of the residual is constant and the residuals values are fluctuating randomly around zero. The BPCW test is performed and the calculated chi-squared corresponds to 0.04 and the p-value of the chi2 is 0.8488 higher than 0.05 which indicates that the null hypothesis that the variance of residuals is constant cannot be rejected.

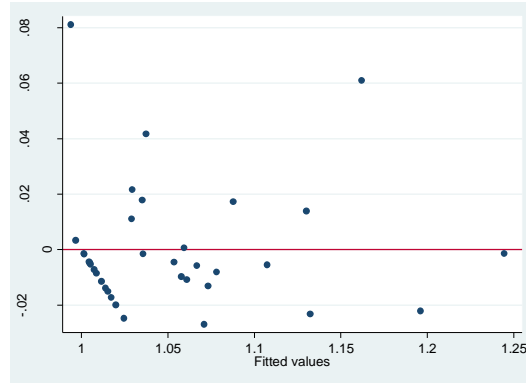


Figure 3.102: Residual plot for the delamination at exit data using MLR

**Test of collinearity:** The collinearity concept is the undesirable state where two variables are highly correlated with each other. This can lead to inaccurate and erroneous results because they are explaining almost the same variability in the output. However, it is beneficial to investigate the relationship of the output by excluding one of the input variables which was done in sections 3.11.1 and 3.11.2; the adjusted  $R^2$  of the thrust force and the cutting force when analyzed separately were respectively 86.25% and 80.42%. Therefore, the same variability is explaining the output. To be able to combine the thrust force and the cutting force within the same analytical model to predict properly the delamination at exit, the test of collinearity must be negative which means absence of correlation is needed for conclusive results.

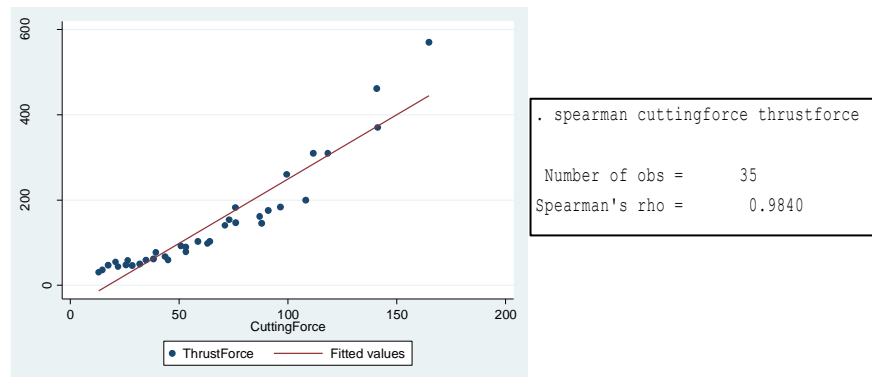


Figure 3.103: Test of collinearity

The visual illustration in **figure 3.103** displays that the cutting force and the thrust force are highly correlated. In 88.66% of the case, the delamination at exit is predicted in the same way. As indicated in **table 3.113**, both coefficients are very similar 0.003197 and 0.003745 for respectively the cutting force and the thrust force. To verify how the variables are correlated, the

spearman test can be used. In this case, the  $r_s$  calculated corresponds to 0.984 which indicated a high positive correlation. Thus, the model developed previously cannot be used due to this issue.

Researchers may use the following models to predict the delamination at exit:

- From the thrust force only:

$$\text{Dex} = 0.9855196 + 0.0004689 \cdot \text{TF}$$

- From the cutting force only:

$$\text{Dex} = 1.007646 - 0.0002544 \cdot \text{CF} + 0.0000108 \cdot \text{CF}^2$$

### 3.12 Delamination at entry analysis versus thrust force and cutting force

#### 3.12.1 Delamination at entry distribution study over thrust force

In this section, the thrust force is analyzed as the only effect on the delamination at entry. As a start, the parameter estimates, ANOVA and coefficient of determination results are generated to understand the effect of the thrust force on the delamination at entry. **Table 3.115** demonstrates that the calculated p-value is near 0, lower than 0.05 which indicates that the thrust force has a significant effect on the delamination at entry. The adjusted  $R^2$  corresponds to 0.6878 which indicates that 68.78% of the variability is explained by the thrust force.

Table 3.115: Parameter estimates and ANOVA results of delamination at entry against thrust force

Source	SS	df	MS	Number of obs = 35		
Model	.542130786	1	.542130786	F( 1, 33) =	75.90	
Residual	.235719122	33	.007143004	Prob > F	=	0.0000
				R-squared	=	0.6970
				Adj R-squared	=	0.6878
Total	.777849908	34	.022877938	Root MSE	=	.08452

delentry	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
thrustforce	.0009992	.0001147	8.71	0.000	.0007658	.0012325
_cons	1.015377	.0219312	46.30	0.000	.9707578	1.059996

The simple linear model is represented by:

$$\text{Dee} = 1.015377 + 0.0009992 \cdot \text{TF}$$

**Figure 3.104** is the plot of this simple linear regression of the thrust force against the delamination at entry. It indicates that the observed values follow the fitted line. **Table C.23 in annex C** presents the data for this plot.

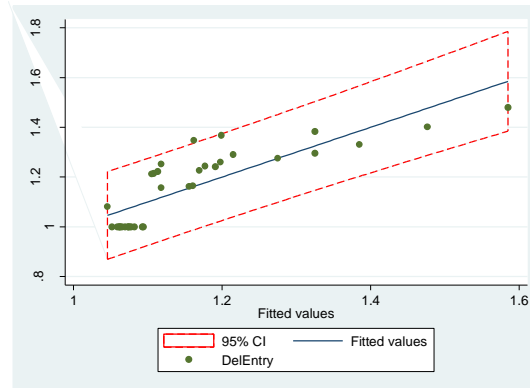


Figure 3.104: Plot of the thrust force against the delamination at entry

To verify if this model fits properly the delamination at entry data, the following assumptions are verified: normality, homoscedasticity and linearity.

**Normality assumption:** The normal probability plot below suggests a deviation of the straight line at the lower, center and upper tail. Also, the Shapiro-Wilk test is performed. As indicated in **table 3.116**, the p-value is 0.00818, lower than 0.05. Consequently, the delamination at entry does not follow a normal distribution.

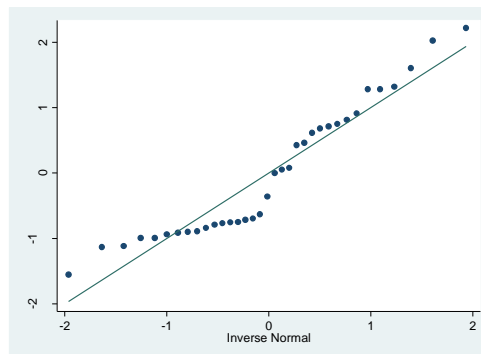


Figure 3.105: Normal probability plot for the delamination at entry using simple regression with thrust force

Table 3.116: Shapiro-Wilk Test for delamination at entry normality using simple regression

Variable	Obs	W	V	z	Prob>z
rdelentrytf	35	0.91150	3.159	2.401	0.00818

**Homoscedasticity assumption:** The plot below demonstrates that the variance of the residual is constant. No specific pattern is recognized. The BPCW test is performed and the calculated chi-squared corresponds to 0.01 and the p-value of the chi2 is 0.93; higher than 0.05 which confirms that the null hypothesis that the variance of residuals is constant cannot be rejected.

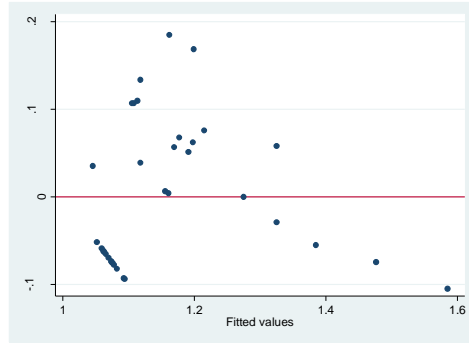


Figure 3.104: Residual plot for the delamination at entry data using simple regression

**Linearity assumption:** Figure 3.105 indicates that the “lowess” outline does not overlap the regression outline which indicates that the existent relationship between the thrust force and the delamination at exit is nonlinear.

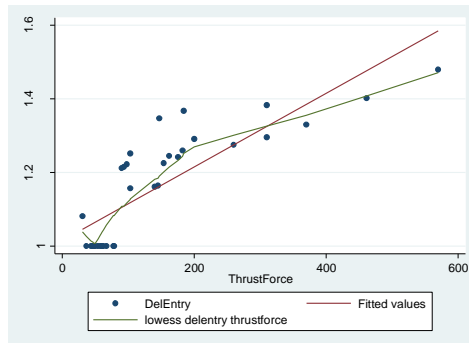


Figure 3.105: Scatterplot of thrust force against delamination at entry

Below is a summary of the assumptions' validation results for the delamination at entry linear regression model (Table 3.117).

Table 3.117: Illustration of the assumption validation results for delamination at entry with simple regression

Assumptions	Valid for this model
Normality	No
Homoscedasticity	Yes
Linearity	No

Based on this series of analysis, this model does not follow the normality assumption and is not linear. To determine an appropriate fitted model, the attempt of transforming the thrust force or the delamination at entry will be completed. From section 3.1.5, it was determined that the log transformation of the thrust force was a best choice than the linear whereas its p-value is the highest one to be over 0.05 and the highest one between all the suggested models. In the

following analysis, the log transformation is verified to construct the delamination at entry prediction model. As a standard procedure, the parameter estimates, ANOVA and coefficient of determination results are generated to analyze the fitted model and the independent variable effect. **Table 3.118** indicates that the logarithmic term of the thrust force is significant: its p-value is near 0; lower than 0.05. Also, the adjusted  $R^2$  (81.72%) is higher after transforming the independent variable compared to the previous regression analysis (68.78%).

Table 3.118: Parameter estimates and ANOVA results of transformed thrust force against delamination at entry

Source	SS	df	MS	Number of obs = 35		
Model	.639868979	1	.639868979	F( 1, 33) =	153.03	
Residual	.137980929	33	.00418124	Prob > F =	0.0000	
				R-squared =	0.8226	
				Adj R-squared =	0.8172	
Total	.777849908	34	.022877938	Root MSE =	.06466	

delentry	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lntf	.1800144	.0145517	12.37	0.000	.1504087	.20962
_cons	.3179846	.0689648	4.61	0.000	.1776746	.4582946

The fitted regression model is represented by:

$$\text{Den} = 0.3179846 + 0.1800144 * \log (\text{Tf})$$

**Figure 3.106** is a graphical illustration of delamination at entry observed vs predicted values against the logarithmic thrust force. Also, **table C.24 in annex C** presents the data of this plot.

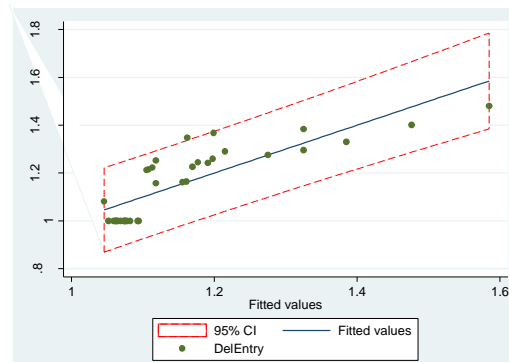


Figure 3.106: Plot of the delamination at entry observed vs predicted values against thrust force transformation

**Normality assumption:** As demonstrated in **figure 3.107**, the graphical representation suggests a deviation of the normal distribution. Also, the Shapiro-Wilk test confirms that the standardized residuals are not normally distributed as p-value corresponds to 0.00818, lower than 0.05.

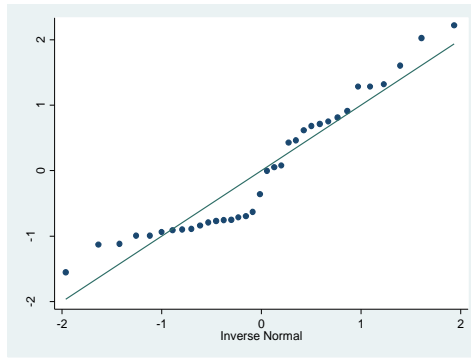


Figure 3.107: Normal plot of the standardized residuals of the delamination at entry

Table 3.119: Shapiro-Wilk test for normal data for delamination at entry

Variable	Obs	W	V	z	Prob>z
rdIntf	35	0.91150	3.159	2.401	0.00818

**Linearity assumption:** Figure 3.108 indicates that the relationship between the delamination at entry and the log (thrust force) is linear.

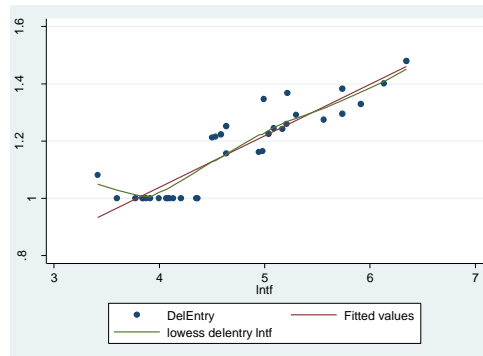


Figure 3.108: Scatterplot of the delamination at entry against the log (thrust force)

**Homoscedasticity assumption:** Figure 3.109 displays that the variance is constant and no specific pattern can be identified. After performing the BPCW, the calculated chi-squared value corresponds to 0.01 and the p-value is at 0.93, higher than 0.05 indicating that  $H_0$  cannot be rejected. In this case, the variance is constant.

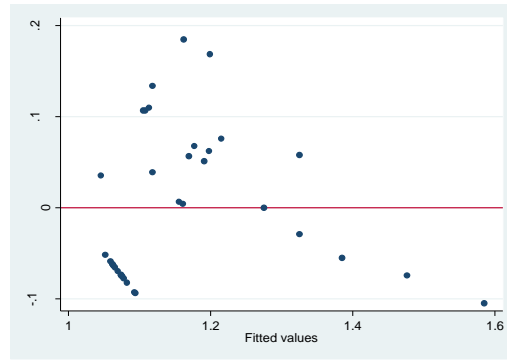


Figure 3.109: Residuals plot for circularity at exit

**Table 3.120** presents a brief description of each assumption's validation.

Table 3.120: Illustration of the assumption validation results

Assumptions	Valid for this model
Normality	No
Linearity	Yes
Homoscedasticity	Yes

Hence, this model can be used to predict the delamination at entry from the thrust force.

### 3.12.2 Delamination at entry distribution study over cutting force

In this section, the cutting force is analyzed over the delamination at entry. As shown in **table 3.121**, the p-value of the cutting force is lower than 0 which is lower than 0.05, which indicates that the cutting force does have an important effect on the delamination at entry trend. The adjusted  $R^2$  is equal to 0.8057 which indicates that only 80.57% of the variability of the delamination at entry is explained by the cutting force.

Table 3.121: Parameter estimates and ANOVA results of cutting force against delamination at entry

Source	SS	df	MS	Number of obs = 35	
Model	.63117225	1	.63117225	F( 1, 33) =	142.00
Residual	.146677658	33	.004444778	Prob > F =	0.0000
Total	.777849908	34	.022877938	R-squared =	0.8114
				Adj R-squared =	0.8057
				Root MSE =	.06667
delentry	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
cuttingforce	.003448	.0002893	11.92	0.000	.0028593 .0040366
_cons	.9350548	.0220094	42.48	0.000	.8902764 .9798333

The simple regression model is represented by:

$$DEE = 0.930548 + 0.003448*CF$$

**Figure 3.110** is the plot of this simple linear regression of the cutting force against the delamination at entry. The overall trend tends to rise. The plot indicates that the scatter of points lie closely to the regression line and are within the 95% confidence of interval of the predicted values. **Table C.25 in annex C** presents the data of this plot.

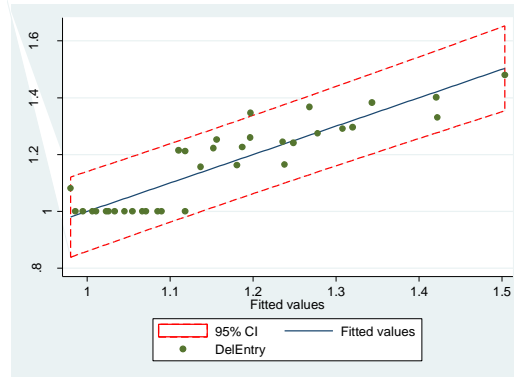


Figure 3.110: Plot of the cutting force against the delamination at entry

**Normality assumption:** The normal probability plot above shows that the overall trend follows the straight line. This indicates that the data do follow a normal distribution. As indicated in **table 3.122**, the calculated p-value is 0.27282, higher than 0.05. Consequently, this test confirms the plot observation.

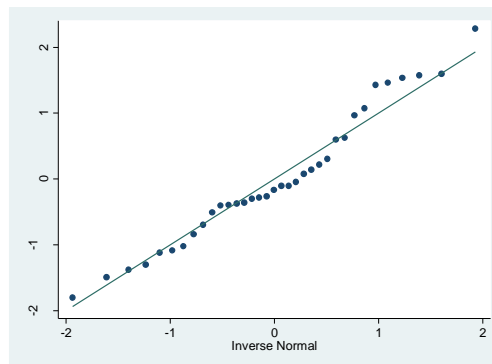


Figure 3.111: Normal probability plot for the delamination at entry vs cutting force data using simple regression

Table 3.122: Shapiro-Wilk Test for cutting force normality using simple regression

Variable	Obs	W	V	z	Prob>z
rdeleentrycf	35	0.96258	1.336	0.604	0.27282

**Homoscedasticity assumption:** The residual plot below confirms that the variance of the residual is constant. The BPCW test is performed and the calculated chi-squared corresponds to

0.03 and the p-value of the  $\chi^2$  is 0.8527 higher than 0.05 which indicates that the null hypothesis that the variance of residuals is constant cannot be rejected.

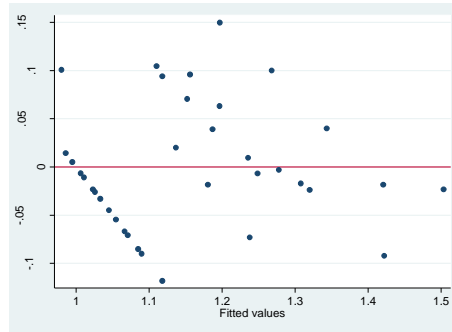


Figure 3.112: Residual plot for the delamination at entry versus cutting force data using simple regression

**Linearity assumption:** The scatterplot of the cutting force against the delamination at entry is illustrated in **figure 3.113**. The “lowess” fitted line suggests a linear shape.

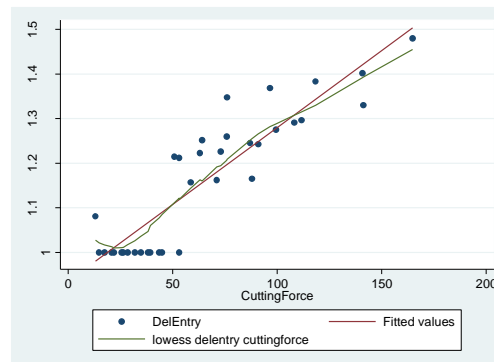


Figure 3.113: Scatterplot of cutting force against delamination at entry

Below is a summary of the assumptions' validation results for the delamination at exit simple regression model (**Table 3.124**). Based on this series of analysis, this model can be used by the experts to predict the delamination at entry.

Table 3.124: Illustration of the assumption validation results for delamination at entry vs cutting force using simple regression

Assumptions	Valid for this model
Normality	Yes
Homoscedasticity	Yes
Linearity	Yes

**Table 3.123** explores all the investigations described in this chapter concerning the delamination at entry.

Table 3.123: Presentation of all analysis completed for the delamination at entry against thrust force and cutting force

Method used	Validity to use	Variability	Significant variables
Simple linear regression with thrust force	Non-normality and non-linear	68.78%	Thrust force
Nonlinear Regression with Thrust force	All valid	81.72%	Logarithmic term of the thrust force
Simple linear regression with cutting force	All valid	80.57%	Cutting force

Researchers may use the following models to predict the delamination at entry:

- From the thrust force:

$$\text{Dee} = 0.3179846 + 0.1800144 * \log(\text{TF})$$

- From the cutting force:

$$\text{Dee} = 0.930548 + 0.003448 * \text{CF}$$

## CHAPTER 4 CONCLUSION

An aircraft is built with thousands of rivets and bolts hovering with thin metal sheets. They are inserted through holes which are drilled by the manufacturer based on exact specifications. Working away from the manufacturing requirements and practices may affect the structural integrity of the aircraft. The end result of the hole's quality is based on many factors: feed rate, spindle speed, tool wear, depth of the cut, drill-bit diameter, etc.

In the preceding chapters, the comprehensive principles of the simple and multiple linear regression analysis have been described and applied to construct a suitable fitted model based on the mechanics of drilling for each one of the following outputs:

- |                          |                                   |
|--------------------------|-----------------------------------|
| 1. Thrust force          | 6. Surface roughness (in microns) |
| 2. Cutting force         | 7. Diameter error at exit (in %)  |
| 3. Torque                | 8. Diameter error at entry (in %) |
| 4. Delamination at entry | 9. Circularity at exit (in %)     |
| 5. Delamination at exit  | 10. Circularity at entry (in %)   |

The inputs under study are the feed rate (micron/rev) and the spindle speed (rpm) at their respective levels [20, 60, 100, 200, 400, 600, 800] and [1500, 5000, 8500, 12000, 15000]. The following is a summary of the main aspects of the simple and multiple regression analysis:

1. Explore the distribution of each output. Then, explore the relationship between the independent variables and the dependents. The simple linear regression provides the contribution of each independent variable to the prediction of the output by analyzing the parameter estimates, ANOVA and coefficient of determination results.
2. Discover the impact of the combined independent variables and their interaction effect on the dependent. In fact, a prediction model will be fitted through MLR. Thus, the important of the feed rate, spindle speed and interaction are judged by comparing their relative p-values to the conventional reference point ( $\alpha=0.05$ ).
3. Generate the regression fitted model by pulling the regression coefficients from the parameter estimates table.
4. Compute the adjusted coefficient of determination. It represents the percent of the data that is the closest to the best fitted line.

5. Examine the plot of the estimated model to visually inspect the fit of the model developed. If the regression line passes roughly through every point on the plot, the model would be able to explain all of the variation. The further the line is away from the predicted points, the less the model is able to explain the variability.
6. Validate the fitted model with assumptions verification.

In today's economic-oriented and competitive marketplace, manufacturing practices simulation based on the mechanics of the drilling helps achieve a better drilling quality and therefore, predict the quality in advance. During this research, the common significant independent variable revealed was the feed rate. The dependence on the spindle speed of the different outputs is not observed on all outputs. **Table 4.1** displays the relationship between each independent variable and the outputs.

Table 4.1: Relationship between independent variables and outputs

Output	Effect increasing feed	Effect the increasing spindle
Thrust force	Linear increase	Linear decrease
Cutting force	Linear increase	Linear decrease
Torque	Linear increase	Linear decrease
Delamination at entry	Cubic increase	Cubic with slightly decrease
Delamination at exit	Linear increase	Linear with decrease
Surface roughness	Linear increase	Linear slightly increase
Diameter error at exit	Linear slightly decrease	Linear slightly increase
Diameter error at entry	Linear slightly decrease	Linear slightly increase
Circularity at exit	Curvilinear bell-shaped	Linear increase
Circularity at entry	Linear slightly increase	None

As explained in **chapter 2**, the settings can be transformed to a uniform coded scale:

Table 4.2: Coded values for the feed rate and spindle seed settings

	Coded (-1)	Coded (0)	Coded (+1)
Feed rate	20	410	800
Spindle speed	1500	6750	15000

**Table 4.3** summarizes the conclusions and the models found for each output. The mathematical models presented in this table are the real settings for the feed rate and the spindle speed. The feed rate ranges from 20 to 800 and the spindle speed from 1500 to 15000. For some outputs as shown in the table below, new models are presented with the coded scale above. These models are centered to 0.

Table 4.3: Models for the outputs studied

Output - type of model	Significant Variables	Variability	Mathematical model to use
Thrust force - Transformation	Feed rate and spindle speed	90.46%	$\text{Log (Tf)} = 4.291 + 0.00247 * F - 0.000045 * SS$
Cutting force - MLR with interaction	Feed rate and interaction	94.78%	$Cf = 29.335 + 0.184 * F - 0.0004 * SS - 0.000007 * F * SS$ $Cf_{centered} = 65.34 + 0.184 * F - 0.0004 * SS - 0.000007 * F * SS$
Torque - MLR with interaction	Feed rate and interaction	94.78%	$T = 146.68 + 0.92 * F - 0.0019 * SS - 0.000035 * F * SS$ $T_{centered} = 326.7 + 0.92 * F - 0.0019 * SS - 0.000035 * F * SS$
Delamination at entry - MLR with interaction	Feed rate and interaction	85.28%	$\text{Den} = 1 + 0.00063 * F + 1.02 \times 10^{-6} * SS - 1.7 \times 10^{-8} * F * SS$ $\text{Den}_{centered} = 1.160 + 0.00063 * F + 1.02 \times 10^{-6} * SS - 1.7 \times 10^{-8} * F * SS$
Delamination at exit - MLR with interaction	Feed rate, spindle speed and interaction	89.44%	$\text{Dex} = 0.98 + 0.00036 * F + 2.8 \times 10^{-6} * SS - 2.3 \times 10^{-8} * F * SS$ $\text{Dex}_{centered} = 1.054 + 0.00036 * F + 2.8 \times 10^{-6} * SS - 2.3 \times 10^{-8} * F * SS$
Surface roughness - Transformation with n=34	Feed rate and interaction	63.77%	No model retained because of outliers
Diameter error at exit - MLR with n=34	Feed rate and spindle speed	38.28%	No model retained because of outliers
Diameter error at entry - MLR with n=34	Feed rate and spindle speed	51.59%	No model retained because of outliers
Circularity at exit - no model	None	NA	No model retained because of outliers and fundamental assumptions are violated
Circularity at entry - no model	None	NA	No model retained; no significant variables found

During this drilling process review and evaluation, the regression method allowed the researchers to mathematically express five of the ten outputs under study to predict their behavior. However, experts are aiming to explain as much variance as possible while controlling as many additional factors as possible. Also, the goal was to be able to predict the ten outputs for a better process control.

When diagnosing the ANOVA results, a coefficient of determination helped define how much variance is explained. Any unexplained variance may be due to random or unconsidered variables. So far, only the feed rate and the spindle speed were explored. Furthermore, additional variables may be added to the original experiment effect to reduce the model variability like the tool wear.

The application of this project can be used to the NRC engineers to understand the relationship between the different variables in the drilling process from a statistical/mathematical point of view. Also, whenever accurate models are found valid using the regression method, the engineers will be able to define the best combination for the optimum to minimize the quality defects while drilling a hole whereas decreases the experimental cost. Because regression models haven't able to be found for five of the ten variables, one of the continuations of this project is to explore the artificial intelligence methods that are proven to be effective by many researchers in different fields (electric, civil and mechanical engineering and medical), such as the:

- ANN: the Artificial Neuron Network [34], a nonlinear statistical data modeling tool, based on neuron networks able to learn and recognize data patterns.
- LAD: the Logical Analysis of Data [35], a data analysis methodology which extracts patterns from a large set of data, first established by Peter L. Hammer in 1986.
- SVM: the Support Vector Machines is a modeling and prediction tool, founded by Vapnik [36] a Russian mathematician in the early 1960s based on the structural risk minimization principle.

These methods will be able to combine the controllable and uncontrollable variables to predict the desired outputs. Also, during their pattern's study, it will recognize the mechanical failure to avoid them.

## REFERENCES

1. L.P. Kollar and G.S. Springer, *Mechanics of composite structure*, Cambridge: Cambridge University Press, 2003.
2. W.C. Chen and C.C. Tsao, "Cutting performance of different coated twist drills," *J Mater Process Technol*, volume 88, issue 1-3, page 203-207, 1999
3. J.Y. Sheikh-Ahmad, *Machining of Polymer Composites*, New York: Springer, 2009.
4. UA Khashaba, "Drilling of polymer matrix composites: A review", *Journal of Composite Materials*, volume 47, issue 15, pages 1817-1832, 2012.
5. J. P. Davim, R. J. Campos and A. M. Abrao, "A Novel Approach Based on Digital Image Analysis to Evaluate the Delamination Factor after Drilling Composite Laminates," *Composites Science and Technology*, volume 67, issue 9, pages 1939-1945, 2007.
6. O. Focke, M. Kalms, C.V. Kopylow and W Jueptner, "NDT of Free Formed CFRP Composites with Laser Ultrasonic," in *12th Asia Pacific Conference on NDT*, Auckland, New Zealand, 2006.
7. C. C. Tsao and H. HoCheng, "Computerised Tomography and C-Scan for Measuring Delamination in the Drilling of Composite Materials Using Various Drills," *International Journal of Machine Tools and Manufacture*, volume 45, issue 11, pages 1282-1287, 2005.
8. M.A. Seif, U.A. Khashaba and R. Rojas-Oviedo, "Measuring Delamination in Graphite/Epoxy Composites Using a Shadow Moire Laser Based Imaging," *Composite Structures*, volume 79, issue 1, pages 113-118, 2007
9. C.C. Tsao, H. Hocheng, "Evaluation of thrust force and surface roughness in drilling composite material using Taguchi analysis and neural network", *Journal of Materials Processing Technology*, Volume 203, Issues 1–3, Pages 342-348, 2008.
10. Engineer's Edge, "Machinery's Handbook," Engineer's Edge [Online] Available at: <http://www.engineersedge.com/manufacturing/images/texture.gif>.  
[Visited on 24/09/2011]

11. M. Jahan, W. Yoke San and M. Rahman, "A comparative experimental investigation of deep-hole micro-EDM drilling capability for cemented carbide (WC-Co) against austenitic stainless steel (SUS 304)," *International Journal of Advanced Manufacturing Technology*, volume 46, issue 9-12, pages 1145-1160, 2010.
12. B. Srinivasa Rao, R. Rudramoorthy, S. Srinivas and B. Nageswara Rao, "Effect of drilling induced damage on notched tensile and pin bearing strengths of woven GFR-epoxy composites," *Materials Science and Engineering: A*, volume 472, issues 1-2, pages 347-352, 2008.
13. S. Jayabal, U. Natarajan, U. Sekar, *Regression modeling and optimization of machinability behavior of glass-coir-polyester hybrid composite using factorial design methodology*, London: Springer, 2010.
14. S. Jayabal and U. Natarajan, "Modelling and optimisation of thrust force, torque and tool wear," *Int. J. of Machining and Machinability of Materials*, volume 9, issue 1-2, pages 149-172, 2011.
15. Reliability Engineering Resources, "Box-Behnken Designs for Optimizing Product Performance" [Online] Available at:  
<http://www.weibull.com/hotwire/issue130/hottopics130.htm>  
 [Visited on 02/01/2013].
16. B. Latha and V. S. Senthilkumar, "Analysis of Thrust Force in Drilling Glass Fiber-Reinforced Plastic Composites Using Fuzzy Logic," *Materials and Manufacturing Processes*, volume 24, issue 4, pages 509-516, 2009.
17. E. U. Enemuoh, A. S. El-Gizawi and A. C. Okafor, "An approach for development of damage-free drilling of carbon fiber reinforced thermosets", *International Journal of Machine Tools and Manufacture*, volume 41, issue 12, pages 1795-1814, 2001.
18. D.C. Montgomery, *Design and Analysis of Experiments*, 7<sup>th</sup> edition, Hoboken: John Wiley & Sons, 2009.
19. J. S. Kwak, "Application of Taguchi and response surface methodologies for geometric error in surface grinding process," *Int.J.Mach.Tools Manuf*, volume 45, issue 3, pages 327-334, 2005.
20. S. Kalidass, P. Palanisamy And V. Muthukumaran, "Prediction of tool wear using regression and artificial neural network models in end milling of AISI 304 Austenitic

Stainless Steel,” *International Journal of Engineering and Innovative Technology*, volume 1, issue 2, pages 29-35, 2012.

21. P. Orth, S. Yacout and L. Adjengue, “Accuracy and robustness of decision making techniques in condition based maintenance,” *Journal of Intelligent Manufacturing*, volume 23, issue 2, pages 255-264, 2012.
22. Li, P. (2003). “An empirical comparison between SVMs and ANNs for speech recognition” [Online]. Available at:  
<http://www.cs.rutgers.edu/~mlittman/courses/ml03/iCML03/papers/li.pdf>  
 [Visited on 24/09/2011].
23. A. K. Gupta, “Predictive modelling of turning operations using response surface methodology, artificial neural networks and support vector regression,” *International Journal of Production Research*, volume 48, issue 3, pages 763-778, 2010.
24. A. Mayyas, A. Qasaimeh, K. Alzoubi, S. Lu, M.T. Hayajneh and A. M. Hassan, “Modeling the Drilling Process of Aluminum Composites Using Multiple Regression Analysis and Artificial Neural Networks,” *Journal of Minerals and Materials Characterization and Engineering*, volume 11, issue 10, pages 1039-1049, 2012.
25. G. K. Smyth, “Nonlinear regression,” in *Encyclopedia of Environmetrics*, volume 3, Hoboken: John Wiley & Sons, 2002, pages 1405-1411.
26. D.C. Montgomery, *Design and Analysis of Experiments*, 4<sup>th</sup> edition, Hoboken: John Wiley & Sons, 1997.
27. J. Neter, M. Kutner, C. Nachtsheim, W. Wasserman, *Applied Linear Statistical Models*, 4<sup>th</sup> Edition, Chicago: Irwin, 1996.
28. R.V. Hogg and J. Ledolter, *Applied Statistics for Engineers and Physical Scientists*, 2<sup>nd</sup> edition, New York: Macmillan Publishing Company, 1992.
29. Military Handbook “Polymer Matrix Composites Guidelines for Characterization of Structural Materials,” U.S. Department of Defense, Washington DC, Handbook, MIL-HDBK-17-1F, 2002 [Online]. Available at:  
<http://app.knovel.com/hotlink/toc/id:kpMHMILH5M/military-handbook-mil-3>  
 [Visited on 20/11/2012].

30. Burnham, K. P., and D. R. Anderson. *Model Selection and Multimodel Inference: a practical information-theoretic approach*, 2nd edition. New York: Springer-Verlag, 2002.
31. R.K. Ranjit, *Design of Experiments Using The Taguchi Approach*, Hoboken: John Wiley & Sons, 2001.
32. S. Rawat, H. Attia, *Wear mechanics and tool life management of WC-Co drills during dry high speed drilling of woven carbon fibre composites*, Science Direct, 2009.
33. Institute for the Digital Research and Education. "Regression with Stata".  
[Online] Available at:  
<http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter2/statareg2.htm>  
[Visited on 03/15/2012].
34. J. Livingston, *Artificial Neural Networks: Methods and Applications*, Totowa: Humana Press, 2008.
35. I. Chikalov, V. Lozin, I. Lozina, M. Moshkov, H. S. Nguyen, A. Skowron, B. Zielosko, *Three Approaches to data analysis*, New York: Springer, 2013.
36. A. Christmann, I. Steinwart. *Support Vector Machines*, New York: Springer, 2008.

## ANNEX A

Below is a recommended outline of the steps to build an appropriate design of experiment [1]:

1. **Recognition of and statement of the problem:** formulate a concise description of the problem and the objectives of the experiment.
2. **Choice of key factors and levels:** determine the significant factors on the outputs and the settings to be used on each factor. It is generally best to keep the numbers of factors levels low (most used is 2 and 3).
3. **Selection of the response variable:** select the outputs to be analyzed.
4. **Choice of experimental design:** determine the design to use and the number of runs to accomplish. This selection involves the sample size and the number of factors, and their settings.
5. **Performing the experiments:** execute each run randomly within the same production environment and record the results.
6. **Data analysis:** analyze the data collected with statistical methods and graphical representations.
7. **Conclusions and recommendations:** document and summarize the results in tabular and graphical forms. Also, present the recommendations and link the conclusions to the problems described in step 1.

---

<sup>1</sup> D. C. Montgomery, Introduction to Statistical Quality Control, 6<sup>th</sup> edition, Hoboken: John Wiley & Sons, 2012

## ANNEX B

### Note 1:

Difference between “Simple” and “Linear” terminologies [2]:

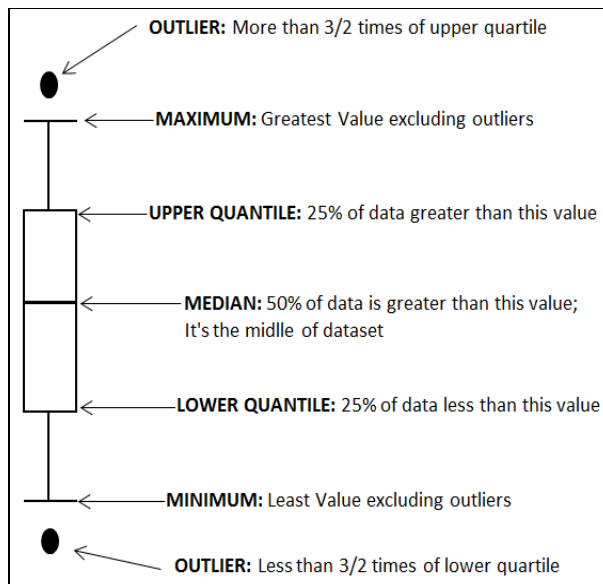
“Simple” implies that the model only contains one predictor variable, and

“Linear” denotes that no parameter appears as an exponent or is multiplied or divided by another parameter and the variable appears only in the first power; it does not refer to the shape of the response surface.

A model that is simple linear is also called “first-order model”.

### Note 2:

Figure B.1: Boxplot representation



### Note 3:

An ANN process is defined by the following:

1. The definition of the most adequate input variables for a most relevant outputs,
2. The training process to learn from the data in order to obtain the expected results,
3. The activation function that predicts the output using an appropriate mapping function,

---

<sup>2</sup> J. Neter, M. Kutner, C. Nachtsheim, W. Wasserman, *Applied Linear Statistical Models*, 4<sup>th</sup> Edition, Chicago: Irwin, 1996.

4. The validation of the predicted values by analyzing the errors, and
5. The selection of the optimal set of values for the input factors that lead to higher quality of outputs.

**Note 4:**

Table B.2: Experimental matrix for a Box-Behnken design for three factors

Run	Factors		
	Factor 1	Factor 2	Factor 3
1	-1	-1	0
2	1	-1	0
3	1	-1	0
4	1	1	0
5	-1	0	-1
6	-1	0	1
7	1	0	-1
8	1	0	1
9	0	-1	-1
10	0	-1	1
11	0	1	-1
12	0	1	1
13	0	0	0
14	0	0	0
15	0	0	0
16	0	0	0
17	0	0	0

**Note 5:**

To determine how much impact each observation has on a particular predictor, the software calculates a measure called DFBETAs. The DFBETA for a predictor and for a specific observation is the difference between the regression coefficient calculated for all of the data and the regression coefficient calculated with the observation deleted, scaled by the standard error calculated with the observation deleted. DFBETAs can be either positive or negative. Any value exceeding the  $\frac{2}{\sqrt{n}}$  where  $n$  is the number of observations, needs further investigation.

## ANNEX C

Table C.1: Experiments data collection [3]

Feed	SpindleSpeed	DelExit	DelEntry	HoleSurfRough	HoleDiaErrExit	HoleDiaErrEntry	HoleCirExit	HoleCirEntry	ThrustForce	CuttingForce	Torque
20	1500	1.000	1.000	0.453	0.032	0.016	0.056	0.074	58.23	26.38	131.90
20	5000	1.000	1.000	1.194	0.138	0.020	0.128	0.112	54.26	20.76	103.80
20	8500	1.000	1.000	2.255	0.182	0.044	0.142	0.118	47.00	17.34	86.70
20	12000	1.000	1.000	3.282	0.088	0.072	0.402	0.202	36.50	14.72	73.60
20	15000	1.075	1.081	2.030	4.800	5.650	0.484	0.594	30.43	13.10	65.50
60	1500	1.000	1.000	0.601	-0.006	-0.104	0.072	0.108	77.50	39.36	196.80
60	5000	1.000	1.000	0.900	-0.004	-0.104	0.154	0.122	58.96	34.72	173.60
60	8500	1.000	1.000	0.720	-0.026	-0.082	0.158	0.136	46.50	28.42	142.10
60	12000	1.000	1.000	0.820	0.038	-0.052	0.170	0.122	43.50	21.98	109.90
60	15000	1.000	1.000	1.060	0.064	0.156	0.104	0.094	48.00	25.56	127.80
100	1500	1.000	1.000	0.972	-0.084	-0.114	0.086	0.168	78.50	53.10	265.50
100	5000	1.000	1.000	1.160	-0.050	-0.104	0.132	0.156	60.00	44.95	224.75
100	8500	1.000	1.000	0.990	-0.034	-0.098	0.156	0.152	67.00	43.53	217.66
100	12000	1.000	1.000	1.200	0.014	-0.068	0.172	0.136	62.30	38.20	191.00
100	15000	1.000	1.000	1.140	0.004	0.210	0.176	0.116	50.00	31.91	159.53
200	1500	1.079	1.252	1.429	-0.052	-0.138	0.128	0.136	103.00	64.06	320.31
200	5000	1.053	1.223	1.305	-0.060	-0.120	0.134	0.166	98.00	63.06	315.30
200	8500	1.051	1.215	1.820	-0.040	-0.064	0.164	0.200	93.00	50.84	254.20
200	12000	1.040	1.212	1.890	-0.046	-0.064	0.178	0.144	90.00	53.10	265.50
200	15000	1.034	1.157	1.870	0.034	0.100	0.154	0.130	103.00	58.58	292.90
400	1500	1.105	1.291	1.620	-0.074	-0.164	0.160	0.216	200.00	108.15	540.75
400	5000	1.060	1.242	1.757	-0.050	-0.110	0.166	0.158	175.65	90.96	454.80
400	8500	1.061	1.245	2.243	-0.070	-0.068	0.182	0.220	162.00	87.16	435.80
400	12000	1.060	1.226	2.140	-0.090	-0.048	0.198	0.158	154.00	73.04	365.20
400	15000	1.049	1.162	2.273	-0.036	0.090	0.208	0.136	140.24	71.14	355.70
600	1500	1.223	1.330	2.000	-0.116	-0.284	0.172	0.250	370.00	141.3	706.50
600	5000	1.144	1.296	1.783	-0.090	-0.144	0.188	0.208	310.00	111.64	558.20
600	8500	1.102	1.275	2.782	-0.136	-0.124	0.204	0.152	260.00	99.46	497.30
600	12000	1.044	1.260	2.256	-0.094	-0.090	0.238	0.162	182.50	75.86	379.30
600	15000	1.050	1.165	1.912	-0.024	0.030	0.312	0.152	145.50	87.92	439.60
800	1500	1.243	1.480	2.492	-1.090	-1.004	0.178	0.198	570.00	164.78	823.90
800	5000	1.174	1.402	1.957	-1.090	-0.968	0.413	0.453	461.40	140.82	704.10
800	8500	1.109	1.383	3.036	-0.068	-0.150	0.556	0.244	310.00	118.36	591.80
800	12000	1.070	1.368	2.937	-0.042	-0.084	0.648	0.202	184.00	96.60	483.00
800	15000	1.048	1.347	2.198	-0.022	0.036	0.756	0.184	147.00	76.02	380.11

<sup>3</sup> Rawat, S. (2006), *The Characterization of Drilling Process of Woven Composites Using Machinability Maps*, Thesis, McGill University, Montreal, QC, Canada.

Table C.2: Thrust force fitted against observed values using MLR

Number	Observed	Fitted
1	58.23	95.81065
2	54.26	65.53903
3	47	35.26741
4	36.5	4.995792
5	30.43	20.95131
6	77.5	110.7649
7	58.96	80.4933
8	46.5	50.22168
9	43.5	19.95006
10	48	5.997041
11	78.5	125.7192
12	60	95.44757
13	67	65.17596
14	62.3	34.90433
15	50	8.95723
16	103	163.1049
17	98	132.8333
18	93	102.5616
19	90	72.29001
20	103	46.34291
21	200	237.8762
22	175.65	207.6046
23	162	177.333
24	154	147.0614
25	140.24	121.1143
26	370	312.6476
27	310	282.3759
28	260	252.1043
29	182.5	221.8327
30	145.5	195.8856
31	570	387.4189
32	461.4	357.1473
33	310	326.8757
34	184	296.6041
35	147	270.657

Table C.3: Thrust force fitted against observed values using MLR with interaction

Number	Observed	Fitted
1	58.23	20.95284
2	54.26	28.65257
3	47	36.35231
4	36.5	44.05204
5	30.43	50.65181
6	77.5	46.18171
7	58.96	48.66969
8	46.5	51.15767
9	43.5	53.64565
10	48	55.7782
11	78.5	71.41058
12	60	68.68681
13	67	65.96304
14	62.3	63.23926
15	50	60.90459
16	103	134.4828
17	98	118.7296
18	93	102.9764
19	90	87.22328
20	103	73.72057
21	200	260.6271
22	175.65	218.8152
23	162	177.0033
24	154	135.1913
25	140.24	99.35252
26	370	386.7715
27	310	318.9008
28	260	251.0301
29	182.5	183.1594
30	145.5	124.9845
31	570	512.9158
32	461.4	418.9864
33	310	325.0569
34	184	231.1274
35	147	150.6164

Table C.4: Table fitted values vs observed for the log-thrust force transformation

Number	Observed	Fitted
1	4.064401	-37.58065
2	3.993787	-1.27903
3	3.850147	11.73259
4	3.597312	31.50421
5	3.415429	51.38131
6	4.350278	-33.26492
7	4.076859	-21.5333
8	3.839452	-3.721683
9	3.772761	23.54994
10	3.871201	53.99704
11	4.363099	-47.21919
12	4.094345	-35.44757
13	4.204693	1.824047
14	4.131961	27.39567
15	3.912023	41.04277
16	4.634729	-60.10487
17	4.584968	-34.83325
18	4.532599	-9.561629
19	4.49981	17.70999
20	4.634729	56.65709
21	5.298317	-37.87622
22	5.168493	-31.95461
23	5.087596	-15.33298
24	5.036952	6.93864
25	4.943355	19.12575
26	5.913503	57.35243
27	5.736572	27.62405
28	5.560682	7.895668
29	5.20675	-39.33271
30	4.980176	-50.38561
31	6.345636	182.5811
32	6.134265	104.2527
33	5.736572	-16.87568
34	5.214936	-112.6041
35	4.990433	-123.657

Table C.5: Observed vs predicted values for cutting force using MLR

Number	Observed	Fitted
1	26.38	46.11267
2	20.76	37.25069
3	17.34	28.38871
4	14.72	19.52673
5	13.1	11.93075
6	39.36	51.14958
7	34.72	42.2876
8	28.42	33.42562
9	21.98	24.56364
10	25.56	16.96766
11	53.1	56.1865
12	44.95	47.32452
13	43.53	38.46254
14	38.2	29.60056
15	31.91	22.00458
16	64.06	68.77879
17	63.06	59.91681
18	50.84	51.05482
19	53.1	42.19284
20	58.58	34.59686
21	108.15	93.96336
22	90.96	85.10138
23	87.16	76.2394
24	73.04	67.37742
25	71.14	59.78144
26	141.3	119.1479
27	111.64	110.2859
28	99.46	101.424
29	75.86	92.56199
30	87.92	84.96601
31	164.78	144.3325
32	140.82	135.4705
33	118.36	126.6085
34	96.6	117.7466
35	76.02	110.1506

Table C.6: Cutting force table fitted vs observed values using MLR with interaction

Number	Observed	Fitted
1	26.38	32.2326
2	20.76	30.41123
3	17.34	28.58987
4	14.72	26.76851
5	13.1	25.20734
6	39.36	39.17462
7	34.72	36.38689
8	28.42	33.59917
9	21.98	30.81145
10	25.56	28.42197
11	53.1	46.11664
12	44.95	42.36256
13	43.53	38.60848
14	38.2	34.8544
15	31.91	31.63661
16	64.06	63.4717
17	63.06	57.30172
18	50.84	51.13174
19	53.1	44.96176
20	58.58	39.67321
21	108.15	98.18181
22	90.96	87.18004
23	87.16	76.17826
24	73.04	65.17648
25	71.14	55.7464
26	141.3	132.8919
27	111.64	117.0583
28	99.46	101.2248
29	75.86	85.39121
30	87.92	71.81958
31	164.78	167.602
32	140.82	146.9367
33	118.36	126.2713
34	96.6	105.6059
35	76.02	87.89277

Table C.7: Torque fitted values versus observed using MLR

Number	Observed	Fitted
1	131.9	230.5641
2	103.8	186.2536
3	86.7	141.9432
4	73.6	97.6327
5	65.5	59.65229
6	196.8	255.7488
7	173.6	211.4383
8	142.1	167.1278
9	109.9	122.8174
10	127.8	84.83695
11	265.5	280.9334
12	224.75	236.623
13	217.66	192.3125
14	191	148.002
15	159.53	110.0216
16	320.31	343.8951
17	315.3	299.5846
18	254.2	255.2741
19	265.5	210.9637
20	292.9	172.9833
21	540.75	469.8184
22	454.8	425.5079
23	435.8	381.1974
24	365.2	336.887
25	355.7	298.9066
26	706.5	595.7417
27	558.2	551.4312
28	497.3	507.1208
29	379.3	462.8103
30	439.6	424.8299
31	823.9	721.665
32	704.1	677.3546
33	591.8	633.0441
34	483	588.7336
35	380.11	550.753

Table C.8: Torque observed against predicted values using MLR with interaction

Number	Observed	Fitted
1	131.9	161.166
2	103.8	152.0574
3	86.7	142.9489
4	73.6	133.8404
5	65.5	126.0331
6	196.8	195.8759
7	173.6	181.9357
8	142.1	167.9956
9	109.9	154.0554
10	127.8	142.1067
11	265.5	230.5858
12	224.75	211.814
13	217.66	193.0422
14	191	174.2704
15	159.53	158.1803
16	320.31	317.3605
17	315.3	286.5096
18	254.2	255.6587
19	265.5	224.8078
20	292.9	198.3642
21	540.75	490.91
22	454.8	435.9009
23	435.8	380.8918
24	365.2	325.8827
25	355.7	278.732
26	706.5	664.4595
27	558.2	585.2922
28	497.3	506.1248
29	379.3	426.9575
30	439.6	359.0998
31	823.9	838.009
32	704.1	734.6835
33	591.8	631.3579
34	483	528.0323
35	380.11	439.4676

Table C.9: Delamination at entry fitted and observed values using MLR

Number	Observed	Fitted
1	1	1.04576
2	1	1.031407
3	1	1.017054
4	1	1.002701
5	1.81	0.9903983
6	1	1.065371
7	1	1.051018
8	1	1.036665
9	1	1.022312
10	1	1.010009
11	1	1.084982
12	1	1.070629
13	1	1.056276
14	1	1.041923
15	1	1.02962
16	1.252	1.134009
17	1.223	1.119656
18	1.215	1.105303
19	1.212	1.09095
20	1.157	1.078647
21	1.291	1.232063
22	1.242	1.21771
23	1.245	1.203357
24	1.226	1.189004
25	1.162	1.176701
26	1.33	1.330117
27	1.296	1.315764
28	1.275	1.301411
29	1.26	1.287058
30	1.165	1.274755
31	1.48	1.428171
32	1.402	1.413818
33	1.383	1.399465
34	1.368	1.385112
35	1.347	1.372809

Table C.10: Delamination at entry fitted and observed values using MLR with interaction

Number	Observed	Fitted
1	1	1.012679
2	1	1.015106
3	1	1.017533
4	1	1.019961
5	1.081	1.022041
6	1	1.036831
7	1	1.036955
8	1	1.037078
9	1	1.037202
10	1	1.037309
11	1	1.060982
12	1	1.058803
13	1	1.056623
14	1	1.054444
15	1	1.052576
16	1.252	1.12136
17	1.223	1.113423
18	1.215	1.105486
19	1.212	1.097549
20	1.157	1.090746
21	1.291	1.242117
22	1.242	1.222664
23	1.245	1.203211
24	1.226	1.183758
25	1.162	1.167084
26	1.33	1.362873
27	1.296	1.331905
28	1.275	1.300936
29	1.26	1.269967
30	1.165	1.243423
31	1.48	1.48363
32	1.402	1.441146
33	1.383	1.398661
34	1.368	1.356176
35	1.347	1.319761

Table C.11: Delamination at exit fitted and observed values using MLR

Number	Observed	Fitted
1	1	1.033303
2	1	1.018358
3	1	1.003414
4	1	0.988469
5	1.075	0.975660
6	1	1.040124
7	1	1.02518
8	1	1.010236
9	1	0.995291
10	1	0.982482
11	1	1.046946
12	1	1.032002
13	1	1.017058
14	1	1.002113
15	1	0.989304
16	1.79	1.064001
17	1.53	1.049056
18	1.51	1.034112
19	1.4	1.019168
20	1.34	1.006359
21	1.105	1.09811
22	1.6	1.083166
23	1.61	1.068221
24	1.6	1.053277
25	1.49	1.040468
26	1.223	1.132219
27	1.144	1.117275
28	1.102	1.10233
29	1.44	1.087386
30	1.5	1.074577
31	1.243	1.166328
32	1.174	1.151384
33	1.109	1.13644
34	1	1.121495
35	1.48	1.108686

Table C.12: Delamination at exit fitted and observed values using MLR with interaction

Number	Observed	Fitted
1	1	0.987654
2	1	0.995865
3	1	1.004076
4	1	1.012286
5	1.075	1.019324
6	1	1.000741
7	1	1.005774
8	1	1.010807
9	1	1.015839
10	1	1.020153
11	1	1.013829
12	1	1.015683
13	1	1.017538
14	1	1.019392
15	1	1.020982
16	1.079	1.046547
17	1.053	1.040456
18	1.051	1.034365
19	1.04	1.028274
20	1.034	1.023054
21	1.105	1.111983
22	1.06	1.090002
23	1.061	1.06802
24	1.06	1.046039
25	1.049	1.027197
26	1.223	1.17742
27	1.144	1.139548
28	1.102	1.101675
29	1.044	1.063803
30	1.05	1.031341
31	1.243	1.242857
32	1.174	1.189094
33	1.109	1.135331
34	1.07	1.081568
35	1.048	1.035485

Table C.13: Surface roughness fitted and observed values using MLR

Number	Observed	Fitted
1	0.453	0.9481253
2	1.194	1.103094
3	2.255	1.258062
4	3.282	1.413031
5	2.03	1.545861
6	0.601	1.013223
7	0.9	1.168191
8	0.72	1.323159
9	0.82	1.478128
10	1.06	1.610958
11	0.972	1.07832
12	1.16	1.233288
13	0.99	1.388257
14	1.2	1.543225
15	1.14	1.676055
16	1.429	1.241063
17	1.305	1.396031
18	1.82	1.551
19	1.89	1.705968
20	1.87	1.838798
21	1.62	1.566549
22	1.757	1.721517
23	2.243	1.876486
24	2.14	2.031454
25	2.273	2.164284
26	2	1.892035
27	1.783	2.047004
28	2.782	2.201972
29	2.256	2.356941
30	1.912	2.48977
31	2.492	2.217521
32	1.957	2.37249
33	3.036	2.527458
34	2.937	2.682426
35	2.198	2.815257

Table C.14: Surface roughness fitted and observed values using MLR with interaction

Number	Observed	Fitted
1	0.453	0.777112
2	1.194	1.018826
3	2.255	1.260541
4	3.282	1.502255
5	2.03	1.709439
6	0.601	0.865682
7	0.9	1.09549
8	0.72	1.325298
9	0.82	1.555106
10	1.06	1.752084
11	0.972	0.954251
12	1.16	1.172153
13	0.99	1.390055
14	1.2	1.607956
15	1.14	1.794729
16	1.429	1.175675
17	1.305	1.363811
18	1.82	1.551947
19	1.89	1.740083
20	1.87	1.901343
21	1.62	1.618523
22	1.757	1.747128
23	2.243	1.875733
24	2.14	2.004337
25	2.273	2.114569
26	2	2.061372
27	1.783	2.130445
28	2.782	2.199518
29	2.256	2.268591
30	1.912	2.327796
31	2.492	2.50422
32	1.957	2.513761
33	3.036	2.523303
34	2.937	2.532845
35	2.198	2.541023

Table C.15: Surface roughness fitted and observed values using MLR with n=34

Number	Observed	Fitted
1	0.453	0.8881406
2	1.194	1.01163
3	2.255	1.13512
4	2.03	1.364458
5	0.601	0.9620539
6	0.9	1.085544
7	0.72	1.209033
8	0.82	1.332523
9	1.06	1.438371
10	0.972	1.035967
11	1.16	1.159457
12	0.99	1.282946
13	1.2	1.406436
14	1.14	1.512284
15	1.429	1.220751
16	1.305	1.34424
17	1.82	1.46773
18	1.89	1.59122
19	1.87	1.697068
20	1.62	1.590317
21	1.757	1.713807
22	2.243	1.837297
23	2.14	1.960786
24	2.273	2.066635
25	2	1.959884
26	1.783	2.083374
27	2.782	2.206864
28	2.256	2.330353
29	1.912	2.436201
30	2.492	2.329451
31	1.957	2.452941
32	3.036	2.57643
33	2.937	2.69992
34	2.198	2.805768

Table C.16: Circularity at exit fitted and observed values using MLR

Number	Observed	Fitted
1	0.056	0.0312234
2	0.128	0.0809572
3	0.142	0.130691
4	0.402	0.1804248
5	0.484	0.2230537
6	0.072	0.0445746
7	0.154	0.0943084
8	0.158	0.1440421
9	0.17	0.193776
10	0.104	0.2364049
11	0.086	0.0579257
12	0.132	0.1076595
13	0.156	0.1573933
14	0.172	0.2071271
15	0.176	0.2497561
16	0.128	0.0913037
17	0.134	0.1410375
18	0.164	0.1907713
19	0.178	0.2405051
20	0.154	0.283134
21	0.16	0.1580596
22	0.166	0.2077933
23	0.182	0.2575271
24	0.198	0.3072609
25	0.208	0.3498899
26	0.172	0.2248154
27	0.188	0.2745492
28	0.204	0.324283
29	0.238	0.3740168
30	0.312	0.4166458
31	0.178	0.2915713
32	0.413	0.3413051
33	0.556	0.3910389
34	0.648	0.4407727
35	0.756	0.4834017

Table C.17: Diameter error at entry fitted and observed values using MLR

Number	Observed	Fitted
1	0.056	0.0578364
2	0.128	0.0981848
3	0.142	0.1385333
4	0.402	0.1788817
5	0.484	0.2134661
6	0.072	0.068784
7	0.154	0.1091324
8	0.158	0.1494808
9	0.17	0.1898293
10	0.104	0.2244136
11	0.086	0.0797315
12	0.132	0.12008
13	0.156	0.1604284
14	0.172	0.2007768
15	0.176	0.2353612
16	0.128	0.1071004
17	0.134	0.1474489
18	0.164	0.1877973
19	0.178	0.2281457
20	0.154	0.2627301
21	0.16	0.1618382
22	0.166	0.2021866
23	0.182	0.2425351
24	0.198	0.2828835
25	0.208	0.3174679
26	0.172	0.216576
27	0.188	0.2569244
28	0.204	0.2972729
29	0.238	0.3376213
30	0.312	0.3722057
31	0.178	0.2713138
32	0.413	0.3116622
33	0.556	0.3520106
34	0.648	0.3923591

Table C.18: Circularity at entry fitted and observed values using MLR

Number	Observed	Fitted
1	0.074	0.1434559
2	0.112	0.1471093
3	0.118	0.1507627
4	0.202	0.1544161
5	0.594	0.1575476
6	0.108	0.1474008
7	0.122	0.1510542
8	0.136	0.1547076
9	0.122	0.158361
10	0.094	0.1614925
11	0.168	0.1513457
12	0.156	0.1549992
13	0.152	0.1586526
14	0.136	0.162306
15	0.116	0.1654375
16	0.136	0.1612081
17	0.166	0.1648615
18	0.2	0.1685149
19	0.144	0.1721683
20	0.13	0.1752998
21	0.216	0.1809328
22	0.158	0.1845862
23	0.22	0.1882396
24	0.158	0.191893
25	0.136	0.1950245
26	0.25	0.2006575
27	0.208	0.2043109
28	0.152	0.2079643
29	0.162	0.2116177
30	0.152	0.2147492
31	0.198	0.2203822
32	0.453	0.2240356
33	0.244	0.227689
34	0.202	0.2313424
35	0.184	0.2344739

Table C.19: The fitted vs observed values of the thrust force against the delamination at  
exit

Number	Observed	Fitted
1	1	1.012821
2	1	1.01096
3	1	1.007556
4	1	1.002633
5	1.075	0.9997867
6	1	1.021856
7	1	1.013163
8	1	1.007321
9	1	1.005915
10	1	1.008024
11	1	1.022324
12	1	1.013651
13	1	1.016933
14	1	1.014729
15	1	1.008962
16	1.079	1.033811
17	1.053	1.031467
18	1.051	1.029123
19	1.04	1.027716
20	1.034	1.033811
21	1.105	1.07929
22	1.06	1.067873
23	1.061	1.061474
24	1.06	1.057723
25	1.049	1.051271
26	1.223	1.158995
27	1.144	1.130864
28	1.102	1.107421
29	1.044	1.071085
30	1.05	1.053738
31	1.243	1.252765
32	1.174	1.201848
33	1.109	1.130864
34	1.07	1.071788
35	1.048	1.054441

Table C.20: The fitted vs observed values of the cutting force against the delamination at  
exit

Number	Observed	Fitted
1	1	0.9970624
2	1	0.988915
3	1	0.983957
4	1	0.9801587
5	1.075	0.9778101
6	1	1.01588
7	1	1.009153
8	1	1.00002
9	1	0.9906837
10	1	0.9958737
11	1	1.035799
12	1	1.023984
13	1	1.021925
14	1	1.014198
15	1	1.005079
16	1.079	1.051688
17	1.053	1.050238
18	1.051	1.032523
19	1.04	1.035799
20	1.034	1.043744
21	1.05	1.115606
22	1.06	1.090686
23	1.061	1.085177
24	1.06	1.064707
25	1.049	1.061952
26	1.223	1.163665
27	1.144	1.120666
28	1.102	1.103008
29	1.044	1.068795
30	1.05	1.086278
31	1.243	1.197704
32	1.174	1.162969
33	1.109	1.130408
34	1.07	1.098862
35	1.048	1.069027

Table C.21: Table fitted values vs observed of the quadratic delamination at exit

Number	Observed	Fitted
1	1	1.008455
2	1	1.007022
3	1	1.006484
4	1	1.006243
5	1.075	1.006168
6	1	1.014373
7	1	1.011839
8	1	1.009144
9	1	1.007275
10	1	1.008203
11	1	1.024604
12	1	1.018043
13	1	1.017047
14	1	1.013696
15	1	1.010531
16	1.079	1.03569
17	1.053	1.034571
18	1.051	1.022641
19	1.04	1.024604
20	1.034	1.029823
21	1.05	1.106512
22	1.06	1.073903
23	1.061	1.067557
24	1.06	1.046708
25	1.049	1.044232
26	1.223	1.187425
27	1.144	1.113912
28	1.102	1.089229
29	1.044	1.050528
30	1.05	1.068801
31	1.243	1.259103
32	1.174	1.186084
33	1.109	1.128902
34	1.07	1.083898
35	1.048	1.050749

Table C.22: The fitted vs observed values of the delamination at exit using MLR

Number	Observed	Fitted
1	1	1.008556
2	1	1.005272
3	1	1.00146
4	1	0.9966894
5	1.075	0.993898
6	1	1.019923
7	1	1.011496
8	1	1.004815
9	1	1.001632
10	1	1.00446
11	1	1.02469
12	1	1.01515
13	1	1.01732
14	1	1.01385
15	1	1.00724
16	1.079	1.03737
17	1.053	1.03517
18	1.051	1.02939
19	1.04	1.02899
20	1.034	1.03561
21	1.05	1.08779
22	1.06	1.07318
23	1.061	1.06685
24	1.06	1.05934
25	1.049	1.05358
26	1.223	1.16206
27	1.144	1.13011
28	1.102	1.10749
29	1.044	1.0709
30	1.05	1.06091
31	1.243	1.24448
32	1.174	1.19614
33	1.109	1.13226
34	1.07	1.07811
35	1.048	1.05767

Table C.23: The fitted vs observed values of the thrust force against the delamination at entry

Number	Observed	Fitted
1	1	1.073559
2	1	1.069593
3	1	1.062338
4	1	1.051847
5	1.081	1.045782
6	1	1.092813
7	1	1.074289
8	1	1.061839
9	1	1.058841
10	1	1.063338
11	1	1.093813
12	1	1.075328
13	1	1.082322
14	1	1.077626
15	1	1.065336
16	1.252	1.118293
17	1.223	1.113297
18	1.215	1.108301
19	1.212	1.105303
20	1.157	1.118293
21	1.291	1.215213
22	1.242	1.190883
23	1.245	1.177244
24	1.226	1.169251
25	1.162	1.155502
26	1.33	1.385073
27	1.296	1.325123
28	1.275	1.275164
29	1.26	1.197727
30	1.165	1.160758
31	1.48	1.584909
32	1.42	1.476398
33	1.383	1.325123
34	1.368	1.199226
35	1.347	1.162256

Table C.24: Delamination at entry observed vs predicted values against thrust force transformation

Number	Observed	Fitted
1	1	1.073559
2	1	1.069593
3	1	1.062338
4	1	1.051847
5	1.081	1.045782
6	1	1.092813
7	1	1.074289
8	1	1.061839
9	1	1.058841
10	1	1.063338
11	1	1.093813
12	1	1.075328
13	1	1.082322
14	1	1.077626
15	1	1.065336
16	1.252	1.118293
17	1.223	1.113297
18	1.215	1.108301
19	1.212	1.105303
20	1.157	1.118293
21	1.291	1.215213
22	1.242	1.190883
23	1.245	1.177244
24	1.226	1.169251
25	1.162	1.155502
26	1.33	1.385073
27	1.296	1.325123
28	1.275	1.275164
29	1.26	1.197727
30	1.165	1.160758
31	1.48	1.584909
32	1.402	1.476398
33	1.383	1.325123
34	1.368	1.199226
35	1.347	1.162256

Table C.25: The fitted vs observed values of the cutting force against the delamination at entry

Number	Observed	Fitted
1	1	1.026012
2	1	1.006635
3	1	0.9948425
4	1	0.9858088
5	1.081	0.9802232
6	1	1.070767
7	1	1.054768
8	1	1.033046
9	1	1.010841
10	1	1.023185
11	1	1.118142
12	1	1.090041
13	1	1.085145
14	1	1.066767
15	1	1.045079
16	1.252	1.155931
17	1.223	1.152483
18	1.215	1.110349
19	1.212	1.118142
20	1.157	1.137037
21	1.291	1.307952
22	1.242	1.248682
23	1.245	1.235579
24	1.226	1.186894
25	1.162	1.180343
26	1.33	1.422252
27	1.296	1.319986
28	1.275	1.277989
29	1.26	1.196617
30	1.165	1.2382
31	1.48	1.50321
32	1.42	1.420597
33	1.383	1.343156
34	1.368	1.268128
35	1.347	1.197169