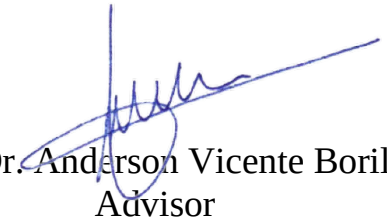
Dissertation presented to the Instituto Tecnológico de Aeronáutica, in partial fulfillment of the requirements for the Degree of Master of Science in the Program of Aeronautics and Mechanical Engineering, in the Area of Materials, Manufacturing, and Automation.

**André Dorigueto Canal**

# SURFACE ROUGHNESS ANALYSIS IN TURNING PROCESSES USING ANN

**Dissertation approved in its final version by the signatories below:**



Prof. Dr. Anderson Vicente Borille
Advisor

Profª. Drª. Emília Villani
Pro-Rector of Post-Graduation

Campo Montenegro
São José dos Campos, SP – Brasil
2022

**Cataloging-in-Publication Data (CIP)**
**Documentation and Information Division**

Canal, André Dorigueto
Surface roughness analysis in turning processes using ANN / André Dorigueto Canal.
São José dos Campos, 2022.
158 p.

Master of Science Dissertation – Postgraduate course on Aeronautics and Mechanical Engineering, Area of Materials, Manufacturing, and Automation – Instituto Tecnológico de Aeronáutica, 2022.
Advisor: Prof. Dr. Anderson Vicente Borille

1. Turning process. 2. Surface roughness. 3. Machine learning. I. Instituto Tecnológico de Aeronáutica. II. Surface roughness analysis in turning processes using ANN.

## BIBLIOGRAPHIC REFERENCE

CANAL, André Dorigueto. **Surface roughness analysis in turning processes using ANN**. 2022. 158 p. Master of Science Dissertation in Materials, Manufacturing, and Automation – Instituto Tecnológico de Aeronáutica, São José dos Campos, 2022.

## CESSION OF RIGHTS

AUTHOR'S NAME: André Dorigueto Canal
PUBLICATION TITLE: Surface roughness analysis in turning processes using ANN
PUBLICATION KIND/YEAR: Dissertation / 2022



_____

André Dorigueto Canal
Praça Marechal Eduardo Gomes, 50 – Vila das Acácias
CEP: 12228-900, São José dos Campos – SP

# SURFACE ROUGHNESS ANALYSIS IN TURNING PROCESSES USING ANN

**André Dorigueto Canal**

Thesis Committee composition:

| | | |
|---|---|---|
| Prof. Dr. Jefferson de Oliveira Gomes | Chairperson | ITA |
| Prof. Dr. Anderson Vicente Borille | Advisor | ITA |
| Prof. Dr. Filipe Alves Neto Verri | | ITA |
| Prof. Dr. Milton Luiz Polli | | UTFPR |

**ITA**

A meus pais, Mauro e Luciana,
pois nada seria sem eles.

# Agradecimentos

Em primeiro lugar, agradeço ao que acredito que olha por mim, de alguma dimensão diferente da nossa, Deus.

À minha família, fundação e exemplo, que construiu comigo meus princípios. Especialmente a meus avós — Vitorino e Honorina, Jacob e Cilma — e a meus pais — Mauro e Luciana.

À minha companheira de vida, Laíssa Carvalho, que por tantas vezes "recarregou minhas baterias" e com quem divido meus sonhos. E aos seus pais, Geraldo e Rita, por terem me acolhido como filho.

Agradeço especialmente àqueles que me acolheram, me abriram portas, que incentivaram meu crescimento e torceram por ele — em ordem cronológica — Profª. Drª. Eliene Lucas, Prof. Dr. Lucas Benini, Dr. Rafael Mundim, Prof. Dr. Ronnie Rego.

Ao meu orientador, Prof. Dr. Anderson Borille, pelos conselhos cirúrgicos e direcionamentos técnicos que me ajudaram a moldar esta dissertação.

Aos amigos que ajudaram diretamente neste trabalho, Mateus Ferreira, Wilson Lara e Carla Cunha. E àqueles que caminharam nessa jornada comigo, cada um em um momento, Givan Junior, Leonardo Miranda, Eduardo Muñoz, Rodolfo Alves, Tédni Goulart, Gabriel Vargas, Natália D'Ávilla, Izabel Criscuolo e tantos outros que formam a especial família do CCM-ITA.

Agradeço aos meus professores, pela dedicação e decisão diária de seguir a admirável vocação de ensinar.

Agradeço com contentamento os funcionários do ITA, que se empenham diariamente para manter tudo em funcionamento, empregando seu ofício de forma atenciosa e receptiva.

É fundamental também que eu agradeça às instituições que apoiaram essa pesquisa: ao ITA e ao CCM, onde me foi fornecida infraestrutura, meus sinceros agradecimentos e admiração; à CAPES que cedeu apoio financeiro; à Villares Metals — através de Roberto Tiburcio — que doou considerável quantidade de aço para experimentação; à SandvikCoromant — através de Aldeci Santos — que cedeu ferramentas de corte e suportes; às parceiras de longa data do CCM que em acordo de comodato, fornecem máquinas-ferramenta de usinagem — Indústrias ROMI — e equipamentos de medição — Mitutoyo — e à Blaser Swisslube, também parceira de muito tempo, que fornece fluido de corte.

Por fim, o presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

*"É pra frente que se anda"*

Jacob Dorigueto — meu avô, que tanto amo.

# Resumo

A rugosidade de superfície é um elemento de qualidade crucial em usinagem e sua previsão tem sido o caso de estudo por muitos anos. Geralmente, três abordagens são realizadas para modelar a rugosidade de superfície: métodos empíricos, métodos teóricos/simulação e métodos de computação (*soft computing*). Neste trabalho, modelos de redes neurais artificiais (RNA) foram treinados com dois conjuntos de dados de treinamento gerados por experimentos de usinagem — torneamento de um aço AISI H13 com fluido de corte. O primeiro experimento com condições de ferramenta teoricamente nova produziu 324 amostras e o segundo com o desgaste de flanco da ferramenta de corte variando em três níveis produziu 288 amostras. Os modelos de RNA foram formados com três camadas ocultas (28-56-14) e a rugosidade Ra como saída. Um modelo usou a profundidade de corte ($a_p$), a velocidade de avanço (f) e a velocidade de corte ($v_c$) como entradas, e um outro modelo usou esses fatores e as componentes da força de usinagem — força de corte ($F_c$), força de avanço ($F_f$), força passiva ($F_p$) e a força resultante. Foi proposta uma estratégia para multiplicar seis vezes os dados disponíveis, aumentando as medidas sem aumentar os experimentos. Os modelos treinados com um conjunto de dados menor foram propensos a sobreajustar e alcançaram menor desempenho do que aqueles com conjuntos de dados maiores. Constatou-se que o modelo treinado com as forças de usinagem teve um desempenho 75% melhor do que o modelo sem elas, produzindo um erro de previsão (MAPE) de menos de 5%. Além disso, os modelos treinados com o primeiro conjunto de dados (sem desgaste) não puderam generalizar para o segundo conjunto de dados (com desgaste), indicando que o desgaste da ferramenta é um fator essencial ao modelar a rugosidade de superfície em processos de torneamento utilizando RNAs. Além disso, os modelos de RNA tiveram melhor desempenho do que uma equação teórica clássica de rugosidade de superfície.

# Abstract

Surface roughness is one crucial quality element in machining production, and its prediction has been the case of study for many years. Generally, three approaches are performed to model it: empirical methods, theoretical/simulation methods, and soft computing methods. In this work, artificial neural network models (ANN) were trained with two training datasets generated by machining experiments—turning of an AISI H13 steel with cutting fluid. The first experiment with theoretically new-tool conditions produced 324 samples, and the second with cutting tool flank wear varying in three levels produced 288 samples. The ANN models were formed with three hidden layers (28-56-14) and Ra as the predicted output. One model used the depth of cut, feed rate, and cutting speed as inputs, and another model used those features and the machining force components—cutting force ($F_c$), feed force ($F_f$), passive force ($F_p$), and the resultant force. A strategy was proposed to augment the available data six times by increasing the measurements without increasing the experiments. The models trained with a smaller dataset were prone to overfitting and achieved less performance than those with larger datasets. The model trained with the machined forces performed 75% better than the model without them, producing a prediction error (MAPE) of less than 5%. Moreover, the models trained with the first dataset (without tool wear) could not generalize to the second dataset (with tool wear), indicating that tool wear is an essential factor when modeling surface roughness in turning processes using ANNs. Also, the ANN models performed better than a classical theoretical surface roughness equation.

# List of Figures

# List of Tables

# List of Abbreviations and Acronyms

| | |
|---|---|
| ABNT | Associação Brasileira de Normas Técnicas |
| Adagrad | Adaptative Gradient (Optimizer) |
| Adam | Adaptative Moment Estimation (Optimizer) |
| AIC | Akaike Information Criterion test |
| AICc | AIC value corrected for small sample sizes |
| AICcWt | Akaike weight (AIC test) |
| AISI | American Iron and Steel Institute |
| ANN | Artificial Neural Network |
| ANOVA | Analysis of variance |
| BRICS | Five major emerging economies: Brazil, Russia, India, China, and South Africa |
| BS | Batch size |
| BUE | Built-up-Edge |
| CAPES | Coordenação de Aperfeiçoamento de Pessoal de Nível Superior |
| CCM | Centro de Competência em Manufatura / Competence Center in Manufacturing |
| ChipA | Chip's sectional area |
| CIP | Dados Internacionais de Catalogação na Publicação / |
| Cum.Wt | The cumulative sum of the AIC weights (AIC test) |
| Delta_AICc | The difference between best and current model AIC values (AIC test) |
| Df | Degrees of freedom |
| DIN | German Institute for Standardization (Deutsches Institut für Normung) |
| DoE | Design of Experiments |
| EIEX | Emerging Industrial Economies (EIE) excluding China (X) |
| EUI | European Union Industrialized Countries |
| EXP or Exp | Experiment |
| F value | Test statistic from the F-test |
| G | Grubbs' test statistic |
| GA | Genetic Algorithm |
| HP | ANN Hyperparameter |
| HV | Hardness Vickers |
| ISO | International Organization for Standardization |
| ITA | Instituto Tecnológico de Aeronáutica / Aeronautics Institute of Technology |
| K | The number of model parameters (AIC test) |
| LL | The log-likelihood of the model (AIC test) |
| MAE | Mean absolute error |
| MAPE | Mean absolute percentage error |
| Max | Maximum |
| Mean Sq | Mean of the Sum of Squares |
| Min | Minimum |
| ML | Machine learning |
| MQL | Minimum quantity lubrication |
| MSE | Mean squared error |
| MVA | Manufacturing Value Added |
| N | Sample size |
| Nadam | Variation of Adam with Nesterov momentum (Optimizer) |
| OECD | Organization for Economic Co-operation and Development |
| P | p-value (Grubbs' test) |
| Pr(>F) | p-value from the F-test |

| | |
|---|---|
| R | Coefficient of correlation |
| $R^2$ | Coefficient of determination |
| R-squared | Coefficient of determination |
| RMSE | Root mean squared error |
| RMSPE | Root mean squared percentage error |
| RMSProp | Root Mean Square Prop (Optimizer) |
| RNA | *Redes Neurais Artificiais* |
| RSM | Response Surface Metodology |
| SGD | Stochastic Gradient Descent (Optimizer) |
| SVM | Support Vector Machines |
| StDev | Standard Deviation |
| Sum Sq | Sum of Squares |
| Tcond | Cutting tool condition |
| UNIDO | United Nations Industrial Development Organization |

# List of Symbols

| | |
|---|---|
| A | Chip's sectional area ($mm^2$) |
| $\alpha$ | Tool orthogonal clearance (°) |
| $a_p$ | Depth of cut (mm) |
| b | Width of cut (mm) |
| $\beta$ | Wedge angle (°) |
| $C_{ks}$ | Cutting pressure |
| d | Diameter (mm) |
| f | Feed rate (mm/rev) |
| F | Resultant force (N) |
| $F_c$ | Cutting force (N) |
| $F_f$ | Feed force (N) |
| $F_p$ | Passive force (N) |
| $F_x$ | Cutting force (N) |
| $F_y$ | Passive force (N) |
| $F_z$ | Feed force (N) |
| g(u) | Activation function |
| G | Slenderness degree ($a_p$/f) |
| $\gamma$ | Rake angle (°) |
| h | Thickness of cut (mm) |
| k | Specific cutting pressure/force ($N/mm^2$) |
| $K_r$ | Tool cutting edge angle (°) |
| $K_r'$ | Tool minor cutting edge angle (°) |
| KT | Crater wear depth (mm) |
| $\lambda$ | Tool cutting edge inclination (°) |
| lt | Measuring length (mm) |
| m | Exponent dependent of the material |
| $\eta$ | Learning rate |
| $\pi$ | Pi |
| Ra | Arithmetic mean deviation (µm) |
| $r_e$ | Corner radius (°) |
| Rku | Kurtosis |
| Rp | Maximum peak height (µm) |
| Rsk | Skewness |
| RSm | Mean width of profile elements (µm) |
| Rt | Total height (µm) |
| Rv | Maximum valley depth (µm) |
| Tc | Cutting time (min) |
| $\theta$ | Bias |
| VB | Width of flank wear (mm) |
| $v_c$ | Cutting speed (m/min) |
| W | Weight |
| x | Input or feature |
| y | Output |

# Contents

# 1 Introduction

The manufacturing added value has globally grown in the last decade[1] (Figure 1) (UNIDO, 2020a). With the COVID-19 crisis' impact on global economies, the urge to reshoring manufacturing has increased due to the manufacturing's strategic role in making nations' industries more resilient to ever more frequent and intensifying disasters[2] (DOMENECH; FOKEER, 2020; UNITED STATES, 2022; WORLD BANK, 2020; WORLD MANUFACTURING FOUNDATION, 2020).

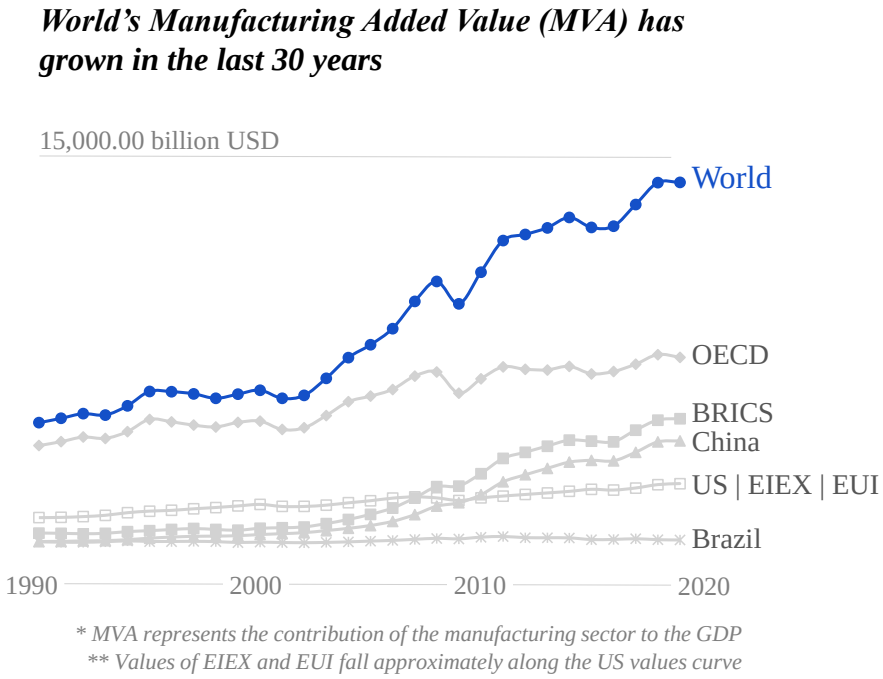



Figure 1. Manufacturing Value Added growth from 1990 to 2020[3] (UNIDO, 2020a).

In addition, the evolving understanding of the gains in using digitalization and related technologies permeates markets and industries around the world. Again, the urge for resilience catalyzed by the pandemic is upon demand—a report shows that the investment in emerging technologies in 2021 is mainly to enhance resilience (GARTNER, 2021). To a broader extent, even government policies and initiatives are being launched to leverage such technologies and

[1] Manufacturing Value Added (MVA) measures the contribution of manufacturing to the gross domestic product (GDP) (UNIDO, 2020b)—MVA is a measure of industrialization (ANDREONI; UPADHYAYA, 2014).
[2] Besides COVID-19, the number of disasters—geophysical, hydrological, meteorological, and climatological—has increased since 1980 (WORLD BANK, 2020).
[3] Available data at UNIDO website.

follow this trend (JAMWAL et al., 2021; LUND; VILDÅSEN, 2022; XU; XU; LI, 2018). Such a movement started with the revolutionary thinking nucleated in Germany in 2013 with the coinage of the term *Industry 4.0*. Following a path toward autonomous adaptability (Figure 2), Industry 4.0 significant achievement is accelerating decision-making and adaptation processes. The term calls worldwide industries to change their internal culture and has the following definition (SCHUH et al., 2017): "real-time, high data volume, multilateral communication, and interconnectedness between cyber-physical systems and people."

From data-rich systems, the goal lies in *seeing* the phenomena ("what is happening?"); *understanding* it ("why is it happening?"); *predicting* ("what will happen?"); and *self-optimizing* it ("how can an autonomous response be achieved?") (SCHUH et al., 2017). The current dissertation lies in these stages: from a data-rich machining process, understanding the correlations, and predicting a dependent variable—a step behind self-optimization.

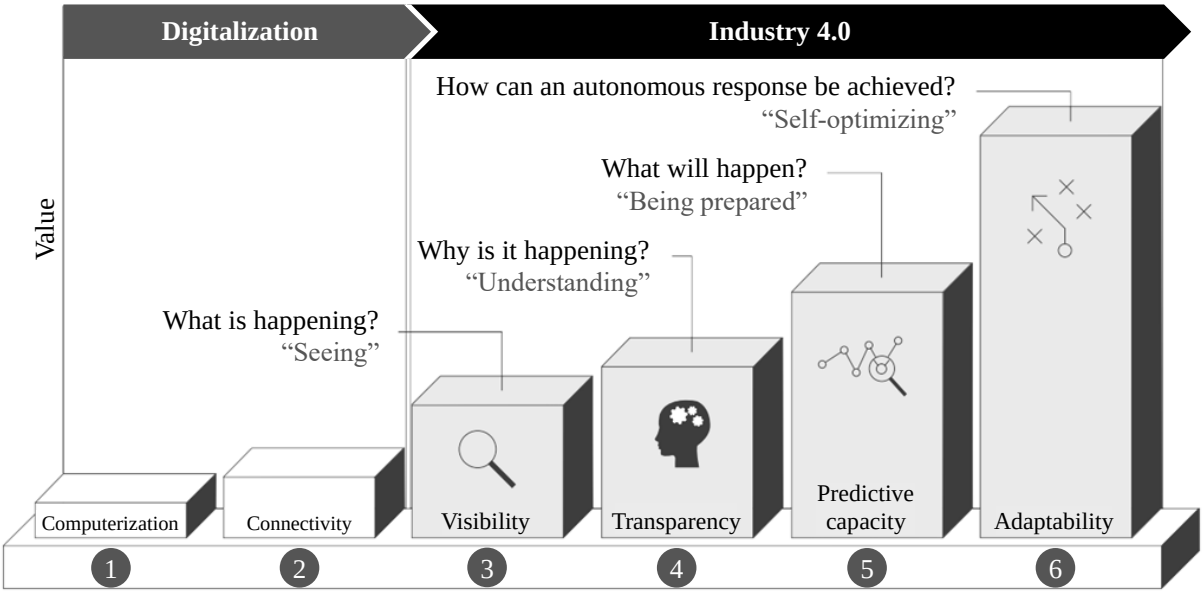



Figure 2. Stages to develop Industry 4.0 (SCHUH et al., 2017).

Many commercially sensing solutions can support process monitoring today, thus enabling factories to see *what is happening*. However, the next step demands a solid understanding. In the machining sector, the phenomena (i.e., cutting forces, surface generation, cutting tool wear, or regenerative vibration) are related to several interrelated factors in a highly complex way. Although much energy has been dedicated to such understandings, there is a long path before reaching solid general knowledge of *why it is happening*. Nevertheless,

understanding and predicting are closely related, and there has already been a rich development of modeling methods aiming at prediction in machining processes (TRENT; WRIGHT, 2000).

Besides such macro industrial background, there is the practical challenge encountered by manufacturing engineers of selecting the correct parameters, predicting the yielded results, and maximizing production with optimum parameters. Generally, this complex mission depends on experienced and acquainted-to-the-task manufacturing practitioners (BENARDOS; VOSNIAKOS, 2003).

Artificial intelligent models could be the agents to support human decisions or, even at a higher level, reduce them to zero, removing subjectivity and dependency of experienced practitioners from factories? This question is the motivation behind Figure 3. By dividing the machining process into four components[4] and using artificial intelligence models, it may be possible to make automated predictions, monitoring, and optimization, fulfilling part of the goals of the Industry 4.0 concept.

Each component of Figure 3 (i, ii, iii, and iv) can correlate and influence themselves during machining. Potential use cases are envisioned, such as using the process parameters to predict the result on the part, using the signal acquired to predict whether the tool is worn and must be exchanged, or defining the best parameters that produce the lower forces inside the desired range of surface roughness.

The three objectives, monitor, predict and optimize (Figure 3, bottom), can converge to autonomous decision-making, which is the final step described in Industry 4.0 ideals. After digitalization, when data is freely available on industrial shop floors and challenges of understanding and foreseen are overcome, the ability to autonomously monitor, predict and optimize a process will be a key to the primary goal of Industry 4.0—the production adaptability.

The current dissertation took one of the objectives addressed in Figure 3 as its objective—the mixed-use of process parameters and acquired signal to predict surface roughness in machining.

[4] The four components would be i: Process parameters that can be defined before the process begins, such as the depth of cut ($a_p$), feed rate (f), cutting speed ($v_c$), tool geometry, and machine-tool properties; ii: Signal acquired collected during the process, such as machining forces, vibration, and acoustic emission; iii: Tool condition post-process collected, such as the flank wear width and crater wear depth; iv: Part features that are previously defined requirements or actual results, such as roughness, waviness, form deviation, geometric deviations, residual stress, and hardness.

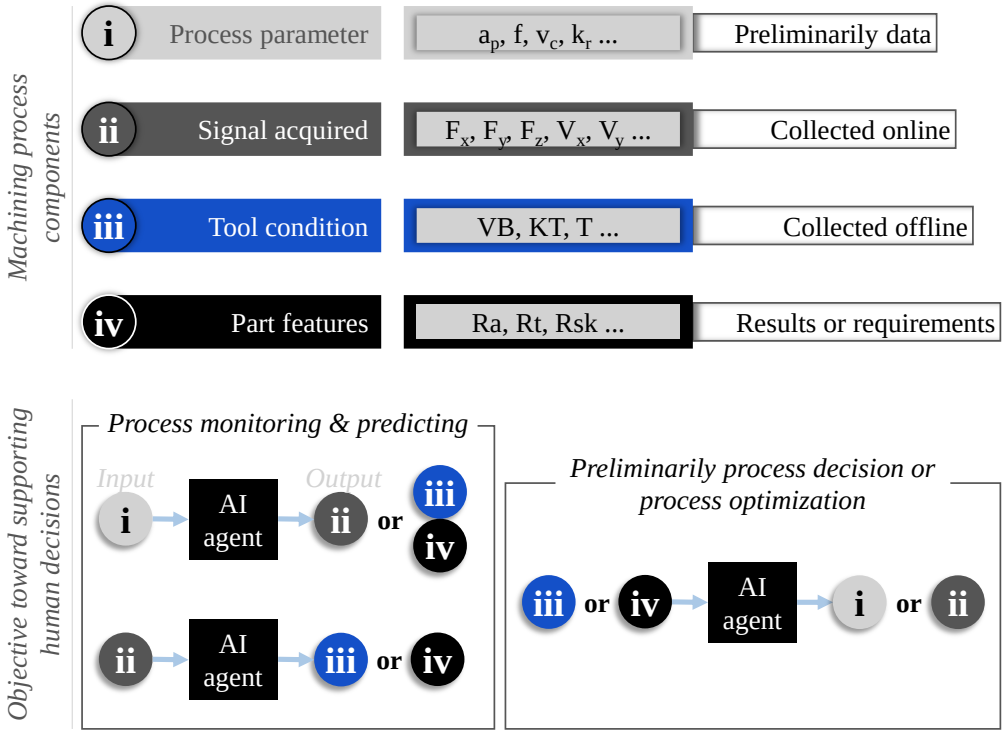



Figure 3. Machining processes division into four components (top). Soft computing technique applications (bottom).

Surface roughness is one classical machining quality requirement and is crucial for some applications, especially when functional characteristics of the final product have to be met (DAVIM, 2010). Several applications depend on the surface roughness, such as bearings fatigue resistance (BHADESHIA, 2012), heat transfer and aerodynamic performance of turbine airfoils (ABUAF; BUNKER; LEE, 1998), contact wear of many mechanical devices (BAYER; SIRICO, 1975), superhydrophobic surfaces sliding angle (MIWA et al., 2000; YANG; TARTAGLINO; PERSSON, 2006), or freeform optics machining (FANG et al., 2013). Moreover, the roughness parameters are related to contact stiffness, fatigue strength, thermal and electrical conductivity, reflexivity, friction and wear, lubrication, mechanical sealing, fatigue corrosion, and assembly tolerances (DAVIM, 2010).

To Trent and Wright (2000), the prediction of surface finish is one of the abilities machining practitioners would like to have in daily operations. Sometimes, though, designing and predicting surface roughness values are not as assertive as expected, and its prediction has been the case of study for many research projects during the last years (BENARDOS; VOSNIAKOS, 2003).

Therefore, such surface roughness prediction ability is one identified motivation, and exploring this theme using machine learning models can likely be applied in several other manufacturing fields.

## 1.1 Driven questions

Through this work's development, some questions drove this researcher. Due to the broad scope of such discussions, they cannot be assigned to specific objectives. However, they will be answered in some way at the end.

To select models' features, understanding the factors' effects on the response is desirable or even obligatory. Therefore, it would be helpful to know how different process parameters influence the machining process. **Q1: How do the varied machining process parameters affect the machined surface finish and cutting forces?**

Cutting tool wear is always present in turning processes, and the wear changes the tool's edge geometry, which can also affect the surface roughness and forces. **Q2: How does cutting tool condition (wear) affects the machined surface finish and cutting forces? Q3: How would models developed in an artificial wear-free condition perform in data generated by wear-existing machining? Q4: Vice versa: How would models developed in a wear-existing condition perform in data generated by artificial wear-free machining?**

Developing regression models usually demands much experimentation, and (especially from a machine learning engineering point of view) a data-hungry characteristic can make the task unfeasible or very expensive. **Q5: Is there a way of augmenting the available data without increasing the experiments? Q6: Does a machine learning model trained with fewer data perform sufficiently well as a model trained with more data?**

Machine learning modeling is not just the modeling part. Preparation is a must. **Q7: Does data preparation consume much effort from the practitioner as proposed in the literature? Q8: Besides data preparation and programming, is there any other human-energy consuming activity that other engineers willing to venture into the machine learning modeling field would appreciate knowing in advance?**

If it is possible to collect machining forces, they can be used as inputs for regression models. If this inclusion improves models' prediction performance, it would indicate that machining forces collecting/monitoring is valuable for engineers designing machining monitoring systems. **Q9: How would the ML model perform when using the cutting**

**parameters as inputs? Q10: How would the ML model perform using the cutting parameters and machining forces as inputs?**

Theoretical models are helpful tools for manufacturing practitioners but are not always assertive. **Q11: How much better a machine learning model can predict surface finish than a theoretical model?**

## 1.2 Objectives

The main goal of this dissertation is to contribute to the understanding of using machine learning to model metal cutting results. More precisely, the goal is to use an artificial neural network model to predict surface finish results in the turning process. Toward this broad objective, the following specific objective stood out:

i. Statistically assess the effect of the factors—depth of cut, feed rate, and tool condition (in terms of flank wear width)—on the responses—surface finish roughness parameters (Ra, Rsk, Rku, RSm, and Rt), and the machining forces;
ii. Develop an ANN model to predict the surface roughness parameter—Ra—and evaluate it with a chosen error metric;
iii. Propose a strategy to augment available data without increasing experiments;
iv. Investigate the influence of the dataset size on the ANN models' prediction performance;
v. Investigate the influence of the tool condition—flank wear width—on the ANN models' prediction performance;
vi. Compare the prediction performance of the ANN models with one theoretical equation.

## 1.3 Document structure

Previously, the first topic presented an introduction and the objectives. Subsequent topics form this work as follows: in the second topic, state of the art is presented for basic knowledge referencing; in the third topic, recently published works are evaluated in a bibliometric study; the fourth topic depicts the primary methodology and materials used; the results are presented in the fifth topic; the sixth topic concludes this work; the seventh topic

enumerates possible actions to future similar work. Subsequently, the bibliography and additional information on four appendices are presented.

Moreover, Figure 4 shows this work's big picture. From left to right, it starts with the contextual aspects connecting this work with the urge for industry resilience, Industry 4.0, and manufacturing, as was discussed in the introduction. The specific issue of this dissertation is the surface roughness modeling, which was historically managed with theoretical equations, empirical methods, computational simulations, and, more recently, soft computing techniques, amongst which the Artificial Neural Network (ANN) is the one used in this work. The procedure to collect data from an actual machining experiment and to perform on this data the necessary manipulation and machine learning pipeline activities, such as feature preparation, building ANN models, and evaluation—using data science tools, such as the Python programming language. The picture ends on the right side, showing the explored problematics: 1. The insufficiency of available data—data generation, dependent on costly machining experiments is an expensive phase that can be optimized; 2. The evaluation of different features, such as machining forces, as inputs to the surface roughness models; 3. The effect of inevitable tool wear—a time domain feature—on the models; and 4. A comparison of the ANN models built in this work with the classical theoretical equation for surface roughness.

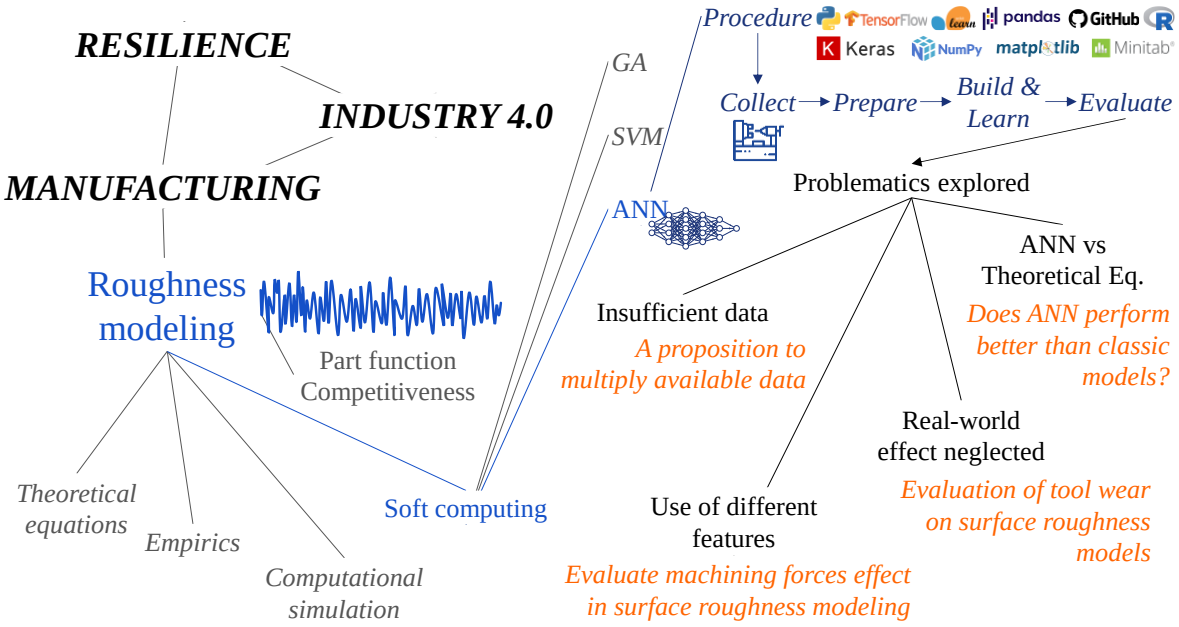



Figure 4. Summarized picture of the dissertation.

This work's context explains the use of such a complex tool as artificial neural networks to solve the problem of surface roughness modeling. The work was done with a broader objective of initiating a research line inside a research group. Furthermore, though some may see the tool as too complex for the task, the research group's objective is to start using this machine learning tool to grow knowledge foundations and use it in applications such as cyber-physical systems and autonomous factories that intrinsically demand these more complex tools.

In advance, this dissertation's products are the discussions over the importance of the tool wear in modeling the surface roughness—it was perceived that considering the tool wear at the experiments affected the produced data for models. In addition, the strategy to multiply data—by measuring more times for each condition—showed good leverage even with the necessity of further validation, the comparison between ANN models and a theoretical equation suggested the supremacy of the first, and the use of machining forces as inputs improved both the understanding of surface roughness analysis and models.

# 2 State of the Art

This chapter presents fundamental concepts naturally necessary to this work's implementation. Given that fundamental concepts can be somewhat basic, the audience already acquainted with the classical manufacturing and machining state of the art can suppress reading such related topics and proceed to the machine learning concepts presented in topic 2.9. If the reader is familiar with essential machine learning and artificial neural networks theory, such related topics can also be suppressed, proceeding to the focused review of topic 3.

## 2.1 Manufacturing and the industrial revolutions

*Manufacture* in the Cambridge Dictionary is defined as: "to produce goods in large numbers, usually in a factory using machines." That is a modern definition, however. The online etymology dictionary defines *Manufacture* based on the join of, from Latin, "*manu,"* the ablative of *manus*, which means "hand," and "*factura,"* which means "to perform." Hence, something made by hand that, in circa the 1560s (the understood period in which this term was first used), is understandable that it was not speaking of machines. Even though the word machine already existed (it is understood that the term was used in the 1540s), it meant a "structure of any kind," not today's machines. The evolution of *Manufacture* to its modern definition followed the evolution of *Machine*, of which the current definition, by the Cambridge Dictionary, is: "a piece of equipment with several moving parts that uses *power* to do a particular type of work." All in all, the evolution of both terms was able mainly when humanity developed *power* from steam, which marked the earth's history with the First Industrial Revolution (1IR).

Three revolutions marked the industry's history, changing peoples' lives and entire societies in their period and further. From the first IR to the current days, the technological inventions can be summarized as follows (MOLL, 2021):

- 1IR: 1760–1850. The invention of steam-powered machines and locomotives promoted faster production systems, commerce, and trade.
- 2IR: 1865–1914. Development of steel, electricity, petrol, industrial chemistry, and production systems (Taylor's and Ford's assembly lines).

- In between period: 1915–1960. Production of rocket technology and nuclear fission. Improvements and developments in several fields: aeronautics, motor, radio technology, plastics, and electronics (invention of the transistor).
- 3IR: 1960–current. Implementation of the Internet, World Wide Web (WWW), Ethernet, robotics, digitization, and computerization.

Some technologies of the 3IR are on focus by what, since 2013, has been addressed as Industry 4.0. Apart from discussing whether, from a historical perspective, Industry 4.0 will be considered the fourth industrial revolution, it refers to a change in how production systems work, aiming for improvement and optimization. The certainty is that the buzzword is already impacting regions from many corners of the world (JAMWAL et al., 2021; LUND; VILDÅSEN, 2022), being the focus of governmental policies (JAMWAL et al., 2021; XU; XU; LI, 2018), or changing the demanded skills from workforces (SONY; MEKOTH, 2022).

## 2.2 Metal cutting

*Cutting* is considered to be the first manufacturing technique invented. The *Oldowan* stone tool, confectioned by the first *Homo* species, was a broken stone used to cut meat and bones: “The history of manufacturing starts with the history of humankind. Somewhere in Africa 2.6 million years ago, the earliest representative of the genus Homo carefully chose a pair of suitable rocks, placed one on the ground, and then repeatedly struck it with the other rock. Through this, he created, piece by piece, a cutting edge.” (PLUMMER, 2004; ROSER, 2016).

As one might suggest, the cutting action can be used in a broad meaning. However, this work focuses on the cut of metal, which happens by removing a thin layer of metal—the chip—from a larger body—the workpiece—by a wedge-shaped tool. In engineering, the term cutting, or machining, is referred to the chip-forming processes and is generally based on metal processing. Even though it can shape other materials, such as wood and plastic, metal processing knowledge cannot be entirely transferred to other materials’ processing (TRENT; WRIGHT, 2000). Therefore, the term cutting is equivalent to machining and metal cutting in this text.

*Metal cutting* (or *machining*) is enclosed in a general manufacturing hierarchy chain as a secondary process, as depicted in Figure 5. The fabrication of a product can be defined in

three steps: primary shaping processes, those that perform the first transformation of the raw material; secondary processes, including material removal processes and other treatments, like heat and surface treatments; and finally, assembly processes to join, assembly and test the product (SWIFT; BOOKER, 2013). The primary shaping processes, such as casting and forming, shape the workpieces conserving their masses. Those processes are material efficient on the one hand, but they are not sufficient regarding surface quality and geometric tolerances, hence the need for removing material processes to a finish stage—the metal cutting processes (BEDDOES; BIBBY, 1999).

*Machining* refers to different processes, but they commonly remove material from a workpiece, each by its manner. In a macro division, some processes use tools with defined geometry (e.g., milling, drilling, turning), and others use tools with non-defined geometry (e.g., grinding, honing, lapping) (DAVIM, 2015). Other material-removal-based processes have more specific applications, like electrical discharge and ultrasonic machining. However, as this work uses the turning process, this review focuses on this technology.

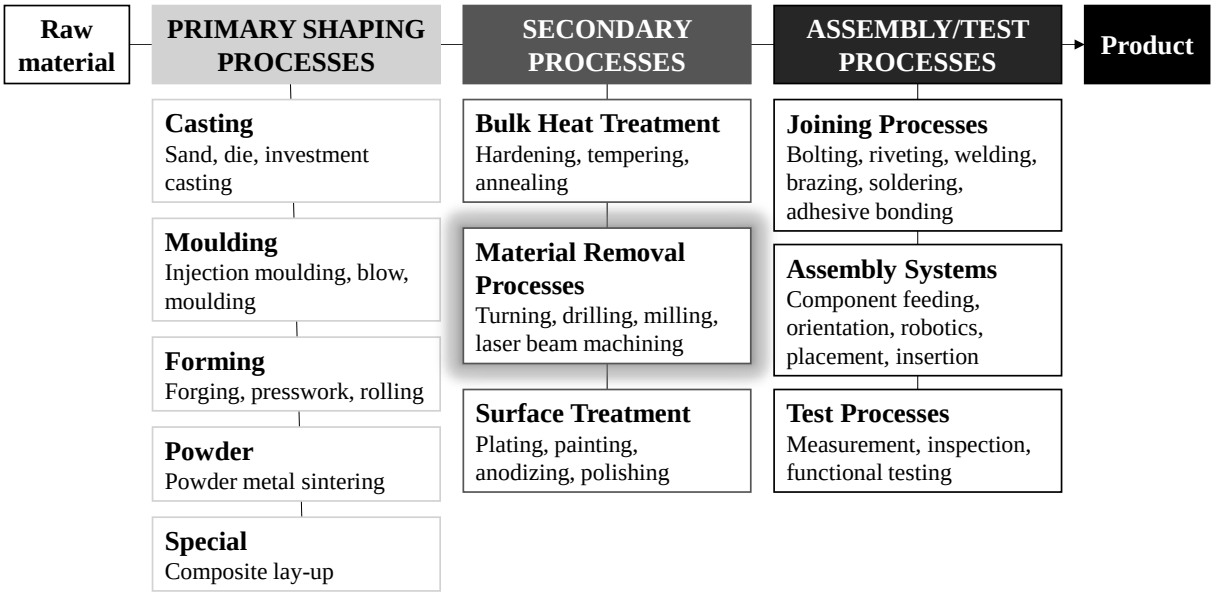



Figure 5. Hierarchy of manufacturing processes (SWIFT; BOOKER, 2013).

## 2.3 Turning process

The turning process uses geometrically defined single-point cutting tools, which means there is theoretically only one cutting point. Moreover, this technology is characterized by a continuous cutting process because the cutting region of the tool is continuously in contact with

the workpiece during machining (MACHADO et al., 2009). The turning process is primarily a rotational process where the workpiece rotation combined with a translatory feed motion results in the material removal (KLOCKE, 2011).

Figure 6 shows a typical single-point cutting tool. It is composed of a major cutting edge (or primary cutting edge) to form the chip that flows through the rake face. The transient surface (depicted on the right side of Figure 7) made by the major cutting edge flows through the flank face of the tool. The flank and rake faces are physically divided by the primary cutting edge (BEDDOES; BIBBY, 1999).

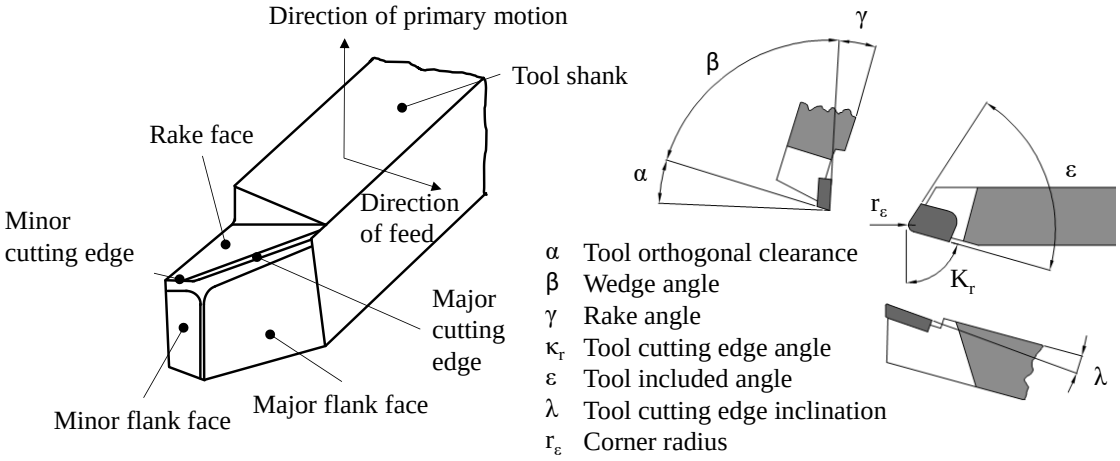



Figure 6. Cutting tool nomenclature and main angles (KLOCKE, 2011; TSCHÄTSCH, 2009).

The tool geometry can influence the process by affecting forces, chip formation, surface roughness, tool durability, and several other aspects. Each angle/inclination has distinct influential characteristics and must be used as a machining process design parameter. The *clearance angle α*—which reduces or avoids friction between the transient surface and the flank face—can significantly improve tool lifetime (STEMMER, 1993). Another important-to-choose angle is the *rake angle γ*, one of the most influential geometric features because it can significantly affect the forces, the demanded power, the surface finish, and the generated heat (STEMMER, 1993). The *tool's cutting edge inclination λ* can direct the chips and control vibration (STEMMER, 1993). The *cutting edge angle* $\kappa_r$ can reduce chatter—a type of prejudicial self-excited vibration—besides tool wear and forces (KLOCKE, 2011). The *corner radius* $r_\varepsilon$ (the radius of the curve uniting the minor cutting edge to the major cutting edge) can increase the demanded power up to 20% (STEMMER, 1993). The $r_\varepsilon$ is also significant to the surface finish, in a way that it is a factor in most of the theoretical surface roughness models (STEMMER, 1993). Moreover, the geometric features must be chosen depending on the

workpiece and tool materials, which are characteristics of each process, and, therefore, must be objects of attention (STEMMER, 1993).

Figure 7 shows, on the left side, a typical lathe machine that performs turning operations, its main components, and axes, and on the right side, a cylindrical turning process, the most generic turning process. The workpiece is clamped on the *chuck*, connected to the motor spindle that rotates the part. When the part is too heavy or too long, the *center-tailstock* can also fix it on both sides. The feed motion is supplied by translating the carriage along feed rods and lead screws (BEDDOES; BIBBY, 1999). The axes of machine tools usually follow the convention: the Z-axis is aligned with the spindle direction, the X-axis is parallel to the longest direction of table movement, and the Y-axis is parallel to the shorter direction of table movement. A, B, and C refer to angular movements around the X, Y, and Z axes, respectively (ALTINTAS, 2012). The schematic fixture and axes system can be translated entirely to the CNC (Computer Numerically Controlled) turning centers.

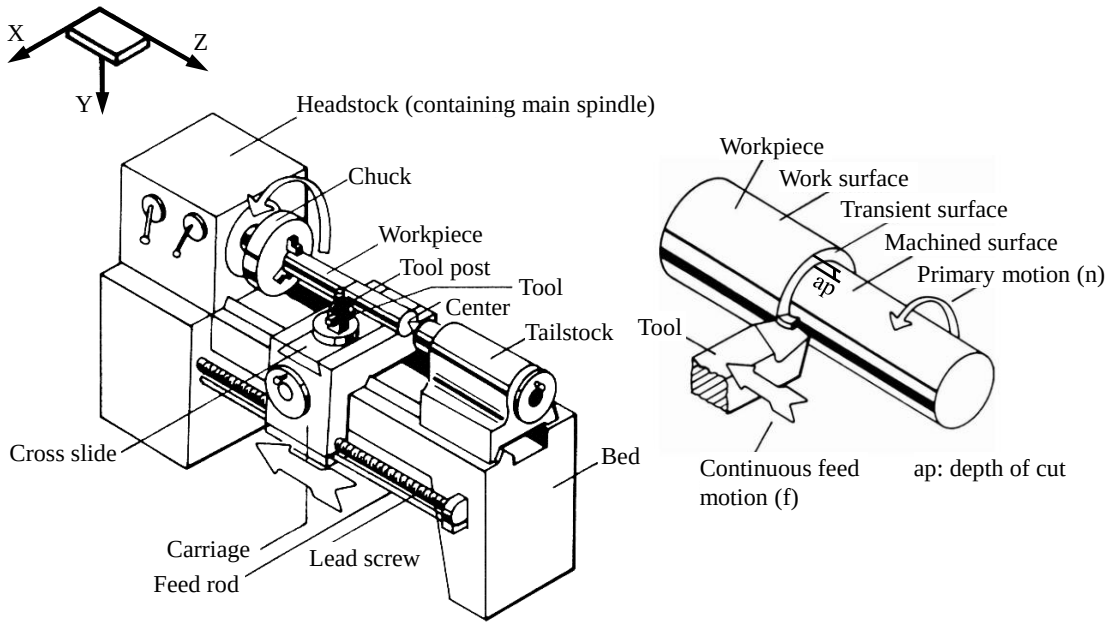



Figure 7. Lathe (left side) and cylindrical turning process (right side) (BOOTHROYD; KNIGHT, 2006).

Some helpful mathematical relations can be used in practice cylindrical turning (Table 1). In this work, the cutting time equation will be used to calculate a given process's time and relate it with tool wear measures. The cutting speed equation was used during the execution of the experiments as hand calculations next to the machine. In addition, the effect in forces of the chips' sectional area—dependent on the width of cut and the uncut chip thickness—will be investigated.

Table 1. Some mathematical relations of cylindrical turning.

| *Item* | *Mathematical relation* | *Unit* | *Eq.* | *Source* |
|---|---|---|---|---|
| Cutting time | $Tc = \frac{L}{f \cdot n}$ | min | 1 | (BOOTHROYD; KNIGHT, 2006) |
| Cutting speed | $v_c = \pi \cdot d \cdot n$ | m/min | 2 | (TSCHÄTSCH, 2009) |
| Width of cut | $b = \frac{a_p}{\sin(\kappa_r)}$ | mm | 3 | (STEMMER, 1993) |
| Uncut chip thickness | $h = f \cdot \sin(\kappa_r)$ | mm | 4 | (STEMMER, 1993) |
| Chip's sectional area | $A = a_p \cdot f = b \cdot h$ | mm² | 5 | (TSCHÄTSCH, 2009) |

## 2.4 Machining forces

Forces during the metal cutting processes are an essential factor. Its magnitudes can be used in the design of machine tools, cutting tools, tool holders, and other cutting systems; also to previously determine the cutting conditions; to understand wear mechanisms after cutting has been performed, in scientific analysis of metal cutting; to monitor the process by identifying tool wear or breakage, and so on (KLOCKE, 2011; TRENT; WRIGHT, 2000; TSCHÄTSCH, 2009). Machining forces can also be used to classify materials regarding their machinability (despite its ambiguous meanings, it means the ease one can cut a specific material in specific conditions; usually, the harder to cut, the larger forces happen) (KLOCKE, 2011).

Figure 8 shows the resultant force and its components for a typical turning process. The forces acting on the tool have the same amplitude as those acting on the chip, in opposite directions (ALTINTAS, 2012). The *resultant force F* can be decomposed into three components that act on the tool (KLOCKE, 2011; TRENT; WRIGHT, 2000):

- *The cutting force $F_c$*: acts contrary to the cutting direction and normal to the cutting edge. It is usually the more significant load in turning processes.
- *The feed force $F_f$*: acts contrary to the feed direction (or normal to the cutting direction). It is also referred to as thrust force Ft.

- *The passive force $F_p$*: acts alongside the tool holder and tends to push the tool away from the workpiece. Sometimes, it is ignored in turning operations because it tends to be the smallest of the three components.

The force components can also be named according to the machine tool axes' names: $F_c$ is equivalent to $F_x$ because it acts along the X-axis; $F_f$ is equivalent to $F_z$ because it acts along the Z-axis; $F_p$ is equivalent to $F_y$ because it acts along the Y-axis. Alternatively, even according to the direction related to the workpiece: $F_c$, $F_f$, and $F_p$ would be named *tangential force*, *axial force*, and *radial force*, respectively (SMITH, 2008).

The cutting force is the most used for evaluating cutting processes, especially when the other two components are negligibly low. However, the specific cutting force, the resultant force, and the specific resultant force may also be used (KLOCKE, 2011). As is typically the greater between the three components, the cutting force is used in the calculus of the machining demanded power. This calculus can be defined as simply the product between cutting force and cutting speed (SMITH, 2008).

The measurement of machining forces is made using *dynamometers*. Many have been developed for this purpose in the last century until the most common dynamometers were made as force platforms using piezoelectric load cells that can measure the force in many directions and calculate the resultants (TRENT; WRIGHT, 2000). A piezoelectric dynamometer usually has high static stiffness and rigidity, high natural frequency, high sensitivity, accuracy and reliability, and low-temperature sensitivity, generating precise results (TSCHÄTSCH, 2009; YOUSSEF; EL-HOFY, 2008). The force measuring Unit is the Newton (N).

The resultant force is the vectorial sum of the three components and can be calculated with Eq. 6, which is essentially a sum of orthogonal vectors using the Pitágoras Theorem.

$$F = \sqrt{F_c^2 + F_f^2 + F_p^2} \qquad \text{Eq. 6}$$

Comparative to other metal processing operations, such as forging (a primary shaping process), the forces in machining tend to be low because of the slight thickness of the material removed. However, for the same reason (small contact areas), the specific pressure (stress) in the cutting region is significant, one of the largest amongst manufacturing processes (TRENT; WRIGHT, 2000).

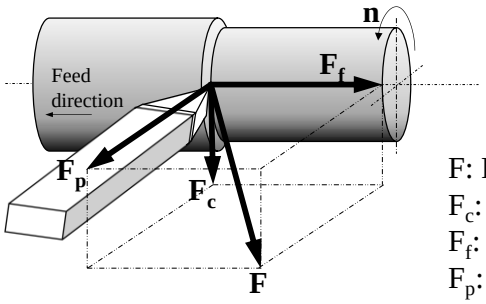



Figure 8. Resultant force and its components for a general turning process (TSCHÄTSCH, 2009).

The force can be empirically linked to the multiplication of specific cutting pressure by the cutting area (STEMMER, 1993). Kienzle proposed a practical machining force model. The Kienzle-Victor model, defined by Eq. 7, uses this relation of the area and pressure with an additional constant established in empirical experimentations and dependent on the workpiece material: *1-mc*. The *specific cutting pressure* $k_{c1.1}$ is the pressure amount that is necessary to remove a chip section of area $A = b \cdot h = 1x1\ mm^2$.

$$F_c = k_{c1.1} \cdot b \cdot h^{1-mc} \qquad \text{Eq. 7}$$

The Kienzle-Victor model depends on the cutting area and the machined material. However, the forces in machining are affected by other factors. Towards this consideration, another model was proposed with more corrective constants in an attempt to represent the numerous influencing factors in the force response. This is the *Kronenberg* model (Eq. 8). In this equation, $C_{ks}$ is the cutting pressure for the removal of a chip section of area $A = b \cdot h = 1x1\ mm^2$ and a slenderness degree $G = a_p/f = 5$, which is obtained empirically. The letters $g$ and $i$ represent p-q/2 and (p+q)/2, respectively. The letters $p$ and $q$ are indicators of the influence of the feed rate (f) and the depth of cut ($a_p$), respectively, and are given in machining standards (STEMMER, 1993).

$$F_c = C_{ks} \cdot \left(\frac{G}{5}\right)^g \cdot A^{1-i} \qquad \text{Eq. 8}$$

For both models and further ones proposed in the literature, there is always the trade-off between the ease of calculation with less demand for empirical data and the preciseness of the predictions with considered more influencing factors.

To extend the prediction of the forces to the milling process, one can adapt the *Kienzle-Victor* equation or use other models proposed in the literature, as discussed by Mundim (2014).

## 2.5 Cutting tool wear

External thermal and mechanical loads actuate on the cutting region (workpiece and cutting tool) during cutting (BRINKSMEIER et al., 2014). Those quantities lead to deformation, separation, and friction, forming a particularly complex load in the cutting edge area of the tool. The constant high compressive and thermal stresses provoke the tool to wear, reaching the end of its useful life: the tool becomes inutile when it does not cut efficiently or even fails to cut at all (KLOCKE, 2011; TRENT; WRIGHT, 2000).

Tool wear is defined as the change of shape in the tool by the loss of material or deformation (ISO, 1993). It is a cause of disturbance in machining processes that can lead to defects and poor final parts quality. Therefore, the wear behavior, its causes, and possible troubleshooting must be considered when designing and performing metal cutting processes (KLOCKE, 2011). Despite other complex causes, tool wear happens mainly by three *mechanisms*: abrasion, adhesion, and diffusion (BOOTHROYD; KNIGHT, 2006; KLOCKE, 2011):

- Abrasion: Hard particles on the chip and workpiece remove tool material by *mechanical* movement. The actuating particles can be strain-hardened remainings of built-up edges, fragments previously removed from the tool, or hard components of the workpiece material.
- Adhesion: *Friction* movement causes micro-weldings between the working surfaces of the tool and workpiece. Later, the fraction of the tool material is carried away by the movement between the two surfaces. The adhesion mechanism is caused by mechanical action—high temperatures provoke high deformable material that tends to bond with the friction movement; or chemical interaction—like diffusion process, electric exchange, or electrical polarization. It depends on the tendency of adhesion between materials.
- Diffusion: Triggered by *high temperatures* in the cutting region, individual atoms from the tool surface migrate to the workpiece surface or vice versa. Although this mechanism affects a narrow region of the interface of contact

between tool and workpiece, it is a critical wear mechanism due to its potential to decrease tool resistance to abrasion. This resistance reduction happens when the diffused atoms are important alloying elements constituents of the tool.

Figure 9 shows how wear causes are dependent on the cutting temperature, which, by its hand, is strongly related to the cutting speed. It is perceived that, initially, with lower temperatures, the predominant causes of tool wear are adhesion and abrasion. Increasing temperature in the cutting region activates other mechanisms, such as diffusion and oxidation. In really high temperatures, diffusion becomes a protagonist cause of tool wear.

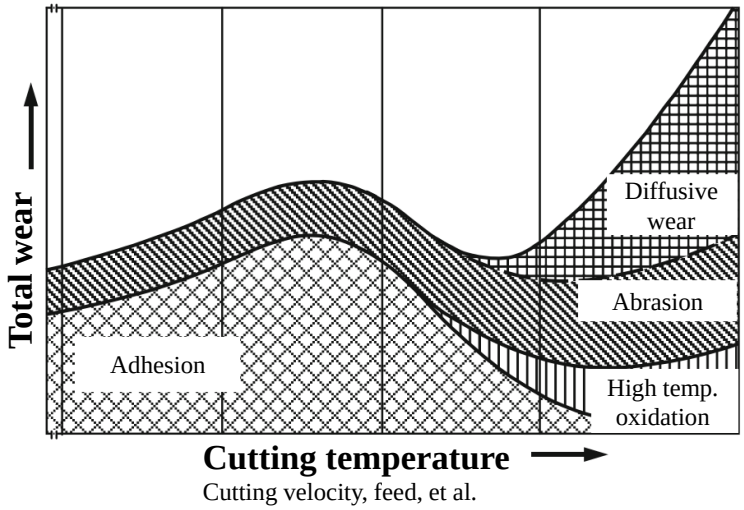



Figure 9. Wear causes dependent on cutting temperature (KLOCKE, 2011).

A practical way of determining the tool's end-life is by detecting distinctive anomalies during cutting, such as fume presence, excessive noise, abnormally severe vibration, changes in the part's surface finish or geometry, or when the cutting tool dimensions change in specified ways. Moreover, identifying the need to exchange or regrind the cutting tool usually depends on experienced operators (TRENT; WRIGHT, 2000).

Wear changes the dimensions of the cutting tool in different aspects. In Figure 10, Klocke (2011) reunited the several forms of wear observed for a typical turning operation: flaking, fragmentation, plastic deformation, smearing, notching, cratering, and flank wear. Emphasis should be addressed to the crater and flank wears because they behave more progressively than others, making it possible to follow their progress during tool-life. In addition, they are the easiest to measure amongst wear forms.

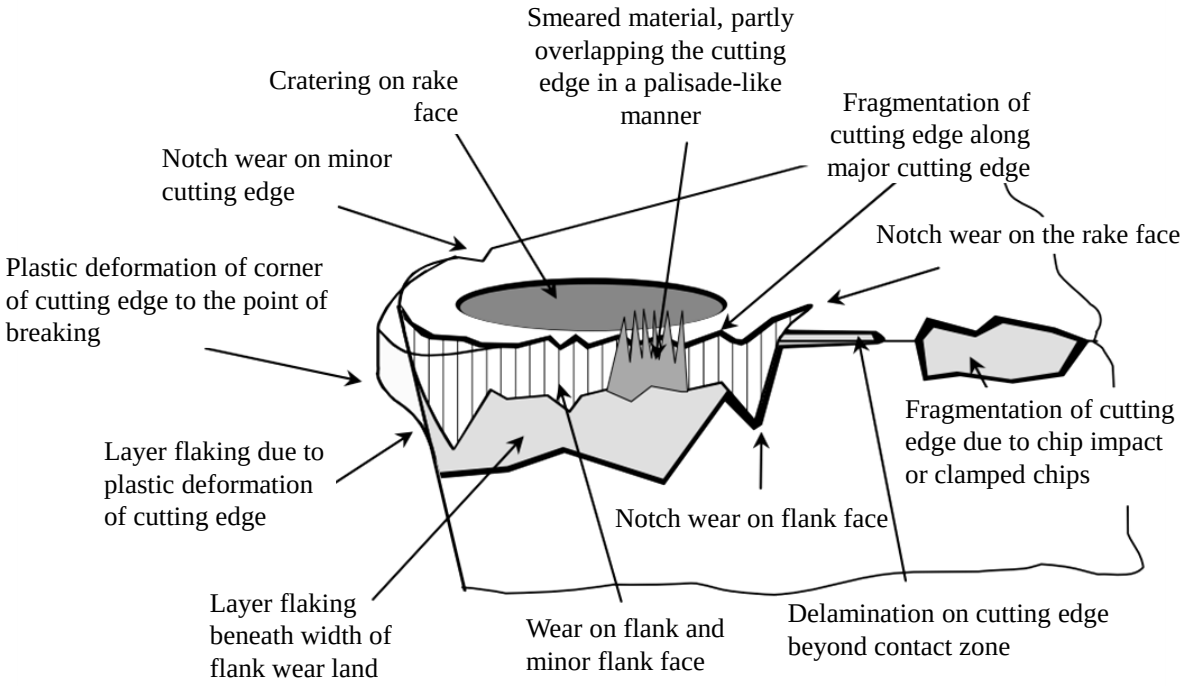



Figure 10. Forms of wear (KLOCKE, 2011).

Although it is possible to see whether a tool is broken or ceased to cut appropriately by observing the anomalies commented above, one can identify the need to exchange the tool before it reaches the end of its useful life, avoiding the adverse effects of a completely worn tool. *Wear measures* are adopted to define the amount of change in the tool's dimensions and can be used as a tool-life criterion (ISO, 1993).

The tool-life criteria based on wear measurands are usually different for different tool types (high-speed, sintered carbide, and ceramics). Commonly, for sintered carbide cutting tools (which is the case for this work), the criteria used are (ISO, 1993): the average width of the flank wear land $VB_B$ = 0.3 mm (Figure 11) when the region is considered to be regularly worn. If it is not regularly worn, one can use the maximum width of the flank wear land $VB_B$ max. = 0.6 mm. Moreover, it also can be used the depth of crater wear KT = 0.2 mm and the breakage of the crater.

The wear measures are related to tool wear *forms*, which visibly change during cutting. Figure 11 shows the *wear measures* defined by the International Standard ISO 3685(E) (1993). From these measures stand out the *width of the flank wear* $VB_B$, the *crater depth* KT, and the *displacements of cutting edge* ($SV_\alpha$ and $SV_\gamma$).

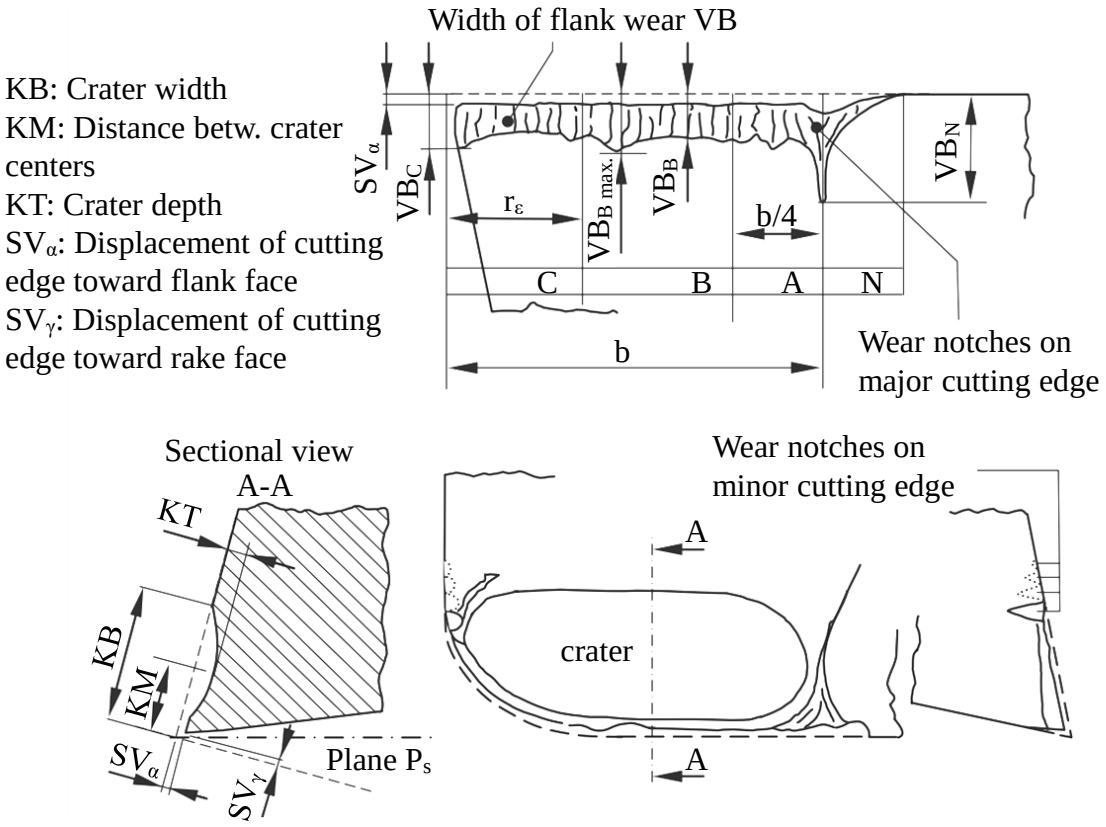



Figure 11. Tool wear measures (ISO, 1993; KLOCKE, 2011).

Notch, defined as $VB_N$ in Figure 11, should not be taken as abrasive wear as in regions A, B, and C. The notch wear usually happens by chemical action and the influence of work-hardening of the workpiece surface generated by the previous cut. It happens outside the interface tool-workpiece and affects both the flank and rake face (ISO, 1993).

Figure 12 shows the two wear forms of crater and flank wear. The interface chip-tool happens on the tool's rake face, which is the region where the type of wear occurs in the form of a crater. Crater wear is the protagonist in tool wear control in high-speed cutting, and flank wear is mainly used under economic conditions (BOOTHROYD; KNIGHT, 2006). Therefore, in most cases, the flank wear is the measure to be assessed when controlling tool condition—which will also be the case in this dissertation. The flank wear happens in the interface tool-workpiece at the tool's flank face. In Figure 12, the dashed lines represent the original state of the tool's rake and flank face before wear.

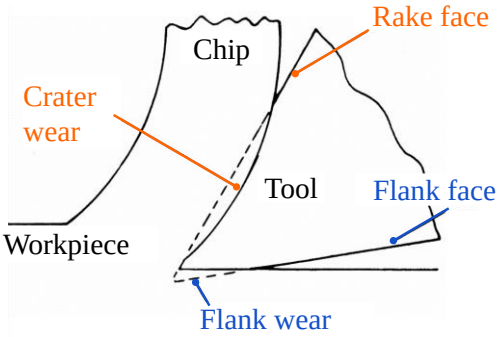



Figure 12. Most used forms to control tool wear (BOOTHROYD; KNIGHT, 2006).

Besides wear measurands, such as the width of flank wear and the depth of crater wear, other tool-life criteria are commonly used, such as a pre-defined limit on surface roughness, geometry deviation, cutting forces, and spindle power, among others. Summarized tool-life criteria are as follows (KLOCKE, 2011):

- Tool wear measurands (e.g., flank wear width limit),
- Workpiece results' measurands (e.g., surface roughness limit),
- Process measurands (e.g., spindle electric power limit).

## 2.6 The machined surface topography

"A surface can be defined as a border between a machined workpiece and its environment" (DAVIM, 2010).

The machined surface is embedded with characteristics that are related to a part's function, performance, or both, in a manner that the surface can imply, for example, in more or less corrosion resistance, improved fatigue strength, or tribological behavior qualities, like improved lubrication and wear resistance or less friction energy loss, among many other. Those surface characteristics are generally referred to as surface integrity aspects and are divided between external—topography—and internal aspects—microstructure, mechanical properties, and residual stress (DAVIM, 2010).

This work mainly treats the external topography aspect of surface integrity: surface roughness. Using the definitions proposed by Davim (2010), a *nominal surface* is given by the original project design without considering any deviation, and a *real profile* is a resultant surface read by measuring assessment equipment (microscopes, profilometers, etc.) that

considers the geometric deviations from the nominal surface, which are listed in Table 2. The first and second orders are named as *form* and *waviness* errors, respectively, and the third, fourth, fifth, and sixth orders sum the *roughness*. Figure 13 depicts the *form*, *waviness*, and *roughness* superposed. The waviness profile (W-Profile) is the primary profile without the roughness—which is removed by a low-pass filter (ISO, 1997). The roughness profile (R-Profile) is the primary profile without the waviness removed by a high-pass filter (ISO, 1997; KLOCKE, 2011).

Table 2. Geometric deviations of machined surfaces according to DIN 4760 (1982).

| *Order* | *Deviation* | *Causes* |
|---|---|---|
| 1st | Form errors | Errors of machine tool slides, elastic deformations, fixation of tool/workpiece, severe tool wear |
| 2nd | Waviness | Eccentric rotation of workpiece/tool, vibrations, tool wear, inhomogeneity of processed material |
| 3rd | Grooves | Tool edge form, process kinematics, chip morphology |
| 4th | Cracks | Tool-nose wear, built-up-edge formation, mode of chip formation, galvanic procedures |
| 5th | Crystalline structure | Crystallization mode, irregularities due to chemical reactions, corrosive damage |
| 6th | Crystalline formation | Physical and chemical alterations in the material's fine structure, deformations of lattice |

Macro geometric deviations, such as the form errors and waviness, can be suppressed by suppressing their causes, but the micro geometric deviations (roughness) will always be present to a greater or lesser extent. Given the influence of surface roughness on function and performance, one might think that the smaller the deviation from the nominal surface, the better. However, it is essential to notice that achieving a low surface roughness deviation means increasing the machining process's effort (time and costs). Therefore, each part's application should be carefully defined to select roughness' essential range (DAVIM, 2010).

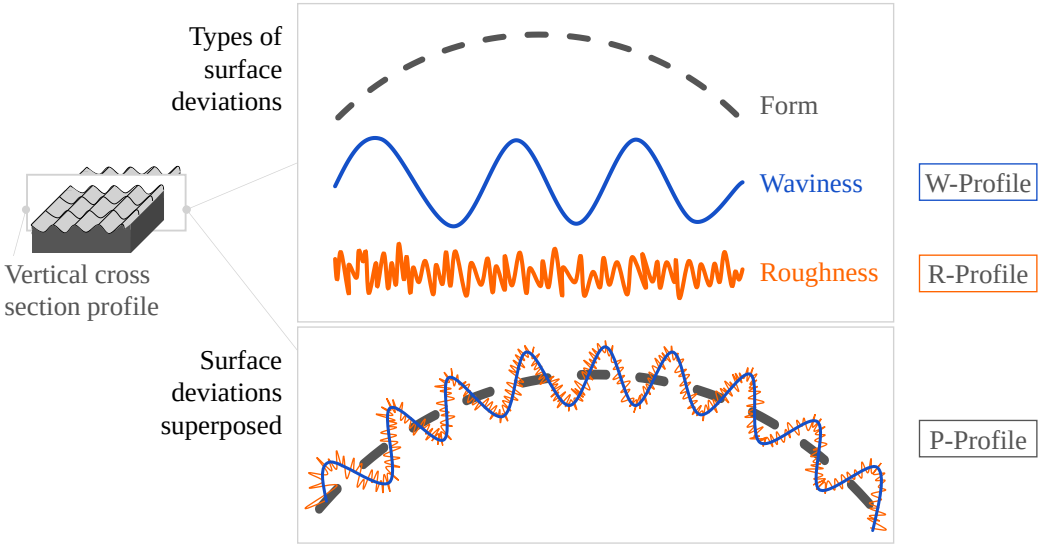



Figure 13. Types of surface geometric deviations (KLOCKE, 2011; REGO, 2020).

## 2.7 Surface parameters

Surface parameters describe surface topography—they must statistically represent the entire surface of interest. In most industrial applications, when inspecting produced parts, the surface values are assessed and addressed as quality control (KLOCKE, 2011). This work used five of the circa 60 existing surface parameters: *Ra: Arithmetic mean deviation, RSm: Mean width of profile elements, Rsk: Skewness, Rku: Kurtosis, and Rt: Total height*. The surface roughness parameters can be divided between three groups according to their functionality (GADELMAWLA et al., 2002):

- Amplitude parameters: show vertical aspects of the profile.
- Spacing parameters: show horizontal aspects of the topography.
- Hybrid parameters: show vertical and horizontal aspects of the topography.

The *arithmetic average roughness Ra* is calculated as the arithmetic mean of the absolute values above and below the mean line (ISO, 1997). It is the most established parameter to assess the roughness of machined surfaces, even adopted by almost all national standards. Its popularity is due to its facility to measure and represent the general profile. However, one should consider its drawbacks depending on the objective, i. e., it makes no distinction between peaks and valleys and is insensitive to slight deviations limiting the understanding of the profile. To this concern, other parameters shall be used in complementarity analysis (DAVIM, 2010;

GADELMAWLA et al., 2002). In addition, the extensively adopted Ra is helpful because it establishes a standard index to compare surfaces quantitatively (BOOTHROYD; KNIGHT, 2006). Table 3 shows the surface roughness parameters used in this work (directly or indirectly), their symbols, and the mathematical representation according to standards.

Table 3. Definition of the surface roughness parameters used in this work according to ISO 4287 (1997)[5].

| *Roughness parameter* | *Symbol* | *Mathematical representation* | *Eq.* |
|---|---|---|---|
| Arithmetic mean deviation | *Ra* | $Ra = \frac{1}{lr}\int_0^{lr} \lvert Z(x)\rvert \, dx$ Where $Z(x)$ are the absolute ordinate values within a sampling length ($lr$). | Eq. 9 |
| Maximum peak height | *Rp* | $Rp = \max Zp_j \ \forall \ Zp_j \in [1, m]$ Where $Zp_j$ is the height of the $j^{th}$ peak within the sampling length, and m is the number of valleys in the sampling length. | Eq. 10 |
| Maximum valley depth | *Rv* | $Rv = \max Zv_j \ \forall \ Zv_j \in [1, m]$ Where $Zv_j$ is the depth of the $j^{th}$ valley within the sampling length, and m is the number of valleys in the sampling length. | Eq. 11 |
| Skewness | *Rsk* | $Rsk = \frac{1}{R_q^3}\left[\frac{1}{lr}\int_0^{lr} Z^3(x)\, dx\right]$ | Eq. 12 |
| Kurtosis | *Rku* | $Rku = \frac{1}{R_q^4}\left[\frac{1}{lr}\int_0^{lr} Z^4(x)\, dx\right]$ | Eq. 13 |
| Mean width of profile elements | *RSm* | $RSm = \frac{1}{m}\sum_{j=1}^{m} Xs_j$ Where $Xs_j$ is the length of the $j^{th}$ profile element within the sampling length ($lr$). | Eq. 14 |
| Total height | *Rt* | $Rt = Rp + Rv$ | Eq. 15 |

[5] A new series of standards—the ISO 21920—was released at the end of 2021 combining and replacing many of the previous surface roughness standards, such as the ISO 4287 that describes the main profile parameters. However, the changes do not affect this dissertation.

The skewness (Rsk) is related to the distribution of peaks and valleys—negative values for Rsk indicate a valley predominance while positive values indicate a peak predominance. Kurtosis is related to surface asperity—values greater than 3 indicate a surface with narrower peaks and valleys, while values smaller than 3 indicate a smoother surface (MUNDIM; LUCAS, 2011).

Moreover, those surface roughness parameters' physical and functional properties are related in Table 4, showing the importance of such parameters to different applications, as commented previously in the introduction topic.

Table 4. Functional properties of the studied surface roughness parameters (DAVIM, 2010).

| *Functional properties* | *Ra* | *Rsk* | *Rku* | *RSm* | *Rt* |
|---|---|---|---|---|---|
| Contact/Contact stiffness | * | * | * | ** | ** |
| Fatigue strength | * | | * | | ** |
| Thermal conductivity | * | | | ** | |
| Electrical conductivity | * | | | * | |
| Reflexivity | | | | | ** |
| Friction and wear | * | ** | ** | * | ** |
| Lubrication | * | ** | * | | ** |
| Mechanical sealing | * | ** | | | ** |
| Fatigue corrosion | * | * | | * | |
| Assembly tolerances | * | | | | ** |

Note: two asterisks indicate a pronounced influence.

## 2.8 Surface roughness modeling in turning processes

In this work, it is essential to understand how surface roughness has been predicted for machined parts. It is a must that classical developed theoretical equations are introduced here to be compared to the models developed in this work. Moreover, it would be helpful to know the general influences of the process parameters on the surface finish to compare with the machining results.

### 2.8.1 Surface roughness modeling methods

The challenge of improving manufacturing performance and still meeting quality requirements makes practitioners and engineers (try to) predict the process behavior based on their experience and conventions that are too general to work in every condition. Research has been done to simulate complex cutting phenomena and establish cause and effect models between independent variables of the process (e.g., feed rate) with desired results' values (e.g., surface roughness) (BENARDOS; VOSNIAKOS, 2003). Regarding surface quality requirements (which is the core of this work), Benardos and Vosniakos (2003) classified the effort in predicting surface roughness into four research approaches:

(a) *Machining theory and simulation*: using numerical computation and machining theory, simulate the cutting process on tool properties and kinematics. Good results are being achieved, though the flaw is that many factors are not considered in the simulations, such as tool wear and temperature, which can affect surface results.
(b) *Experimental/empirical investigations*: extensively test the correlation between changing parameters/conditions and the surface response. This approach is the most conventional—test-observe-conclude—with good historical results. However, it is usually not replicable to other process conditions due to the high complexity of the phenomenon involved in the surface generation, which can lead to differences with minor changes in the factors.
(c) *Design of experiments*: a similar approach to (b) in the experimental and analysis procedure essence, however, being much more systematic and using the experimental planning to use the available resources optimally. Response surface methodology (RSM) and Taguchi techniques for "design of experiments" (DoE) are often used in this approach.
(d) *Artificial intelligence*: uses mainly artificial neural networks (ANNs) and genetic algorithms (GAs) to predict surface roughness in machining. Those approaches produce excellent results and stand out as having potential for use in online monitoring and decision systems.

Complementary, He, Zong, and Zhang (2018) divided the surface roughness prediction approaches into two categories:

- *Theoretical modeling methods*: models fulfilled using the physical phenomena observed in turning processes, such as the periodic duplication of cutting tool edge/waviness, material spring back, plastic side flow, and other surface formation mechanisms.
- *Empirical parametric modeling methods*: performed using data mining regarding process parameters, such as the feed rate and surface roughness responses. The most typical are the Response Surface Methodology (RSM), Artificial Neural Networks (ANN), and Support Vector Machines (SVM). Other data mining correlating factors and responses can be fitted in this category.

Both separations proposed by Benardos and Vosniakos (2003) and proposed by He, Zong, and Zhang (2018) fundamentally use the same concepts, as depicted in Table 5. However, the one proposed by He, Zong, and Zhang (2018) is more helpful to the present work due to its synthesized format. In the light of the separation proposed by He, Zong, and Zhang (2018), the models proposed in this work are empirical parametric models.

Table 5. Complementarity of propositions for the surface roughness prediction effort's separation.

| *He, Zong, and Zhang (2018) proposition* | *Benardos and Vosniakos (2003) proposition* |
|---|---|
| Theoretical modeling methods | (a) Machining theory and simulation |
| Empirical parametric modeling methods | (b) Experimental/empirical investigations<br>(c) Design of experiments<br>(d) Artificial intelligence |

### 2.8.2 Surface roughness influencing factors

An actual surface generation is a complex event in machining. It involves several factors, such as the machining method and the machining parameters (DAVIM, 2010). Usually, the values of actual surface roughness are higher than the calculated with theoretical equations due to various causes, such as chip-formation characteristics, tool wear, chatter on the machine-tool system, improper run-out on the part, and varying feed motion (DAVIM, 2010). One can

enumerate some factors that (perceived empirically) influence the response of surface roughness (HE; ZONG; ZHANG, 2018):

- Factors relative to the machine tool: depth of cut ($a_p$), feed rate (f), cutting speed ($v_c$), vibration amplitude acceleration (A/a), and vibration frequency (ω);
- Factors relative to the cutting tool: tool material, angles/inclinations, corner radius ($r_\varepsilon$), and tool edge waviness ($w_1$);
- Factors related to the workpiece material: material hardness, Young's modulus, grain size, orientation, inclusions, etc.;
- Factors related to the environment: coolant (type, pressure, etc.), temperature, and dampness.

In turning processes, the feed rate (f) is the primary influencer of the average surface roughness Ra, with a positive correlation (higher the feed, higher Ra). Figure 14 shows the general effect of feed and cutting speed on Ra. The lower values of cutting speed provoke what is called a built-up-edge (BUE) formation—workpiece material adhesion on the tool's tip that becomes harder than the tool. This formation is significantly prejudicial for the surface roughness and the tool condition. The depth of cut does not appear to have a consistent effect on surface roughness Ra (DAVIM, 2010).

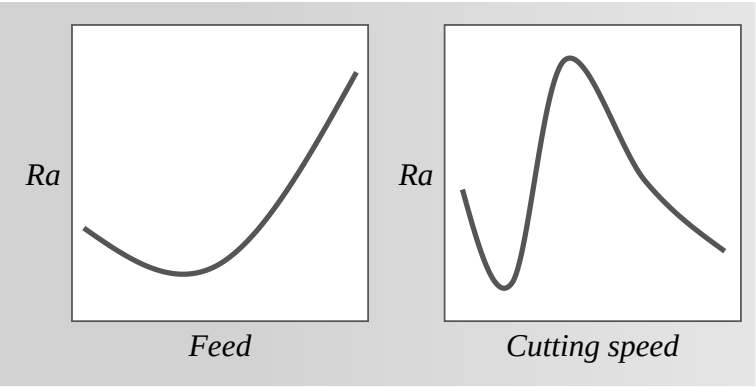



Figure 14. The general effect of the feed rate (f) and the cutting speed ($v_c$) in the average surface roughness (Ra) (DAVIM, 2010).

### 2.8.3 Surface roughness theoretical models

Surface roughness theoretical models were proposed during the last decades, mainly to the *average height of the profile Ra* and the *total height of the profile Rt.* Some of those most

known models are listed in Table 6 (BOOTHROYD; KNIGHT, 2006; DAVIM, 2010; HE; ZONG; ZHANG, 2018). Those models' factors are the feed rate (f), the corner radius ($r_\varepsilon$), and the minor and principal cutting-edge angles ($\kappa_r$ and $\kappa_r$'). Figure 15 shows ideal surfaces generated with a theoretical perfectly sharp tool (left) and a rounded tool (right). Visually, the feed and tool geometry (especially its corner radius) exert an essential role in the machined surface.

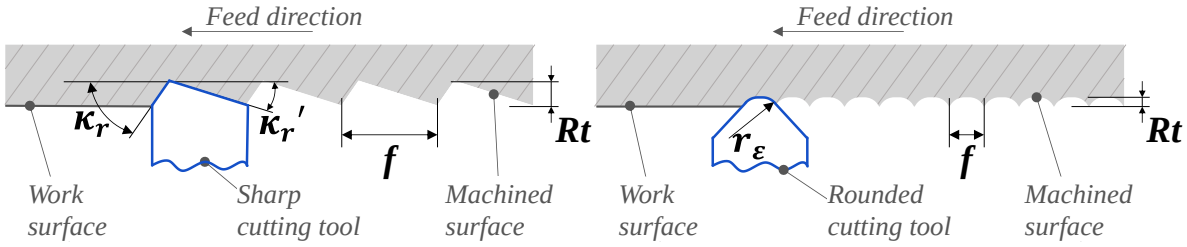



Figure 15. Surface generation influencing factors for a theoretical perfectly sharp (left) and rounded (right) cutting tool (DAVIM, 2010).

Other models, more recent, less known, and more complex in terms of computational effort, are also related to cutting dynamics factors, such as the vibration amplitude acceleration (A/a), vibration frequency (ω), minimum width chip thickness ($h_{min}$), and even an artificial variable named SE indicating cutting stability (HE; ZONG; ZHANG, 2018). In this work, however, the intention is to propose empirical parametric models, as described in the previous topic, and to compare these new models to classical simplistic theoretical models.

Table 6. Some theoretical equations of surface roughness Ra and Rt (BOOTHROYD; KNIGHT, 2006; DAVIM, 2010; HE; ZONG; ZHANG, 2018).

| *Theoretical equation* | *Description* | *Equation* |
|---|---|---|
| $Ra = \frac{0.0321 \cdot f^2}{r_\varepsilon}$ | Ra for a rounded tool and given feed rate (f) and corner radius ($r_\varepsilon$) | Eq. 16 |
| $Rt = \frac{1}{8} \cdot \frac{f^2}{r_\varepsilon}$ | Rt for a rounded tool | Eq. 17 |
| $Ra = \frac{f \cdot \tan(\kappa_r')}{4 \cdot (1 + \tan(\kappa_r'))}$ | Ra for a perfectly sharp tool. $\kappa_r$' is the minor tool cutting edge angle | Eq. 18 |
| $Ra = \frac{f}{\cot(\kappa_r) + \cot(\kappa_r')}$ | Rt for a perfectly sharp tool. $\kappa_r$ is the tool cutting edge angle | Eq. 19 |

## 2.9 Machine learning

The suggesting name of a machine that learns is precise. The concept is to program a computer to learn with data (GÉRON, 2021). Several citations of different authors are typically used to define the term. However, the definition ends up being interpreted like this most of the time. The fact that one program running in a computer—a machine—can decide whether an e-mail is or is not *spam* is an example of a machine that has learned with data. This practical tool is, actually, the example of the first machine learning application that became widespread worldwide (GÉRON, 2021).

Furthermore, traditional algorithmic programming with a detailed list of tasks is not the case for machine learning applications. Due to those systems' complexity, setting a to-do list is impossible. Machine learning applications can adapt and *learn* from their experience (SHALEV-SHWARTZ; BEN-DAVID, 2014).

Machine learning is part of the great field of artificial intelligence (AI) study, of which a good definition was proposed by Elaine Rich (1983): "Artificial Intelligence is the study of how to make computers do things at which, at the moment, people are better." For Wolfgang Ertel (2017), this definition will fit the term (AI) for decades to come. The author also elaborates that machine learning is a core component of AI because humans are still much better than computers in adaptative learning—this introduces the importance of neuroscience to the field. Intelligent systems must be built under a deep understanding of human reasoning (ERTEL, 2017). Those intelligent systems are the core of the AI goal: "AI is a field of study that attempts not only to understand how humans think but to build similar intelligent entities" (RUSSEL; NORVIG, 2010). In this context, machine learning attempts to build intelligent systems that can learn through data more deeply and with more complexity than humans, following a core proposition in the AI field of studies.

Historically, the hype of having a machine that could learn dates to long decades ago, possibly even before computers. However, not so much was developed until the 1980s. The lack of computerization capacity, follow-up research, and funding agencies support has made this technology stationary for many years. In the 1970s, the idea of machine learning appeared as it is today. To give instructions to a computer indirectly by examples, using algorithmic programming, to teach the computer how to deal with a determined challenge. At this time, the possibility of a machine with learning ability has attracted the world's attention. In 1983, a collection of scientific papers about the theme was published and followed by several other publications, such as books, research papers, doctoral thesis, and even the creation of a scientific

journal (KUBAT, 2017). Today, machine learning is offered as an undergraduate course in many universities worldwide and surrounds us as the basis for several daily-using technologies, leading high-tech products, web research, video recommendation, and even starting to drive cars. Moreover, it is used as a tool in many scientific research fields, such as bioinformatics, medicine, and astronomy (GÉRON, 2021; SHALEV-SHWARTZ; BEN-DAVID, 2014).

The fact that lots of research were done in the 1990s (KUBAT, 2017) does not mean that the most numerous publications are dated to this period. The current escalation of research based on machine learning technology is perceived in the exponential-like curve depicted in Figure 16. The flourishing curve shows the importance of this technology to the global community and people's consciousness about it. Moreover, it suggests a prominence of the field, with much more to develop and discover. Machine learning is growing fast, partially because the available data is getting bigger and partially because the theory to turn data into knowledge improves yearly (ALPAYDIN, 2014).

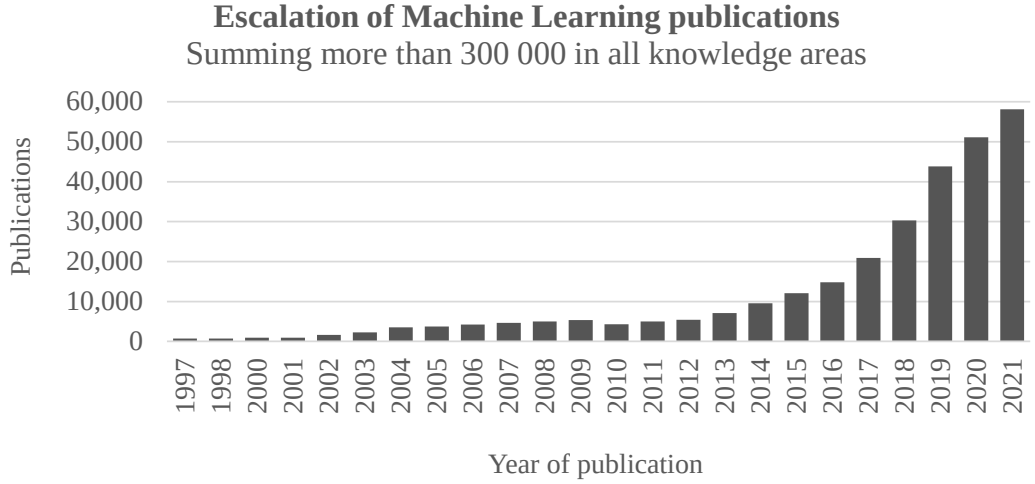



Figure 16. Publications regarding Machine Learning thematic through the years (Web of Science).

### 2.9.1 Machine learning basics

One can categorize any machine learning system by three methods: type for *learning*, *incrementing* and *generalizing* (GÉRON, 2021). Although the explanation of each follows below, in advance: the current dissertation uses a supervised, mini-batch, instance-based machine learning system.

The *learning* method (the *learning paradigm*) defines whether a machine is taught with or without supervision, and can be supervised, unsupervised, semisupervised, and

Reinforcement learning (GÉRON, 2021). *Supervised learning* means giving labeled examples, each as a unique input signal and a corresponding response so that the model can learn (HAYKIN, 1999). Linear regression is a case of an algorithm that uses supervised learning. Also, many machine learning algorithms use supervised learning, such as *K-Nearest Neighbors*, *Support Vector Machines*, *Decision Trees* and *Random Forest*, and *Artificial Neural Networks* (GÉRON, 2021).

As the name suggests, not supervising (unsupervised learning) means feeding data to the algorithm without answering the questions. Géron (2021) comments that unsupervised learning is "the system learning without a teacher." Many other machine learning algorithms use this type of learning, such as *K-Means* for clustering (dividing groups), *t-distributed Stochastic Neighbor Embedding* (t-SNE) for visualization, *Principal Component Analysis* for dimensionality reduction, and *Apriori* for association rule learning. For the manufacturing industry, a helpful method of unsupervised learning is anomaly detection: training the model with standard data to posteriorly detect whether new data fed to it is normal or not, potentially detecting anomalies in machines or processes (GÉRON, 2021).

Semisupervised learning mixes both types mentioned above, usually having many unlabeled examples and a few of the examples given to the model with one label to each feature. Reinforcement learning is a type of learning that uses "penalties" and "rewards" as a basis for learning (GÉRON, 2021). This model type intends to output a list of actions (named *policy*) to reach the objective correctly (ALPAYDIN, 2014).

The *incrementing* method defines whether the system is fed with information during the learning process (*online learning*) or whether it receives data before and works on the process without learning new information (offline incrementation or *batch-learning*). The model is trained with the entire dataset in batch learning before applying it. When new data must be learned, the model must be trained from zero again (and offline) to fit the new data. Only then can it be applied again (GÉRON, 2021). Nevertheless, there is more meaning-information about online and offline/batch learning than its suggesting name. Batch learning starts from the principle that there is a function correlating the inputs ($X$) with the outputs ($y$), and it generalizes the response using the whole training dataset of labeled examples ($X$, $y$). Online learning uses one input (or small sets of samples, also called mini-batches) at a time to predict the output. It then calculates the *actual y* and the *predicted y* difference error and corrects the prediction model for the following samples (DEKEL, 2008). Therefore, online learning is usually performed offline (outside live production) or online (inside live production).

The *generalizing* method is another criterion for group machine learning types. It can be instance-based or model-based. Instance-based learning means the pairs of labeled examples are given to the machine, so it learns which response (a.k.a., variable, output, *y*) is the more correct to each question (a.k.a., feature, input, *x*). Then, it chooses the respective response for a new similar question. Model-based learning means creating a function that correlates the inputs and outputs to predict responses to new data. The model can be evaluated and self-optimized by increasing a *utility function* (what can be done to improve the model's fitness) or by decreasing a *cost function* (what can be done to decrease the model's errors) (GÉRON, 2021).

## 2.10 Artificial neural networks

Although the first works regarding this topic date back to the 1960s, the most prominent research studies started in the 1990s (SILVA; SPATTI; FLAUZINO, 2016). The motivation comes from understanding the outstanding efficiency and capacity of the brain and from the attempt to, using artificial computing systems, mimic the way it performs a particular task. Remarkably, the human brain behaves much differently from a standard computer. The brain is a highly complex, nonlinear, and parallel computing, and neural systems (even in small insects) are the most efficient and adaptable information-processing devices (HAYKIN, 1999; ROZENBERG; BÄCK; KOK, 2012). Biological neurons' organization can do several computations, such as face recognition and motor control, in much less time than a conventional digital computer (HAYKIN, 1999).

Moreover, brain *plasticity* allows the neural system to constantly self-develop and adapt to its surroundings. This behavior is as crucial to the natural neural system as it is to the artificial ones constructed by artificial neurons (or processing units) (HAYKIN, 1999).

Artificial neural networks (ANN) can be described as a machine made of simple processing units that acquire experiential knowledge from their surroundings through (a process called) *learning*, storing it by neuron-neuron connection strengths called *weights* (HAYKIN, 1999).

To fully comprehend natural neural networks, it would be necessary to have multidisciplinary knowledge, collecting information from biology, mathematics, and artificial intelligence. However, despite the historical inspiration for the name, many are the differences between artificial and biological neural nets. The brain's neural system modeling is not the goal of ANNs, and some ANN can even be dissociated entirely from the biological neural structure

(NASCIMENTO JR; YONEYAMA, 2000). While artificial neural networks can have a few neurons or many hundreds, dividing the brain into 1000 modules, imagine each module containing circa 500 million neurons with many involved in ordinary activities. This perspective can be used to understand the magnitude of the difference between artificial and biological neural nets. It is not just about quantity, however: the Eccles law that applies to natural neural nets (each neuron either excites or inhibits its connected neighbors) does not apply to artificial nets; the speed processing is different; there are several different neuron types in the brain, while typically, the ANN is implemented with only one type of processor unit; the biological synapsis can occur in different neuron locations and can be by chemical or electrical action. Those are some differences between natural and artificial neural systems (EBERHART; DOBBINS, 1990; ROZENBERG; BÄCK; KOK, 2012).

Given the differences, it is clear that it is possible to use artificial neural networks as just another analysis tool to solve problems without a deep understanding of its biological inspiration (it will undoubtedly require programming knowledge and effort, though) (EBERHART; DOBBINS, 1990). Neural network power as a tool comes from its parallel processing structure and ability to learn and generalize. *Generalization* is the skill to produce reasonable outputs from never-seen-before inputs (the machine is learning) and is an important feature that highlights this tool from others (HAYKIN, 1999).

Moreover, to visualize neural networks as a tool is also to put in perspective that the effort (or the code lines) of ANNs is only a part of the whole problem-solving. Preprocessing and manipulating data and further analysis and visualization have a crucial impact on the entire figure. Figure 17 shows the proportional modeling role (that extends to ANNs) in machine learning applications. However, despite being small in perspective, it is indispensable (EBERHART; DOBBINS, 1990).

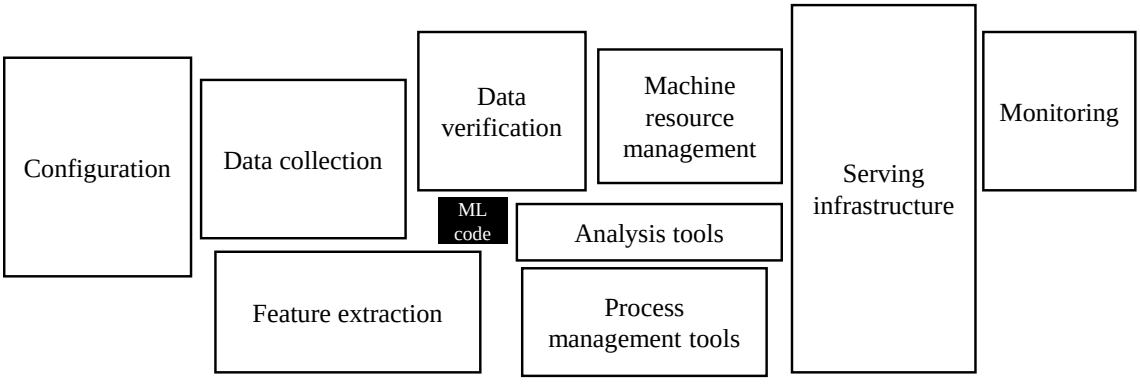



Figure 17. Representation of real-world vast and complex machine learning systems. Modeling is a small part of the whole picture (SCULLEY et al., 2015).

As commented above, there are many differences between natural and artificial neural networks. However, it is helpful to visualize the typical natural neuron and the general working mechanism of the neural system to understand the function of artificial ones. The biological nervous system works like the three-stage diagram depicted in Figure 18. Stimuli from the human organism or the exterior surroundings (inputs) are converted into electrical impulses by the receptors and transmitted to the neural net (the brain) that makes decisions. The effectors convert electrical impulses into responses to the organism. Information is fed forward and back through the system during processing (HAYKIN, 1999).

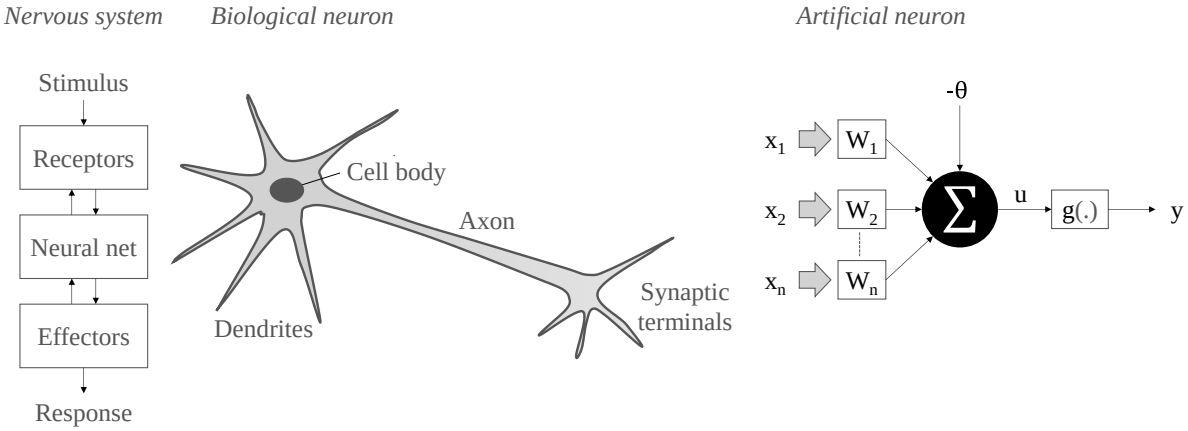



Figure 18. Representation of the nervous system (left), the biological neuron (middle), and the artificial neuron (processing unit) (HAYKIN, 1999; SILVA; SPATTI; FLAUZINO, 2016).

The neuron, depicted in the middle of Figure 18, is the elementary cell of the nervous system. Its job is to transmit the electrical impulses through the network. The dendrites receive input from the exterior or adjacent neurons and transmit the impulse to the neuron cell body that processes information and generates an activation potential. The impulse is then forwarded along the axon to the synaptic terminal to go to another neuron or directly to the organism (SILVA; SPATTI; FLAUZINO, 2016).

The *artificial neuron model* is based on a simple biological neuron and can be described using the diagram on the right side of Figure 18. The *x* represents an *input*, just as the stimulus received by the dendrites in a biological neural system. The *W* represents the synaptic *weights* that correspond to the strength or relevance of the input signal for the neuron functionality. The signal *u* gets out as a *weighted sum* of the inputs. An externally applied *bias* $\theta$ is responsible for setting a plateau of activation. If it is negative, the neuron produces inhibitory potential, and if it is positive, it produces an excitatory potential. The transformation of *u* by an *activation function g()* produces the *output y* (SILVA; SPATTI; FLAUZINO, 2016). The structure of the

artificial neuron can be represented mathematically by both Eq. 20 and Eq. 21 (SILVA; SPATTI; FLAUZINO, 2016).

$$u = \sum_{i=1}^{n} W_i \cdot x_i - \theta \quad \text{Eq. 20}$$

$$y = g(u) \quad \text{Eq. 21}$$

Artificial neural networks can be classified in three manners: the activation function that transforms the weighted sum of inputs into output; the architecture representing how processing units organize themselves in a given net; and the learning algorithm used for the ANN training (EBERHART; DOBBINS, 1990).

### 2.10.1 Activation functions

The artificial neuron ruled by Eq. 20 essentially multiplies the weights by the inputs and adds the sum to a bias value, which is essentially a linear model like Eq. 22 (where *y* and *x* are variables and *a* and *b* are any real numbers). Therefore the activation function has a crucial role in artificial neural network models: it introduces nonlinearity (JAIN, 2019; LECUN et al., 2012; SHARMA; SHARMA; ATHAIYA, 2020). The ability to model nonlinear relations is one of the most significant benefits of ANNs that make them learn very complex phenomena (HAYKIN, 1999). An ANN that does not use a nonlinear activation function performs as a simple linear regression model (SHARMA; SHARMA; ATHAIYA, 2020).

$$y = a \cdot x + b \quad \text{Eq. 22}$$

The activation function plays other roles. It can restrict the output limits, and more importantly, activate the neuron (SHARMA, 2017; SILVA; SPATTI; FLAUZINO, 2016). Simply put, the activation function decides whether that neuron is essential to the prediction model. If it is essential, it becomes activated (or more activated)—usually a value close to 1 or higher—and if not, it becomes inactive (or less activated)—usually a value close to zero.

Some of the most usually implemented functions and their mathematical representation are listed in Table 7. Their graphical representations are depicted in Figure 19. Since the beginning of ANN research, the most used functions were sigmoidal functions (hyperbolic tangent is also a sigmoidal function). Recently, the rectified linear unit (ReLU) that was proposed by Nair and Hinton (2010) has become very popular and one of the most widely used activation functions, mainly due to its faster learning in neural nets with many layers (LECUN;

BENGIO; HINTON, 2015). The ReLU is a linear function for positive inputs and zero for negative ones. This function generally works well for all hidden layers (GÉRON, 2021).

Table 7. Activation functions and their mathematical representation.

| *Function* | *Acronym* | *Mathematical representation* | *Equation* |
|---|---|---|---|
| Sigmoid | - | $y = g(u) = \frac{1}{1 + e^{u}}$<br>Where $g$ is the activation function, $y$ is the neuron output, and the $u$ is the weighted sum of the inputs. | Eq. 23 |
| Hyperbolic tangent | Tanh | $g(u) = \frac{e^{u} - e^{-u}}{e^{u} + e^{-u}}$ | Eq. 24 |
| Rectified Linear Unit | ReLU | $g(u) = \max(0, u) = \begin{cases} u, & u > 0 \\ 0, & u \leq 0 \end{cases}$ | Eq. 25 |
| Exponential Linear Units | ELU | $g(u) = \begin{cases} u, & u > 0 \\ \alpha \cdot (e^{u} - 1), & u \leq 0 \end{cases}$<br>Where $\alpha \approx 1.67326324$ | Eq. 26 |
| Scaled Exponential Linear Units | SELU | $g(u) = \begin{cases} \lambda u, & u > 0 \\ \lambda \cdot \alpha \cdot (e^{u} - 1), & u \leq 0 \end{cases}$<br>Where $\lambda \approx 1.05070098$ | Eq. 27 |

After ReLU was first proposed in 2010, the activation functions became a theme of intense research (URBAN, 2018), and today, new functions are being tested. This work used classical functions such as the sigmoid and the popular ReLU and its derivatives—ELU and SELU—readily available at the *Python Keras API*. The ELU can output negative values—avoiding the "dying ReLU" problem[6]—and is addressed to be faster and with better generalization performance than the ReLU (NWANKPA et al., 2018). The SELU has the curious function of self-normalization (NWANKPA et al., 2018), which is an attractive characteristic (in this work, this feature did not configure an advantage because the inputs were previously normalized).

[6] The "dying ReLU" problem is a typical problem of this activation function, in which some neurons can be permanently inactive (dead), generally related to high learning rates (LEUNG, 2021; LU et al., 2020).

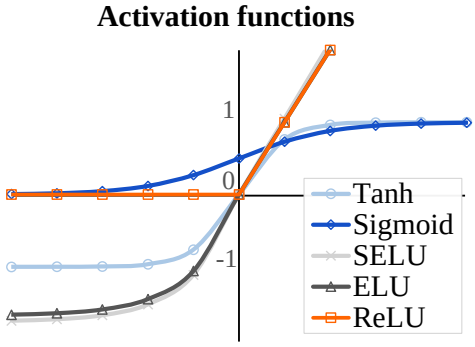



Figure 19. Activation functions and their graphical representation.

### 2.10.2 Architecture

Network architecture is how neurons are structured inside the net (HAYKIN, 1999). Three architectures are generally used: single-layer feedforward networks, multiple-layer feedforward networks, and recurrent networks (HAYKIN, 1999).

The present work makes use of a multiple-layer feedforward network. This type of ANN is typically represented as depicted in Figure 20. The multilayer stack is composed of the input layer that receives the inputs vector and has the same number of neurons as the number of inputs, the output layer that presents the outputs and has the same number of neurons as the number of outputs, and the layer (or layers) in-between—the hidden layers—where most of the ANN processing is performed. The hidden layers increase the capability of the ANN to extract higher-order statistics and to perform a more complex mapping between the inputs and the outputs due to the additional synaptic connections and more neuron interaction (HAYKIN, 1999; SILVA; SPATTI; FLAUZINO, 2016). Each layer's neurons are usually subject to learning and computing non-linear input-output mappings like the artificial neuron depicted in Figure 18 (LECUN; BENGIO; HINTON, 2015).

The neural net dimension can be represented by size, depth, and width (SHALEV-SHWARTZ; BEN-DAVID, 2014). The size is the sum of all neurons, the depth is the sum of all hidden layers and the output layer, and the width is the maximum number of neurons in the hidden layers. In the example of the multilayer feedforward neural net in Figure 20, the size is 7, the depth is 2, and the width is 3.

This work conventionally presented the topology as “number of input neurons–number of neurons on the $n^{th}$ hidden layer–number of neurons in the output layer.” Therefore, the neural net of Figure 20 topology is 3-3-1.

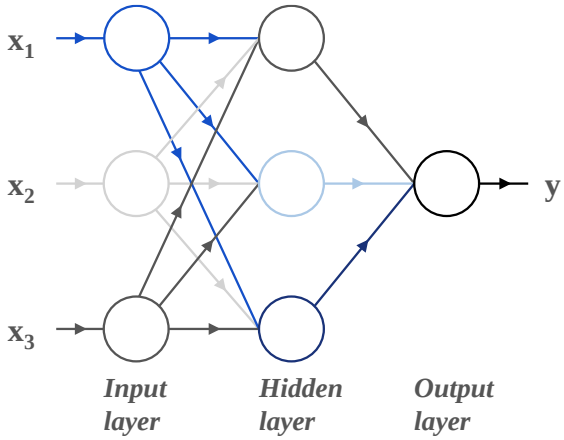



Figure 20. Example of multilayer feedforward neural network with one hidden layer.

### 2.10.3 The learning process

The essence of learning (Figure 21) is finding the network’s suitable adjustable parameters (weights and biases) that minimize the error. A cost function computes this error regarding the predicted and desired outputs. The learning problem is approached generally with gradient-based learning methods (LECUN et al., 2012).

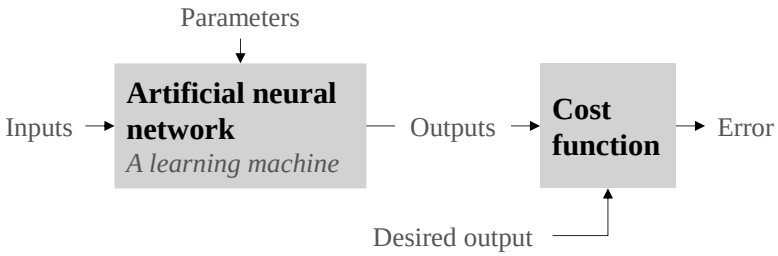



Figure 21. The simplest form of the learning problem for a learning machine (LECUN et al., 2012).

A part of the learning algorithm used in this work is the back-propagation learning, considered the pillar of modern artificial neural network learning (KOSTADINOV, 2019; NIELSEN, 2015). The back-propagation algorithm idea appeared independently in several

research papers from the 1970s to the 1980s, and most of its importance was fully consolidated by Rumelhart, Hinton, and Williams (1986) (LECUN; BENGIO; HINTON, 2015).

The back-propagation learning algorithm essentially performs (HAYKIN, 1999): 1. a forward pass (after the inputs' vector is applied to the input layer) *fixing* the weights (generally by randomly selecting values close to zero) layer by layer until a set of outputs is generated; then, 2. it calculates the difference between the predicted and the actual output producing an error signal—through a cost function that for regression problems can be the mean squared error; and then 3. it performs a backward pass calculating the cost function gradient that will be subsequently used by another algorithm to update the network parameters aiming at minimizing the error.

A fundamental characteristic of the back-propagation algorithm is that it calculates the cost function gradients (first-order derivatives of the cost function regarding the weights) in the backward pass. Attention must be paid to this characteristic in some cases to avoid problems like "vanishing gradients," "dying ReLUs," and "exploding gradients" (KARPATHY, 2016).

While the back-propagation algorithm computes the gradients that indicate the direction of weights update, there is the need for another algorithm to implement those weight changes, and this is done with the optimizer algorithm. While the back-propagation is built-in with the *Keras Sequential API* used in this work, the optimizer must be chosen within a variety of options readily available.

An optimizer is an algorithm that changes the learnable parameters of the network (using the gradient calculated by the back-propagation algorithm), i.e., weights and biases. Of the several options of optimizers, the stochastic gradient descent (SGD) and its variants are most likely the more popular for deep learning applications (GOODFELLOW; BENGIO; COURVILLE, 2016). The SGD interactively updates the network parameters using a subset of the training dataset, called mini-batch. The mini-batch size is called the batch size (ALBANESI; VAMOSSY, 2019).

Optimizers have changeable features which can significantly affect the learning process. For the more general case, in the SGD, two features can be edited: the learning rate ($\eta$) and the momentum. The learning rate is a constant that directly affects the algorithm's speed towards the cost function minimum. Generally, if this constant is too small, the learning can converge too slowly, and if it is too large, maybe the algorithm never finds the minimum (LECUN et al., 2012). Learning rates usually range from 0.0001 to 1 (TAYLOR, 2017). The momentum increases the learning speed by changing the learning rate during the process.

This work intends not to show (or manually use) the learning process's mathematical operations. However, an exception will be open, for it is helpful to see the intuitive equation of the weights updates by the SGD optimization algorithm (Eq. 28). From it, one can understand how fundamentally the weights are updated and how the learning rate directly affects the learning process speed—or, in other words, how a small learning rate would lead to more epochs of neural network training.

In Eq. 28 (TAYLOR, 2017), the new weight $W_{new}$ is calculated as the former weight W subtracted by the product of the learning rate η and the partial derivative of the error regarding the former weight.

$$W_{new} = W - \eta \cdot \frac{\partial E}{\partial W} \qquad \text{Eq. 28}$$

To summarize: the learning algorithm is generally formed by 1. the back-propagation algorithm, responsible mainly for computing the gradient of the cost function, and 2. an optimization algorithm (usually based on gradient descent), responsible mainly for updating the network parameters using this gradient (GOODFELLOW; BENGIO; COURVILLE, 2016). Then, it can be said that a given model uses "stochastic gradient descent with back-propagation" as its learning algorithm.

### 2.10.4 Usual application procedures and precautions

While essentially all the mathematical operations are available as off-the-shelf software and do not have to be performed by the machine learning practitioner, some procedures, decision-making, and precautions must be taken when designing and training machine learning models.

When building a model, a practitioner must deal with hyper-tuning. A neural network is composed of parameters (weights and biases) adjusted by the learning algorithm, as discussed in the previous topic 2.10.3. However, several other features can be tweaked and tuned to optimize the network's performance. Those other features are called hyperparameters (HUTTER; KOTTHOFF; VANSCHOREN, 2019). While the network can learn the parameters, the hyperparameters must be manually set (hyper-tuned) (SKANSI, 2018).

Some hyperparameters can be (MONTAVON; ORR; MÜLLER, 2012): the learning rate, the batch size, the momentum, the number of hidden units in a hidden layer, the number of hidden layers, the activation function, the random seed, or the features preprocessing method.

One must think that selecting the right hyperparameter (HP) would be testing all the possible combinations. This procedure is called grid search. However, for there can be several different hyperparameters, the experiment can escalate quickly. For instance, with a range of 5 HP, each varying in 3 levels, 243 tests would have to be performed. This experiment already seems to have a large but feasible set of tests. It happens that, for hyperparameters, the number of factors and levels is usually higher. One can find the need to test 10 HP varying in 5 levels, a combination that would demand almost 10 million tests.

Therefore, training a model for the whole set of different hyperparameters is often impractical. Nevertheless, a reasonable estimate of the broader behavior of the model to a given hyperparameter can be achieved in a feasible time by training the model with a shorter number of epochs and observing the results at the early stages of learning (AGGARWAL, 2018; SMITH, 2018). Several other approaches are described in the literature, such as *Random Search* and *Bayesian Optimization*, examples that can be performed using the *Python Keras API*. In conclusion, general performance depends significantly on good hyperparameter choosing—however, there is no simple and straightforward way of selecting hyperparameters yet (SMITH, 2018).

Another recommended procedure before training the model is preprocessing the features (inputs). The preprocessing is usually made through *standardization* (transforming the data into a set of zero mean and unit variance) or *normalization* (scaling the data inside an arbitrary range of generally 0–1) (AGGARWAL, 2018). Besides being necessary for some occasions where the features have significant differences in their ranges, it can increase learning convergence speed (LECUN et al., 2012).

The holdout method, another general procedure, divides the dataset into three subsets (Figure 22 (B)). Those subsets are the training set: values used by the learning algorithm; The validation set: used to evaluate the model generalization during learning—and avoid overfitting; and the test set: used to evaluate the model prediction performance (TAYLOR, 2017). The learning algorithm only “sees” the “training dataset without validation set” so that it can be tested in never-seen-before values.

A precaution must be taken to avoid overfitting and underfitting in machine learning. Learning algorithms are mainly based on gradient descent, and while the learning convergence can optimize the loss of the training data, it does not mean that it optimizes the loss of never-seen-before data (AGGARWAL, 2018). For this reason, as mentioned in the above paragraph, one might want to separate a validation dataset to check the model even before the final test on the test dataset (MONTAVON; ORR; MÜLLER, 2012). The validation loss computed during

learning can be graphically compared to the training loss and can be used to implement the *early stopping* technique—stopping the training as soon as the validation error increases (MONTAVON; ORR; MÜLLER, 2012).

Figure 22 (A) shows two graphs. From left to right, the learning curves of a good model (the validation loss converging similarly to the training loss); and the three stages of learning, underfitting, good fitting, and overfitting—that happens when the validation loss starts to increase while the training error continues to decrease or stays at the horizontal plateau (GÉRON, 2021; MONTAVON; ORR; MÜLLER, 2012; SMITH, 2018).

Figure 22 (B) shows the usual division of the dataset. The first split divides the original labeled dataset into training and test sets; then, the training set is split into training without validation—used for learning—and validation sets—used for hyperparameter tuning. In the end, the test dataset is used to evaluate the model with unseen values.

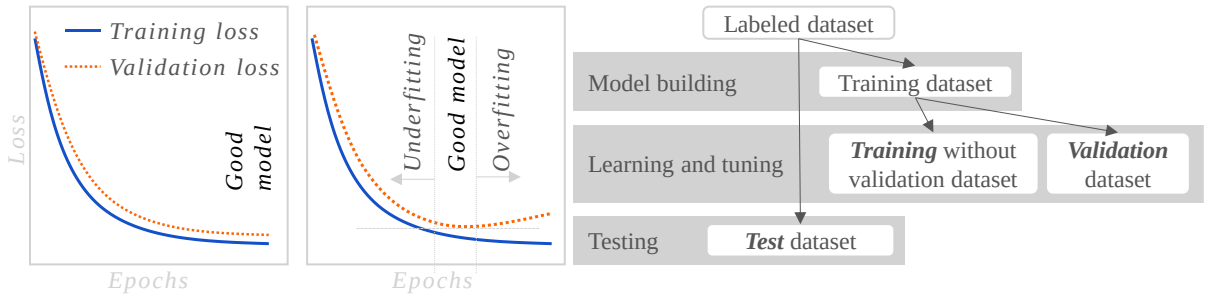


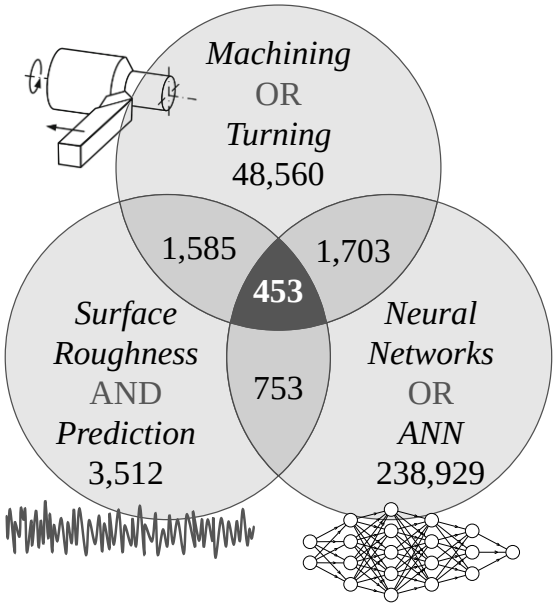

Figure 22. Practical observations of learning curves (A) (GÉRON, 2021; MONTAVON; ORR; MÜLLER, 2012; SMITH, 2018) and usual dataset divisions (B) (AGGARWAL, 2018).

# 3 Bibliometric Study and Examination of Recent Developments

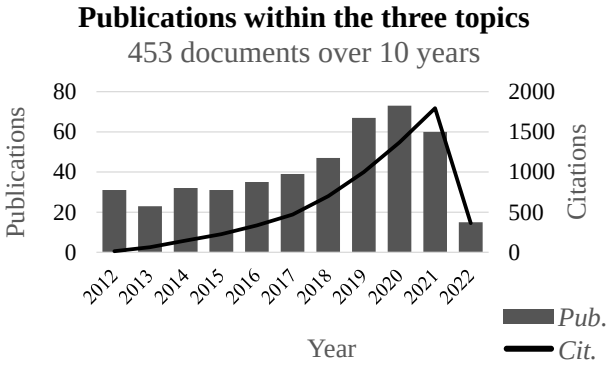
In order to perform proper research work on the thematic of surface roughness prediction in turning using ANNs, it is essential to look up what people in scientific institutes worldwide were capable of doing. A good approach would be to find an opportunity to contribute to the international research history line. Luckily, this is facilitated by several research bases, such as Scopus, Web of Science, and many others. Then, a bibliometric search was made to understand some characteristics of the thematic, the agents and the prominent publishers, the status quo, and what can be explored in the current studies regarding surface roughness prediction. The framework adopted for this literature review was based on the one proposed by Mundim (2018) in his doctorate thesis.

With bibliometric results (Figure 23), one can see that there is, for example, much work being done with neural networks in the fields of engineering and materials science, which follows the exponential-like curve seen in Figure 16 that indicates a substantial growth in machine learning developments during the last ten years. Publications within the three topics (the interception group), equal to this study's thematic, make up 453 documents, and slight growth in the number of publications can be seen from 2012 to 2021 (2022 goes only to February). Moreover, using only the term "turning" as a search entry, this number of publications drops to 172. Web of Science makes it easy to look at the countries and affiliations producing the publications. India and China currently lead the research in this thematic, and the affiliation with more published papers is the National Institute of Technology (NIT)—which forms a network of institutes across India.

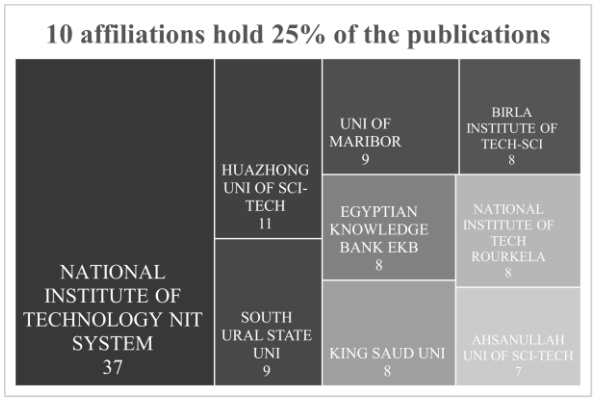
Web of Science makes it easy also to analyze how those 453 papers have been cited. Such analysis indicates that besides the growing production (bar-graph), this topic has a crescent consumption (represented by the crescent curve of citations). In addition, the treemap at the bottom of Figure 23 shows where comes from the consumption of such publications. Despite India and China together summing up to more than 50% of the citations, other 23 countries have 50 or more citations, and other 73 countries cited at least once—suggesting that the topic is of interest also from the rest of the world.

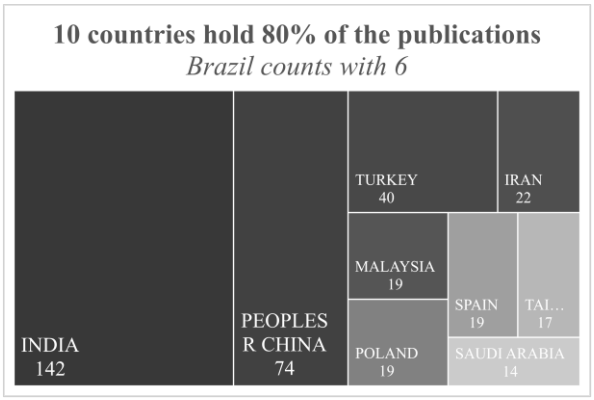



**Research criteria**
**Data base**: Web of Science (Clarivate)
**Period**: 01.2012 – 02.2022
**Research areas**: Engineering, Materials science, Computer science




*From where publications come from (production)*




*From where citations come from (consumption)*

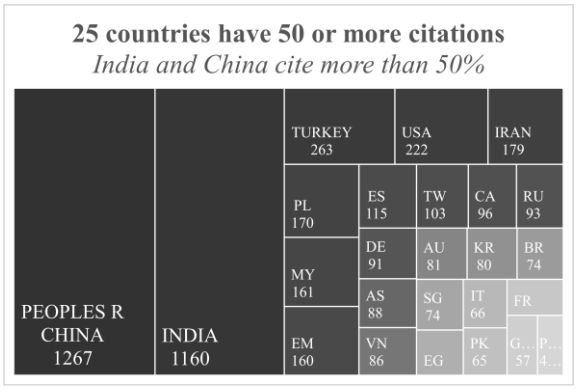



Figure 23. The bibliometric search of "surface roughness prediction in machining using neural networks."

Reading 453 papers is not an easy task, and a good literature review depends on reading the works that contribute most to the thesis, hence the importance of choosing good criteria to refine the number of relevant publications.

One criterion to refine could be to choose the publications more recently published. However, the publication history depicted on the right side of Figure 23 indicates that not many discoveries are being delivered from year to year. Therefore, excluding the papers from 2012–2016 could lead to the missed opportunity to read actual papers. A publication history like the machine learning topic (back in Figure 16) indicates that much effort and novelties are being published yearly because of the exponential behavior.

In the end, a mix of criteria was used. The first filter (Figure 24) retained 278 papers and let pass 175 by reading their abstracts and subjectively deciding their relevance to this thesis. The second filter used criteria such as the machine learning approach, the machining process, the workpiece material, whether it measured machining forces and tool wear, the number of citations, and the novelty aspect of the papers. With this, 55 documents were selected to be read entirely, and at last, 19 papers will appear in this literature review. Also, in the middle of the reading process, some other cited papers published outside the observation period were relevant and were also cited. From reading 19 papers, some conclusions were drawn in the subsequent paragraphs.

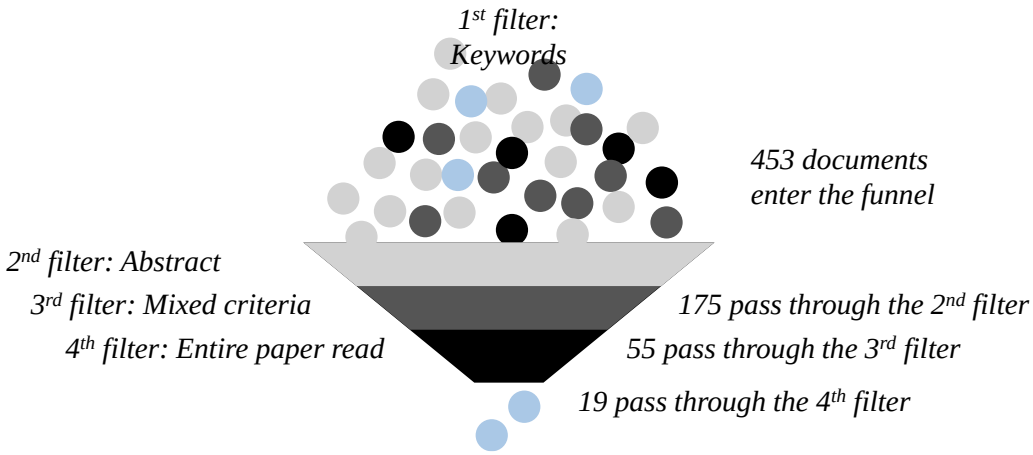



Figure 24. Choosing fitting papers to read for literature review.

Of all studied works, the errors achieved by the ANN prediction on never seen data were relatively low. Çaydaş and Ekici (2010), Mia et al. (2017), and Mia and Dhar (2016) declared coefficients of correlation (R) of more than 0.90. Garg and Tai (2014), Jurkovic et al. (2018), Laouissi et al. (2018), and Meddour et al. (2018) reached mean absolute percentage errors

(MAPE) between 5% and 15%. Some works achieved surprisingly low errors: Meddour et al. (2018) achieved a root mean squared error RMSE = 0.0005 and an R-squared of 0.99 using only 23 samples to train the model; Asiltürk (2012) achieved an MSE = 0.0001 and R-squared of 0.99 using a training dataset of 600 samples; Kara et al. (2020) achieved a root mean squared percentage error RMSPE = 0.07% with a training dataset of 130 samples; Beatrice et al. (2014) achieved RMSE = 0.02 using only 23 training samples.

Several authors have used a training set smaller than 55 samples, achieving low error results (BEATRICE et al., 2014; ÇAYDAŞ; EKICI, 2012; CHEN et al., 2017; GARG; TAI, 2014; LAOUISSI et al., 2019; MEDDOUR et al., 2018; MIA et al., 2017, 2018; MIA; DHAR, 2016; MUTHUKRISHNAN; DAVIM, 2009; SENTHILKUMAAR; SELVARANI; ARUNACHALAM, 2012). However, Pontes *et al*. (2012) compared the performance results of different ANN models, each of which used a different number of labeled training examples (varying from 20 to 500 samples), and suggested that, generally, the performance of an ANN in predicting Ra is improved when more training samples are available.

Most often, papers clearly showed the ratio between training and test datasets. However, most of them did not use a validation dataset or have not mentioned its use. Very few made clear the use of validation datasets to test generalization errors during training (ABBAS et al., 2018; CHEN et al., 2017; SENTHILKUMAAR; SELVARANI; ARUNACHALAM, 2012).

Most of the papers used neural networks with only one hidden layer with up to 17 hidden neurons (ABBAS et al., 2018; ÇAYDAŞ; EKICI, 2012; CHEN et al., 2017; GARG; TAI, 2014; JAFARIAN; TAGHIPOUR; AMIRABADI, 2013; JURKOVIC et al., 2018; KARA et al., 2020; LAOUISSI et al., 2019; MIA et al., 2017; MIA; DHAR, 2016; MUTHUKRISHNAN; DAVIM, 2009; SENTHILKUMAAR; SELVARANI; ARUNACHALAM, 2012). Some works used more than one hidden layer or more than 20 hidden neurons: Asiltürk (2012) used one hidden layer with 42 hidden neurons, and Beatrice et al. (2014) and Meddour et al. (2018) used two hidden layers with seven hidden neurons each and ten hidden neurons each, respectively.

In the majority, it was used the same three inputs: depth of cut ($a_p$), feed rate (f), and cutting speed ($v_c$) to predict some of the surface roughness parameters (ABBAS et al., 2018; BEATRICE et al., 2014; ÇAYDAŞ; EKICI, 2012; JAFARIAN; TAGHIPOUR; AMIRABADI, 2013; JURKOVIC et al., 2018; LAOUISSI et al., 2019; MUTHUKRISHNAN; DAVIM, 2009; PONTES et al., 2012; SENTHILKUMAAR; SELVARANI; ARUNACHALAM, 2012). However, some added different factors: Mia et al. (2017), Mia and Dhar (2016), and Saric, Simunovic, and Simunovic (2013) also included the way of cooling as an input—wet and dry cut, coolant through the tool or outside, or different uses of MQL; Kara et al. (2020) also used

different cutting tools and workpieces with varying heat treatments as inputs; Garg and Tai (2014) used the cutting time, depth of cut and cutting speed as inputs; Asiltürk (2012) and Meddour et al. (2018) also included the cutting tool edge radius ($r_\varepsilon$) as input; and Chen et al. (2017) used the depth of cut, feed rate, and cutting speed to predict the three components of the machining force ($F_x$, $F_y$, $F_z$) and the three components of tool vibration ($V_x$, $V_y$, $V_z$) as an enclosed ANN, and subsequently used all the nine factors as inputs to calculate the surface roughness Ra.

Regarding the output, most of the reviewed works built ANN models to predict the surface roughness Ra (ABBAS et al., 2018; ASILTÜRK, 2012; BEATRICE et al., 2014; CHEN et al., 2017; JAFARIAN; TAGHIPOUR; AMIRABADI, 2013; JURKOVIC et al., 2018; KARA et al., 2020; LAOUISSI et al., 2019; MEDDOUR et al., 2018; MIA; DHAR, 2016; MUTHUKRISHNAN; DAVIM, 2009; PONTES et al., 2012; SARIC; SIMUNOVIC; SIMUNOVIC, 2013; SENTHILKUMAAR; SELVARANI; ARUNACHALAM, 2012), following the logic that Ra is the surface roughness parameter most used in practice—hence the more demanded prediction model.

Some papers, though, included other outputs to the prediction models: Jafarian, Taghipour, and Amirabadi (2013) made the model with different inputs: surface roughness Ra, the resultant machining force, and the cutting tool flank wear width (VB); Senthilkumaar, Selvarani, and Arunachalam (2012) also predicted the tool condition in terms of flank wear width (VB) besides the surface roughness Ra; and Abbas et al. (2018) also predicted the machining time and cost.

Several works employed a trial-and-error procedure to select the hyperparameters of the neural network (BEATRICE et al., 2014; CHEN et al., 2017; GARG; TAI, 2014; MEDDOUR et al., 2018; MIA et al., 2017); some used design of experiments strategies (KARA et al., 2020; PONTES et al., 2012; SARIC; SIMUNOVIC; SIMUNOVIC, 2013); and cross-fold-validation (ASILTÜRK, 2012). Not to mention those that left the approach to hyper-tuning unclear (ABBAS et al., 2018; ÇAYDAŞ; EKICI, 2012; JURKOVIC et al., 2018; LAOUISSI et al., 2019; MIA; DHAR, 2016; MUTHUKRISHNAN; DAVIM, 2009; SENTHILKUMAAR; SELVARANI; ARUNACHALAM, 2012).

The sigmoidal variations are the most used activation functions (in hidden or last layers) (ABBAS et al., 2018; ASILTÜRK, 2012; BEATRICE et al., 2014; CHEN et al., 2017; GARG; TAI, 2014; JURKOVIC et al., 2018; KARA et al., 2020; LAOUISSI et al., 2019; MEDDOUR et al., 2018; MIA; DHAR, 2016; MUTHUKRISHNAN; DAVIM, 2009).

The RMSE and MSE were the most often used cost functions among the studied papers (ABBAS et al., 2018; ASILTÜRK, 2012; BEATRICE et al., 2014; CHEN et al., 2017; JURKOVIC et al., 2018; KARA et al., 2020; MEDDOUR et al., 2018; MIA et al., 2017; MIA; DHAR, 2016; MUTHUKRISHNAN; DAVIM, 2009; SARIC; SIMUNOVIC; SIMUNOVIC, 2013).

Most of the works used *Matlab neural networks toolbox* to implement the ANN models (ABBAS et al., 2018; ASILTÜRK, 2012; BEATRICE et al., 2014; ÇAYDAŞ; EKICI, 2012; GARG; TAI, 2014; KARA et al., 2020; MIA; DHAR, 2016; MUTHUKRISHNAN; DAVIM, 2009). Other software used were the *Statistica* (PONTES et al., 2012), *C++* (JURKOVIC et al., 2018), and *NeuralWare* (SENTHILKUMAAR; SELVARANI; ARUNACHALAM, 2012). Although its popularity and open-source aspect, *Python* was not used (or not mentioned) in any of the reviewed papers.

## 3.1 Bibliometric remarks

From the reviewed recent developments, some remarks and trends were noted:

- The process parameters depth of cut, feed rate, and cutting speed were the most-used features;
- The average height of the profile (Ra) was the most modeled parameter;
- The RMSE and MSE were the most-used cost functions;
- The sigmoidal variations were the most-used activation functions;
- The topology was, in the majority, composed of one hidden layer and less than 20 hidden neurons.

Besides the technical aspects, a lack of transparency was observed, which could hamper reproducibility. For instance, mentioning all hyperparameters values does not occur in most reviewed papers. In this regard, Donoho (2017)—in a dense review paper about data science history and trends—states that the reproducibility of soft computing research developments will be more practiced in the future, and code and data must be universally citable and systematically retrievable. The current dissertation intends to make the process of building the ANN as transparent as possible, with full access to the data, code, and values of the used hyperparameters.

From the read papers, it is clear that surface roughness modeling with ANN is possible, given that several researchers have performed such a task. However, the big picture suggests the existence of two opportunities that the current work will try to seize—and thus contribute to this research line.

The first opportunity is based on the suggestion that there is a search for continuously producing more precise models (mean absolute percentage errors ranging from 1% to 10%) with fewer and fewer training samples. The problem with this tendency is that training machine learning models with very few training samples and reaching performances of minimal errors (MAPE < 2%) suggest an overfitting situation. Shi and Gindy (2007) commented on "obvious overfitting" in artificial neural networks developed with few samples. On the one hand, the (challenging) data-hungry aspect of machine learning modeling is classical common sense (GÉRON, 2021)[7]. On the other hand, it is a sensitive question in a scenario of evolving machine learning techniques and more experienced practitioners—it may be that the models are getting better and less dependent on massive datasets. **Therefore, a strategy of multiplying the data without increasing experiments configures an *opportunity* of making a way out of this ambiguous scenario.**

The second opportunity is based on the fact that the built ANN models from the reviewed literature generally do not give proper attention to the cutting tool wear. Since it is virtually impossible to cut metals without wearing the cutting tool, it is plausible that this phenomenon—which potentially changes the cutting kinematics (HE; ZONG; ZHANG, 2018)—must be counted. In their review of modeling surface roughness methods, He, Zong, and Zhang (2018) comment that influencing factors in time dimension like cutting tool wear are generally neglected while spatial dimension factors are considered—emphasizing that a surface roughness model considering tool wear must be developed in the future. **Therefore, investigating how tool wear affects surface roughness ANN modeling configures the second *opportunity*.**

[7] Artificial neural networks need relatively rich data to learn from, while not needing much theory knowledge of the phenomenon—ANN essentially learns from past data, not from rules (KASABOV, 1996; SHANMUGANATHAN; SAMARASINGHE, 2016).

# 4 Material and Methods

This work's activities can be divided into machining experiments, statistical analysis, and artificial neural network developments (Figure 25). The top of the figure shows the three datasets acquired through three machining procedures. Experiment 1 varied three factors in three levels, reaching a dataset of 324 labeled examples (factors, responses). The Preparation Phase has reached 212 labeled examples. Experiment 2 varied three factors, each at different levels, producing a dataset of 288 labeled examples.

The results of the experimental phase were assessed and statistically analyzed. Subsequently, the data from Exp1 and Exp2 was used in neural network modeling. The Preparation Phase was used to produce damaged and partially damaged tools for Exp2.

The ANN formulations comprised the development of six similar models. They differ in the inputs, the size of training datasets, and the dataset used in training. Subsequently, the models were investigated regarding the influence of training set size and cutting tool wear in the training data and compared to a theoretical equation.

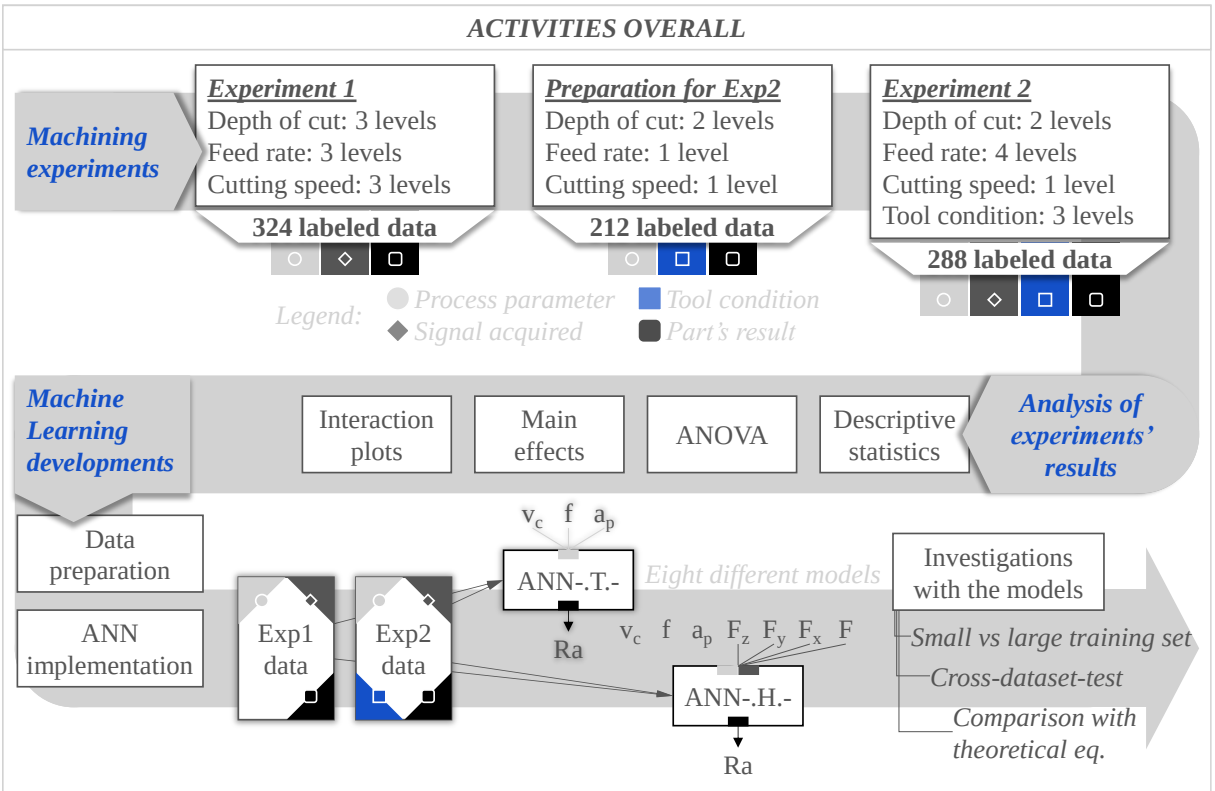



Figure 25. The overall activities. Machining experiments generated data; the results were statistically analyzed; machine learning models were developed and investigated.

## 4.1 Machining experimental procedures

Despite the computation science-oriented content, this is fundamentally a manufacturing work. To be more precise, a turning process investigation. Therefore, machining experiments will be explained in detail to corroborate results that would be achieved later with the computer procedures. Two experiments and one intermediate preparation were performed, as depicted in Figure 25. The machining was conducted in a CNC turning center with cutting fluid. All experiments were performed using the Competence Center in Manufacturing (CCM) infrastructure, a laboratory from the Aeronautics Institute of Technology (ITA).

### 4.1.1 Experiment 1

The first experiment equally explored the influence of three process parameters on the surface finish result of the steel workpiece. The experiment was designed as depicted in Table 8. With cutting speed ($v_c$), feed rate (f), and the depth of cut ($a_p$) varying in three levels in a full factorial design $3^3$, 27 different conditions were tested. Moreover, two replicas were performed, totalizing 54 runs. Factorial experiments are generally the most efficient approach when two or more factors affect a response. This experiment strategy guarantees that every possible combination of factors' levels is tested (MONTGOMERY, 2013). The levels' range for each parameter was chosen according to the tool manufacturer's recommendation to replicate actual production conditions and avoid unknown disturbances possible in out-of-range parameters. After all, a factorial analysis of variance (ANOVA) was performed to understand the effect of the factors on the surface finish and forces results.

Table 8. DoE of Experiment 1.

*Independent variables (factors)*

| *Factors* | *Symbol* | *Units* | *Level* | | | *No. of levels* |
|---|---|---|---|---|---|---|
| Cutting speed | ($v_c$) | m/min | 310 | 350 | 390 | 3 |
| Feed rate | (f) | mm/rev | 0.07 | 0.1 | 0.13 | 3 |
| Depth of cut | ($a_p$) | mm | 0.25 | 0.5 | 0.8 | 3 |

The workpiece was prepared so that each condition (process parameters combination) could be tested individually as a single machining run, avoiding the influence of one run on the others. The part was prepared with grooves separating each machining space related to each

condition. This technique also avoided the measurement in the wrong area because the test parts were physically divided. Figure 26 shows the scheme of the experiment and the acquired signal.

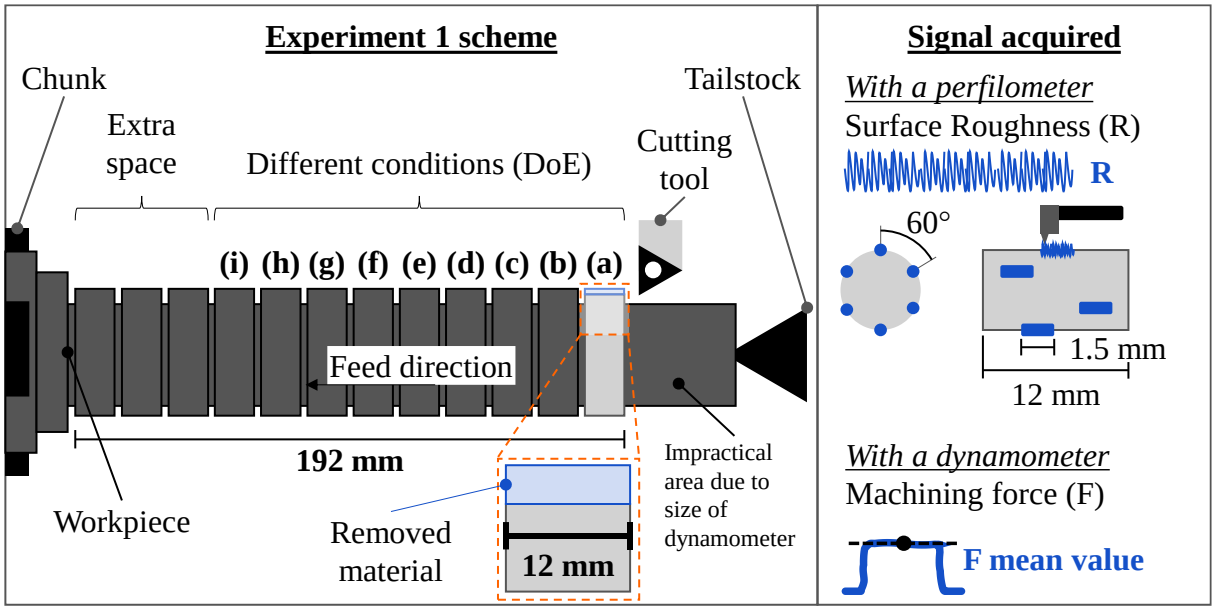



Figure 26. Experiment 1 Scheme (left) and the signal acquired (right).

Each machining space, testing area, or testing length (for example, (a) in Figure 26) was 12 mm long. According to ISO 4288 (1996), for surfaces with 0.04 mm ≤ RSm ≤ 0.13 mm, the measuring test length shall be lt = 1.5 mm. The test length (lt) jumps to 4.8 mm when RSm > 0.13 mm and 15 mm when RSm > 0.4 mm. Using a machining length of 12 mm avoids the problem of insufficient space when the surface roughness parameter RSm is higher than 0.13 mm. Since the theoretical roughness profile has a periodical behavior in a longitudinal turning operation, the mean width of profile elements RSm depends and is almost linearly correlated to the feed rate: in normal conditions, the feed and the RSm are approximately equal (TOMOV; KUZINOVSKI; CICHOSZ, 2016). Therefore, the maximum RSm expected for this experiment was 0.13 mm—the same as the maximum feed rate—which makes the testing length of 12 mm enough because it is already higher than the 4.8 mm length standardized to this value of RSm.

Another reason to use a testing length of 12 mm was that the author was unsure of the effect of the cutting-edge entry impact (BOUZAKIS et al., 2012a, 2012b; REHORN; JIANG; ORBAN, 2005) on the machining force peaks or surface finish. Therefore, a larger testing area was deliberately defined to avoid this influence, whether the phenomenon has a significant effect or not.

The surface roughness was measured in six different places in the testing area along the surface of the workpiece to follow a strategy of augmenting the available data. The measurements were away from each other by 60º, considering the center of the workpiece's face as the origin (Figure 26, right side). The measurement was taken normal to the surface lay as recommended in standards and practical manuals (ISO, 1996; MITUTOYO, 2016) and randomly throughout the 12 mm length to generate more representative results and avoid systematic errors. The possibility of measuring in random spaces is another reason to use a testing length over 4.8 mm.

The grooves were machined with the exact width of the grooving tool of 4 mm. From the whole part length, it was possible to keep 12 testing areas, nine used for the experiment, and the three extras were kept to be used in cases where it could be necessary to repeat some of the runs. The area between the tailstock (or the workpiece's right face), and the first testing area, was impractical due to the size of the dynamometer plate. Therefore, this area had to be machined and could not be used.

The depth of cut was one factor that was to be varied, and it would be a problem if it were randomized because when a forward test area is machined with a smaller diameter than the previous one, this could cause a type of collision on the next series of runs (Figure 27, Case 2). Although the four-millimeter space between each sample area would be enough to approach the cutting tool at every run, avoiding collision risk was preferred. Therefore, the depths of cut were kept constant for each of the nine conditions throughout the workpiece length (Figure 27, Case 1). Keeping the depth of cut constant was made by grouping the experiment design into "9 + 9 + 9" groups.

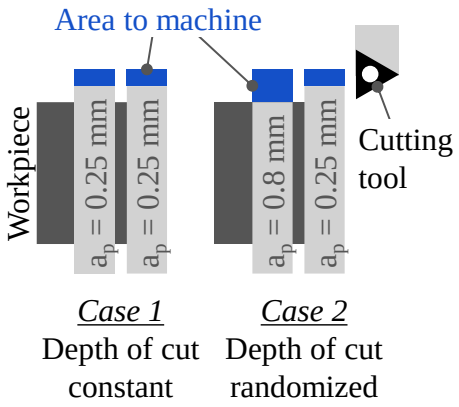



Figure 27. Case 1 (left) depicts the adopted procedure, and Case 2 (right) depicts a collision risk situation.

The experiment was randomized to avoid possible extraneous factors effects and keep observations (or errors) as independently distributed random variables (MONTGOMERY, 2013). Figure 28 shows the actual sequence of conditions used in Experiment 1. The randomization was made inside and between groups. The cutting speed, with three levels, was used to build sub-groups (orange box) inside the depth of cut groups (blue box). Moreover, the feed rate was randomized inside the sub-groups. After all: internally to each sub-group, the effect of the factor feed rate is investigated; between sub-groups, the effect of the cutting speed is investigated; and between groups, the effect of the depth of cut is investigated.

The nine testing areas were machined separately with a stop in-between runs. Starting from group 1, in Figure 28, the following steps were to be performed:

i. Start cutting with the process parameters of the current condition (in first: condition 4, position a);
ii. Measure forces during the cut;
iii. After the cut has ended, measure surface roughness six times, varying measurement position;
iv. Proceed to cut the next testing area until there is no condition left;
v. Change to the following group and repeat steps 1-5 until no group is left.

| | Group 1 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Sub-group 1 | | | Sub-group 3 | | | Sub-group 2 | | |
| f | 0.13 | 0.07 | 0.10 | 0.07 | 0.13 | 0.10 | 0.10 | 0.13 | 0.07 |
| $v_c$ | 310 | 310 | 310 | 390 | 390 | 390 | 350 | 350 | 350 |
| $a_p$ | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 |
| Cond. | 3 | 1 | 2 | 7 | 9 | 8 | 5 | 6 | 4 |


| | Group 1 (2nd replica) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Sub-group 1 | | | Sub-group 3 | | | Sub-group 2 | | |
| f | 0.07 | 0.13 | 0.10 | 0.07 | 0.10 | 0.13 | 0.07 | 0.13 | 0.10 |
| $v_c$ | 310 | 310 | 310 | 390 | 390 | 390 | 350 | 350 | 350 |
| $a_p$ | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 |
| Cond. | 1 | 3 | 2 | 7 | 8 | 9 | 4 | 6 | 5 |

| | Group 2 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Sub-group 5 | | | Sub-group 6 | | | Sub-group 4 | | |
| f | 0.07 | 0.13 | 0.10 | 0.13 | 0.07 | 0.10 | 0.10 | 0.13 | 0.07 |
| $v_c$ | 350 | 350 | 350 | 390 | 390 | 390 | 310 | 310 | 310 |
| $a_p$ | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| Cond. | 13 | 15 | 14 | 18 | 16 | 17 | 11 | 12 | 10 |

| | Group 2 (2nd replica) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Sub-group 6 | | | Sub-group 5 | | | Sub-group 4 | | |
| f | 0.13 | 0.10 | 0.07 | 0.07 | 0.10 | 0.13 | 0.07 | 0.13 | 0.10 |
| $v_c$ | 390 | 390 | 390 | 350 | 350 | 350 | 310 | 310 | 310 |
| $a_p$ | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| Cond. | 18 | 17 | 16 | 13 | 14 | 15 | 10 | 12 | 11 |

| | Group 3 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Sub-group 7 | | | Sub-group 9 | | | Sub-group 8 | | |
| f | 0.10 | 0.07 | 0.13 | 0.13 | 0.10 | 0.07 | 0.10 | 0.13 | 0.07 |
| $v_c$ | 310 | 310 | 310 | 390 | 390 | 390 | 350 | 350 | 350 |
| $a_p$ | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 |
| Cond. | 20 | 19 | 21 | 27 | 26 | 25 | 23 | 24 | 22 |

| | Group 3 (2nd replica) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Sub-group 9 | | | Sub-group 8 | | | Sub-group 7 | | |
| f | 0.10 | 0.13 | 0.07 | 0.10 | 0.13 | 0.07 | 0.13 | 0.07 | 0.10 |
| $v_c$ | 390 | 390 | 390 | 350 | 350 | 350 | 310 | 310 | 310 |
| $a_p$ | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 |
| Cond. | 26 | 27 | 25 | 23 | 24 | 22 | 21 | 19 | 20 |

Figure 28. Randomization of the experiment and groups' division.

### 4.1.2 The Preparation Phase for Experiment 2

The preparation phase for the second experiment was driven by the necessity of having cutting tools with defined wear. Figure 29 shows this phase's scheme (left side)—a cylindrical turning—the information acquired (middle) and its steps (right side). Moreover, during Experiment 2, the tools' wear needed to change the less possible. Therefore, the preparation phase's first step was to conduct a tool life test to roughly define a time region where the tool wear behaves with more stability (phase II in Figure 29). For the tool life characterization, the criterion selected to identify tool life ending was the flank wear VB ≥ 0.3 mm, a standard criterion (ISO, 1993).

After the tool life curve was depicted, the preparation phase continued with the following steps, preparing other tools for Experiment 2. Two tools were prepared to reach the mid-life phase of wear, which corresponds to the beginning of the stable phase defined in the first step (begin of phase II); and two tools were prepared to reach the end-of-life criterion of flank wear ≥ 0.3 mm (end of phase III).

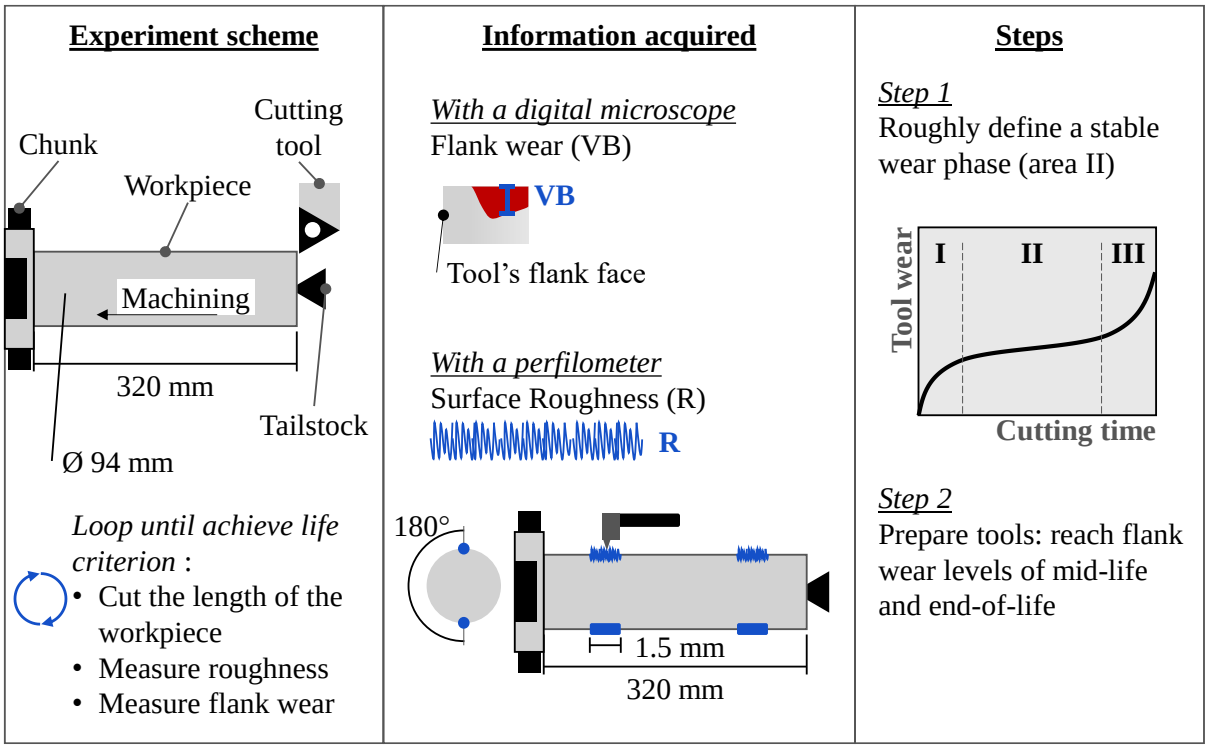



Figure 29. Preparation for Experiment 2 scheme (left), information acquired (center), and preparation steps (right)[8].

[8] The tool life curve with three wear stages—(I) Initial (slight rapid), (II) Intermediary (uniform), and (III) Final (extremely rapid)—is presented by Vaughn (1966) (and others).

During the entire preparation phase, which includes Step 1 and Step 2, the surface values were collected to analyze the influence of the varying tool condition on the surface roughness results. The data was multiplied by measuring the surface roughness in different positions for the same factor combination (the same idea presented in Experiment 1). In this case, the surface roughness was collected in four places of the workpiece after each cut. Two hundred twelve labeled (212) pairs of inputs/responses for each surface roughness parameter were achieved by the end of this phase. With five surface roughness parameters of interest, 1060 pairs of labeled examples were collected in the Preparation Phase.

### 4.1.3 Experiment 2

Experiment 2 was characterized by investigating the influence of the tool condition (measured by the flank wear height VB) on the surface finish results. It had the exact setup configuration of Experiment 1. Therefore, in Figure 26 (left side), it is possible to see the experiment scheme, with the testing areas physically divided by the grooves. The second experiment used fewer testing areas, getting until the position (h) in the workpiece. Regardless of the same machining configuration, Experiment 2 used one more measurement system needed to assess the tool's wear. The information acquired in this experiment is shown in Figure 30.

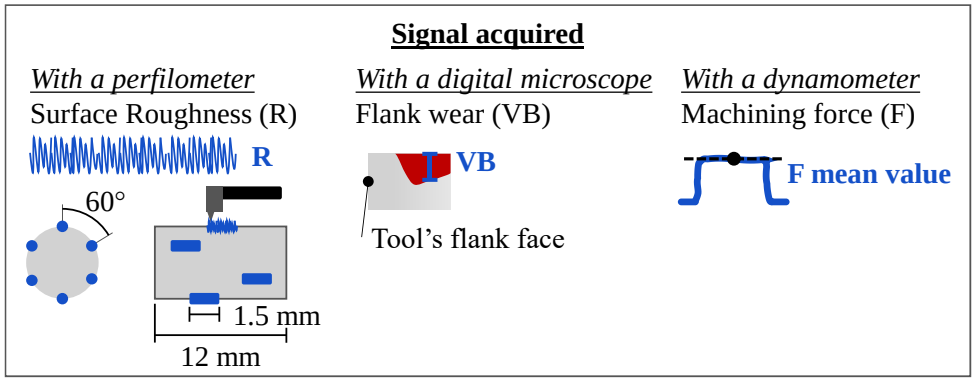



Figure 30. Information acquired at Experiment 2.

This experiment was designed using a pre-analysis of the first experiment results; hence, cutting speed was kept constant due to its apparent low contribution to the surface finish response. The depth of cut was varied only in two levels, for it had a lesser influence on the surface finish than the feed rate. Nevertheless, the depth of cut was varied more than the cutting speed due to its influence on the machining forces and because it was discovered to be more

influential than the cutting speed. The feed rate was varied four times because it was the factor that influenced more on the surface finish. The design of Experiment 2 is depicted in Table 9.

Table 9. DoE of Experiment 2.

*Independent variables (factors)*

| *Factors* | *Symbol* | *Units* | *Level* | | | | *No. of levels* |
|---|---|---|---|---|---|---|---|
| *Cutting factors* | | | | | | | |
| Cutting speed | ($v_c$) | m/min | 350 | | | | 1 |
| Feed rate | (f) | mm/rev | 0.07 | 0.09 | 0.11 | 0.13 | 4 |
| Depth of cut | ($a_p$) | mm | 0.25 | 0.5 | | | 2 |
| *Cutting tool factors* | | | | | | | |
| Tool condition | (VB) | mm | 0 | 0.1 | 0.3 | | |

Later, an ANOVA was performed to check the contribution of each parameter in the surface finish and forces responses. In this case, it was possible to see the influence of the tool condition as a factor.

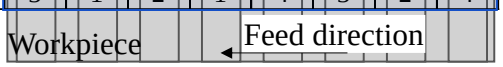
The levels of tool condition (VB) are approximations for simplification. Each cutting tool's actual flank wear size was slightly around these categorical level values. For the record, new tools: tool 44 and 53 had VB = 0.000 mm; mid-life tools: tool 13 and 33 had VB = 0.102 mm and VB = 0.096, respectively; and end-life tools: tool 81 and 23 had VB = 0.300 mm and 0.305 mm, respectively.

The current sequence of conditions tested during Experiment 2 is shown in Figure 31. The runs were randomized in the same manner as in Experiment 1, and the depth of cut was kept constant on each level to avoid a collision, the same reason explained in Figure 27. The replica conditions were grouped and randomized together inside each level.

The tool condition was tested through the levels depicted in Figure 31. Level 1 and Level 2 used a new cutting tool; Level 3 and Level 4 used mid-life tools, with flank wear $0.1 \leq VB \leq 0.15$ mm, a condition of more durable wear-state defined at Item 4.1.2, in the Preparation phase; Level 5 and Level 6 used end-life cutting tools with a flank wear $VB \geq 0.3$ mm, but not higher than 0.35 mm.

New tool

| | Level 1 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| f | 0.11 | 0.07 | 0.09 | 0.07 | 0.13 | 0.11 | 0.09 | 0.13 |
| $a_p$ | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 |
| Cond. | 3 | 1 | 2 | 1 | 4 | 3 | 2 | 4 |




| | Level 2 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| f | 0.07 | 0.13 | 0.09 | 0.07 | 0.11 | 0.09 | 0.11 | 0.13 |
| $a_p$ | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| Cond. | 5 | 8 | 6 | 5 | 7 | 6 | 7 | 8 |

Mid-life

| | Level 3 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| f | 0.09 | 0.07 | 0.09 | 0.13 | 0.13 | 0.07 | 0.11 | 0.11 |
| $a_p$ | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 |
| Cond. | 10 | 9 | 10 | 12 | 12 | 9 | 11 | 11 |

Mid-life

| | Level 4 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| f | 0.11 | 0.09 | 0.07 | 0.11 | 0.13 | 0.13 | 0.07 | 0.09 |
| $a_p$ | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| Cond. | 15 | 14 | 13 | 15 | 16 | 16 | 13 | 14 |

End-life

| | Level 5 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| f | 0.13 | 0.09 | 0.11 | 0.11 | 0.13 | 0.07 | 0.09 | 0.07 |
| $a_p$ | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 | 0.25 |
| Cond. | 20 | 18 | 19 | 19 | 20 | 17 | 18 | 17 |

End-life

| | Level 6 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| f | 0.09 | 0.13 | 0.11 | 0.07 | 0.09 | 0.13 | 0.07 | 0.11 |
| $a_p$ | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| Cond. | 22 | 24 | 23 | 21 | 22 | 24 | 21 | 23 |

Figure 31. The sequence of conditions tested in Experiment 2. Three different tool conditions were tested through the levels.

Figure 32 shows the experimental setup used in Experiment 2. The tool wear assessment (represented in blue) was made possible with a digital microscope that, with a magnetic base, was held in the same position after each run and pointed to the cutting tool's flank. The surface roughness measurement proceeded as in Experiment 1. Figure 32 also represents the setup for Experiment 1, except for the tool wear assessment system.

In Experiment 2, the data was multiplied using the same idea presented in Exp1—measuring the surface roughness in six different positions after each cut. Two hundred eighty-eight labeled (288) pairs of inputs/responses were acquired for each surface roughness parameter. Given that five surface roughness parameters were analyzed, 1440 pairs of labeled examples were collected in the Preparation Phase.

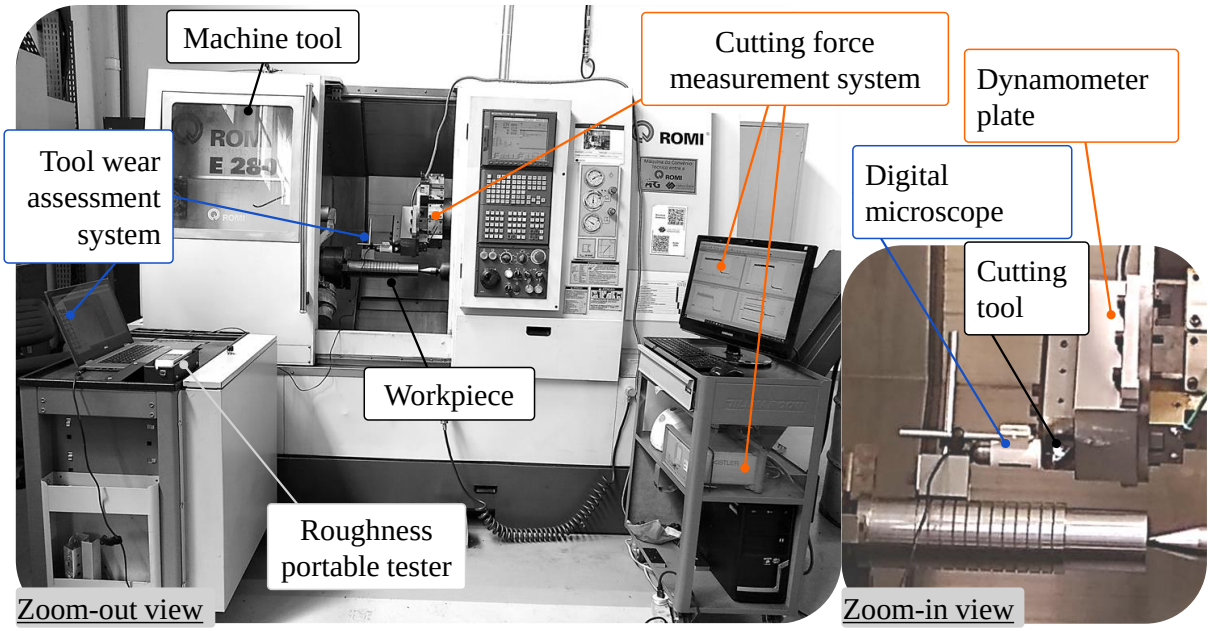



Figure 32. Experiment 2 setup. Zoom-out view (left) shows the whole setup, and the zoom-in (right) shows the digital microscope pointing at the tool to measure flank wear.

### 4.1.4 Workpiece material and cutting tool

This topic lists the material and the cutting tool used in Experiment 1, the Preparation phase, and Experiment 2.

The material used was the AISI H13 steel, fabricated by *Villares Metals*. As provided by the steelmaker, its chemical composition is indicated in Table 10. Its mean hardness is 200HV (Vickers hardness scale), measured transversally to the bar. It is a tooling steel for hot work, usually used in plastic injection molds. It is commonly characterized by good resistance to crack propagation and thermal failure; good mechanical properties; good machinability, good dimensional stability during heat treatment; and good wear resistance (VILLARES METALS, 2006).

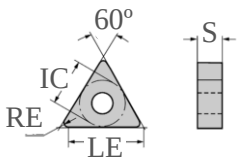
Table 10. AISI H13 steel chemical composition (VILLARES METALS, 2006).

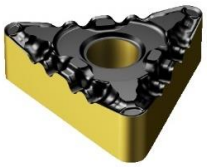
| *Element* | *C* | *Si* | *Mn* | *Cr* | *Mo* | *V* |
|---|---|---|---|---|---|---|
| Percentage (%) | 0.40 | 1.00 | 0.35 | 5.20 | 1.50 | 0.90 |

The cutting tool used is a commercialized model fabricated by *Sandvik Coromant*, with its core substrate of hard metal covered with TiCN, Al2O3, and TiN. This tool is recommended

for finishing applications, and its triangular geometry with two working sides gives it six cutting edges. More specifications of the used tool are shown in Table 11.

Table 11. Cutting tool specifications.



Cutting tool commercial code:

*ISO*
**TNMG 16 04 04-PF 4425**
*ANSI*
**TNMG 331-PF 4425**

*Geometric specifications*

| *Item* | *Symbol* | *Description* |
|---|---|---|
| Diameter of the inner circle | IC | 9.525 mm |
| Edge radius | RE | 0.397 mm |
| Effective cutting length | LE | 16.098 mm |
| Width | S | 4.762 mm |

*Functional specifications*

| *Item* | *Description* |
|---|---|
| Operation | Finishing |
| Size and shape | TN1604 |
| Hand | N |
| Substrate | HC |
| Grade | 4425 |
| Coating | CVD TiCN+Al2O3+TiN |
| Clearance angle major | 0 |

The tool shank (holder) used has code ISO: MTJNL 2020K 16M1 and is also manufactured by *SandivikCoromant*. The grooving tool for preparing the workpiece was a *TaeguTec* TDJ 3 TT9080.

### 4.1.5 Machining forces processing

The force measurement acquisition system was composed of one dynamometer *Kistler* Type 9265B, connected to a charge amplifier *Kistler* Type 5070 and an acquisition software *Kistler Dynoware* Type 2825A, one computer, and peripherical item: a highly insulated cable, Peripherical Component Interconnect (PCI interface), connection cable, and acquisition plate (A/D), as depicted in Figure 33.

For Experiment 1 and Experiment 2, the machining forces were collected during machining. The measurements were posteriorly processed in the *Kistler DynoWare* software Type 2825A, as depicted in Figure 34. Removed the start and end of cutting—to avoid increasing/decreasing forces derived from the acceleration of the cutting tool and tool-workpiece engagement—the mean value can be assessed in the software. The mean value is sufficient to describe the forces actuating in the cutting zone due to the continuous characteristic of the turning process. Moreover, there is no need for smoothing or filtering because the higher

and lower values are reasonably equally distributed around the mean value line. Smoothing and filtering techniques were tested, and the observed mean remained with the same value, indicating that the techniques have removed equally distributed noise around the mainline of the measured forces.

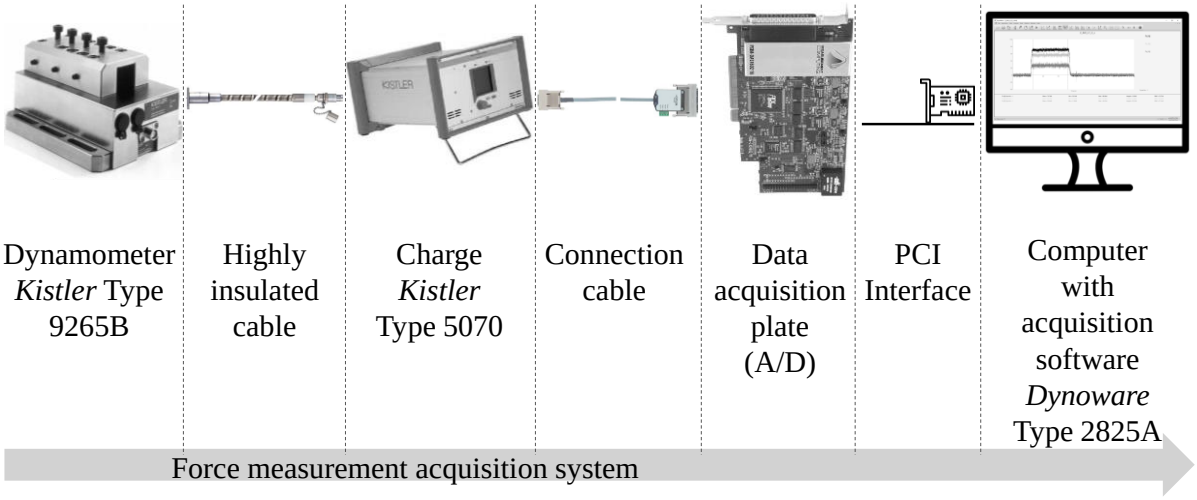



Figure 33. Force measurement acquisition system: connection configuration.

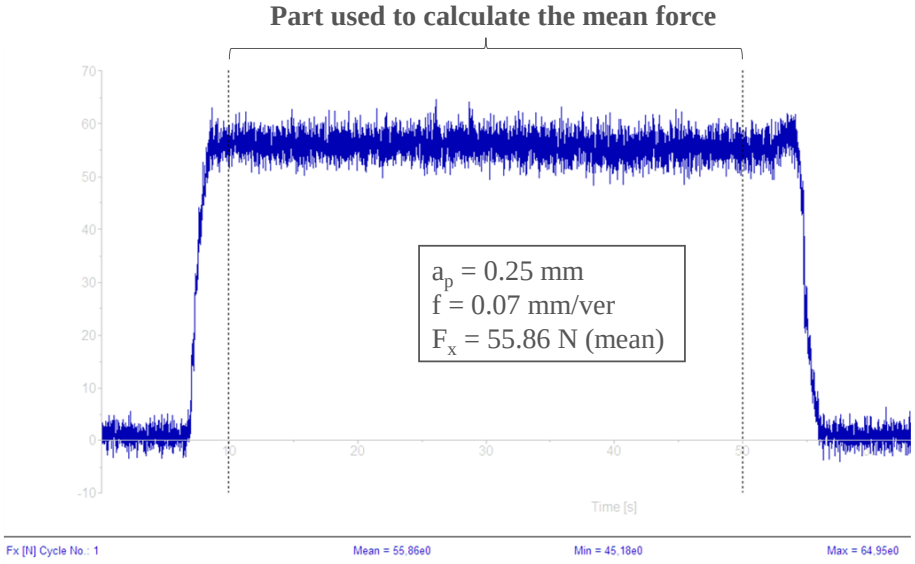



Figure 34. Machining forces visualization and means assessment. A discrete value was selected from the force time-series measurements: the mean.

The dynamometer collects the three components of the resultant force, as depicted in Figure 35: the cutting force ($F_c$) on X-axis, the passive force ($F_p$) on Y-axis, and the feed force ($F_f$) on Z-axis. After their mean was collected at the acquisition/analysis software, the resultant force was calculated using Eq. 6.

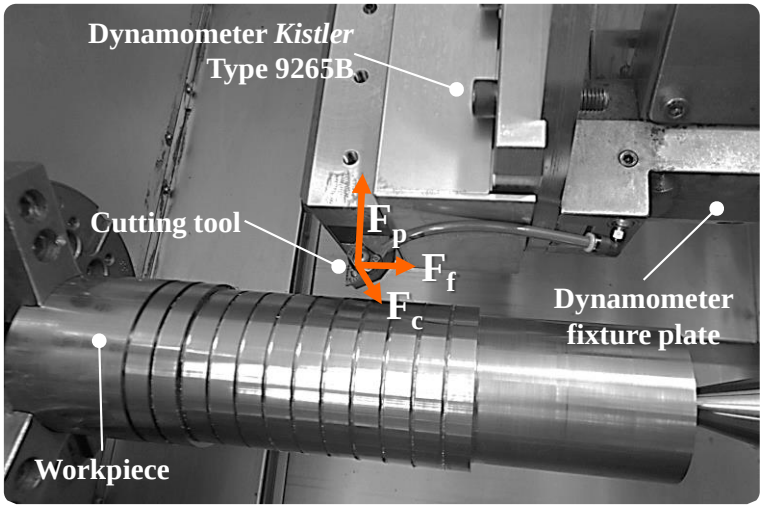



Figure 35. The machining force components measured in experiments 1 and 2.

### 4.1.6 Tool wear assessment

The tool flank wear assessment was possible with a digital microscope *Dino-Lite* model *AM4113ZT* (Figure 36). The microscope was positioned directly on the workpiece using a magnetic base to enable measurement of the tool wear without removing the tool from the machine.

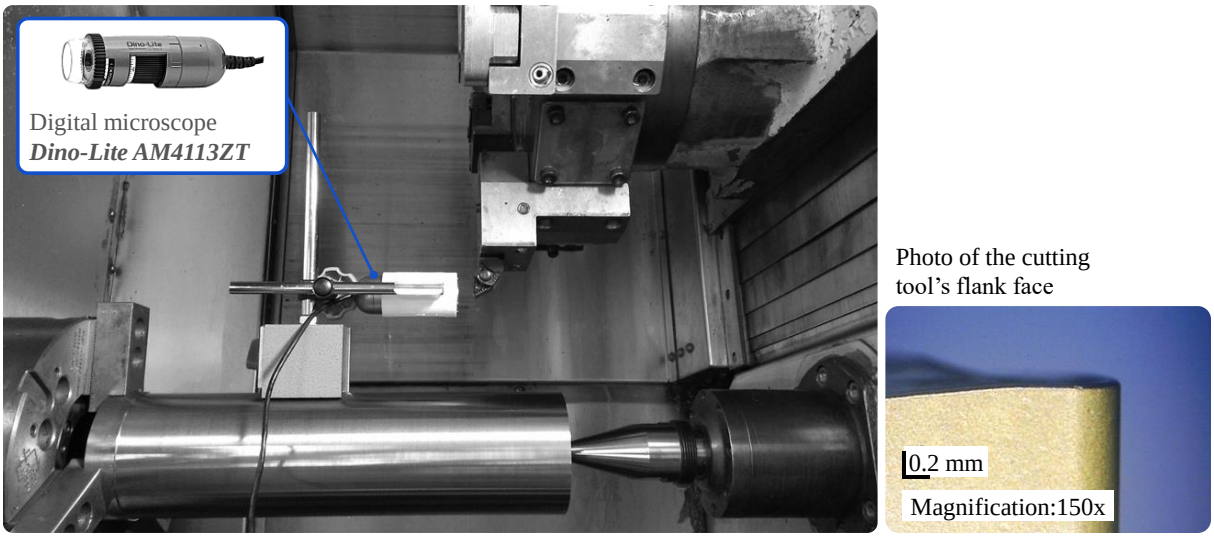



Figure 36. Tool wear assessment with a digital microscope. Measurement setup (left) and result example (right).

### 4.1.7 Surface roughness processing

A *Mitutoyo* portable roughness tester model *Surftest SJ-210*, with a diamond tip of 2 µm radius and 60º angle (Figure 37), was used to measure the surface roughness. The surface roughness was measured perpendicular to the theoretical surface lay—alongside the feed rate direction.


Figure 37. *Mitutoyo* portable roughness tester model Surftest SJ-210. Measurement setup (left) and result example (right).

The measured surface roughness values must be organized in tables or data frames so that the neural network models can read and interpret them. However, the surface roughness measured in Experiment 1, in the preparation phase, and Experiment 2 was first acquired by the portable roughness tester and saved in text format (TXT) in a memory card with several non-numeric characters attached to them. Each measurement corresponded to one document, hence 824 documents. Moreover, five surface roughness parameters were chosen for each measurement, multiplying 824 by 5, which means collecting 4120 values from the text files to data frames.

An automated solution is proposed to avoid the effort and the probability of collecting wrong values, causing undesired influencing outliers. Each surface roughness value was surrounded by the same characters and located in specific lines throughout the text document. With this pattern known, an automated program was developed in *Python* to read and select the values from the TXT files and write them into *Excel* and comma-separated value format files (XLSX and CSV). The workflow from measuring to accessing the individual surface roughness values is shown in Figure 38.

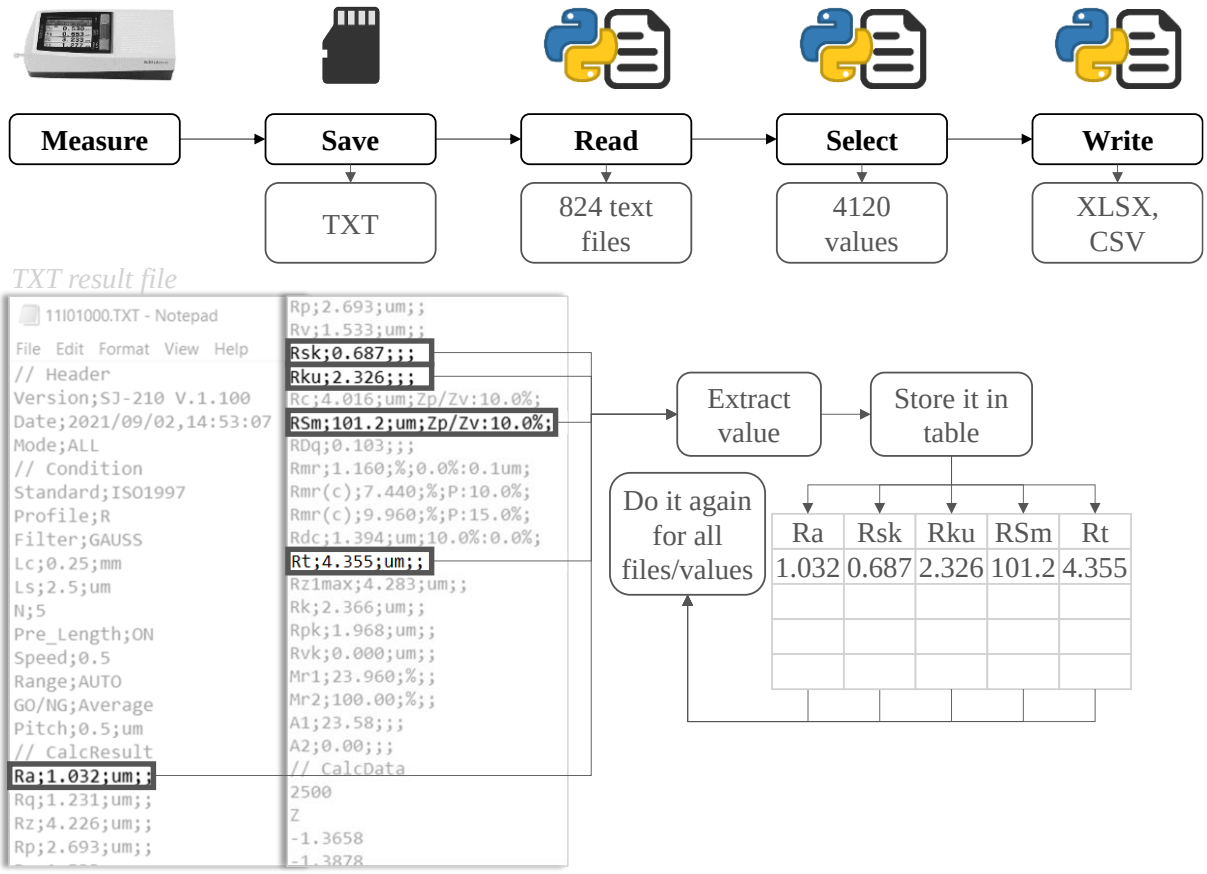



Figure 38. Surface roughness data processing-to-machine-learning workflow.

### 4.1.8 Strategy to increase available data

A strategy was used in Exp1 and Exp2 to multiply the available data without increasing the number of experiments. The procedure was already commented on throughout each experiment topic. However, it is highlighted here to clarify: the strategy idea was to measure the surface roughness (the primary dependent variable) six times for each tested combination—so that the experimentation was not increased.

## 4.2 Soft computing procedures

Different artificial neural networks have been implemented in *Python 3.9* during the machine learning development phase, as depicted in Figure 25. Their identification nomenclature follows the pattern shown in Figure 39. The first number after the ANN acronym is a reference to the machining experiment that generated the data; the subsequent letter is a reference to the number of inputs—T is for three inputs (following the Greek prefix *tri-*), and

H is for seven inputs (following the Greek prefix *hepta*-)—and the last letter is a reference to the dataset size—S for small and L for large. The procedures narrated in this chapter happen similarly to all of those models, and the Results section will discuss the outcomes in pairs for models trained with Exp1 and Exp2.

Sometimes, hyphens are used instead of letters/numbers for generalization. For instance, the name "ANN1.-.-" represents all models trained with the dataset generated in Exp1, regardless of the inputs' number or dataset size. Another procedure is to name models "ANN1.-.-" as "models 1". The same can be made for the "T/H-models" and "S/L-models."

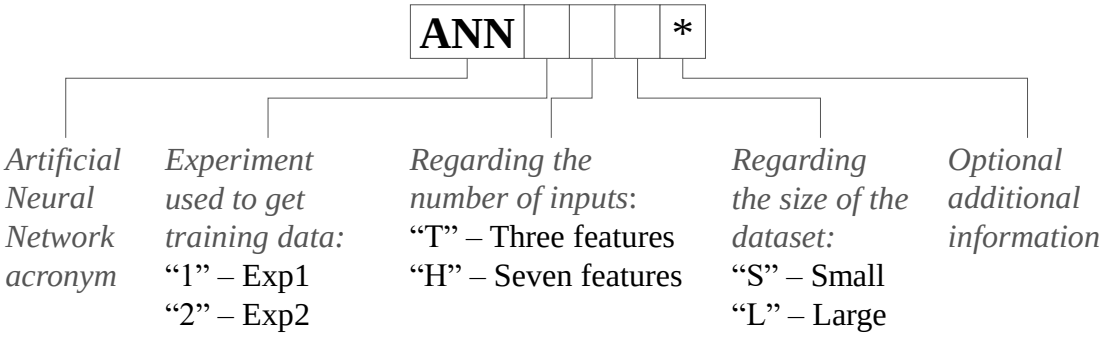



Figure 39. Machine learning models nomenclature.

Most machine learning models, especially artificial neural networks, do not work when the data sets have problems with non-numeric values, missing values, and features with too different ranges. Therefore, some preparation steps are needed to introduce data into a machine learning model. This pre-model phase is represented by the two left gray boards in Figure 40. Figure 40 depicts the general steps taken in this work for the ANN application. Previous preparation takes place in the measurements' software to extract the responses. Both surface roughness' and forces extraction were explained in previous chapters. This phase uses secondary *Python* coding and *Excel* sheets to organize and clean the data, suppressing non-numeric values and outliers.

The middle board steps are inside the main *Python* code where the ANN application is developed. After loading the data, small preparation steps are made, such as dropping unnecessary values (experiment identification info and other non-numeric data). Inside the main *Python* program, essential preparation steps are:

- To split *train* and *test* values into two data sets;
- Define which *labeled examples* will be used;

- Split *train* set and *validation* set;
- Scale the *features*.

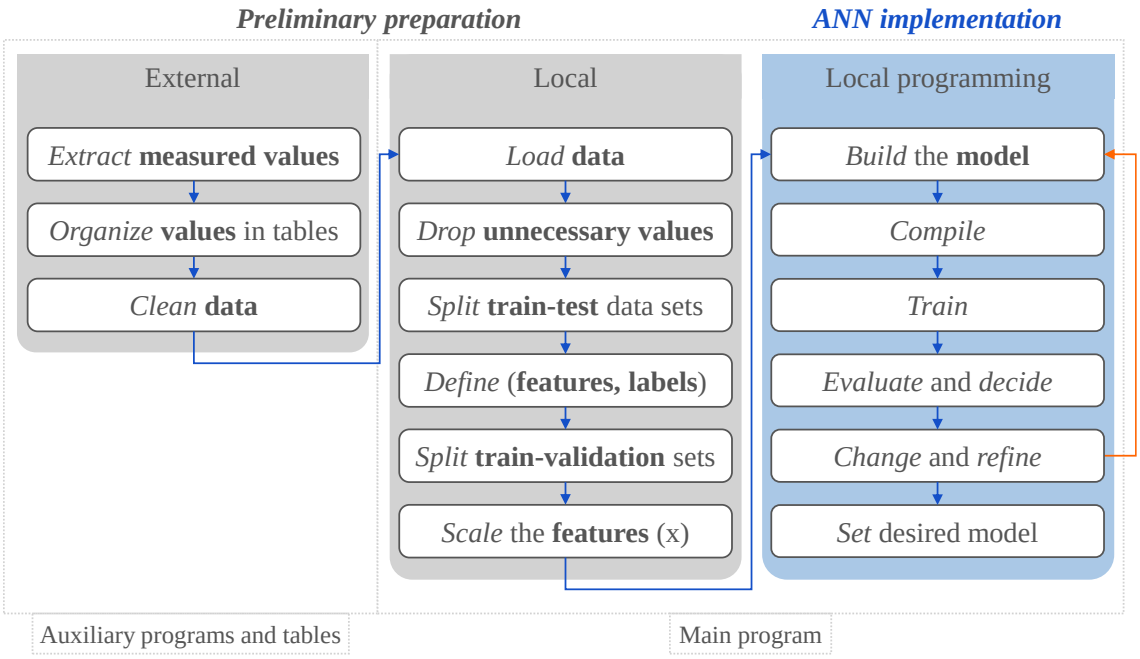



Figure 40. Steps from preparing the data to defining the model.

The blue board on the right side of the figure shows the steps to develop the ANN application. It begins by building the model, compiling, training, evaluating its result, and deciding if there is the need to enhance or if it is sufficient. The main changes to optimize the model were made in the build phase by changing its hyperparameters. Other minor changes were made to optimize, such as changing the *seed* of random splitting train-test-validation data sets and manually organizing these sets. However, the orange arrow indicates the primary source of optimization-driven activities.

### 4.2.1 Preliminary preparation

The two gray boxes from Figure 40 comprise the preliminary preparation steps. The prime step was to organize the measured responses (a.k.a. dependent variables, output, or labels) and the factors (a.k.a. independent variables, inputs, or features) inside the same tables, ready to be read by the program. This phase involved extracting surface roughness and forces, as shown previously in 4.1.5 and 4.1.7, and generating sheets in *Comma Separated Values*

(CSV) format with *Python*. These sheets contained 1. the experiments' informative values, 2. the experimental factors, and 3. the responses.

Table 12 shows the headline of experiments' informative values: experiment identification, replica, the cutting tool used, group and subgroup (see Figure 28), position in the workpiece length (see Figure 26), condition, tool condition (that in Experiment 1 was always zero because all tools were wedge-new (VB = 0.0 mm)), test length, initial and machined diameters, cutting time (non-applicable values for Experiment 1), surface roughness measurement angle indicating the angular position in which the surface roughness was accessed on the workpiece (see Figure 26), and lastly the run identification (that comprises the information of the other columns).

The information values, shown in Table 12, keep data traceability and are necessary for previous data organization and when a value must be accessed. However, they are not used throughout the ANN application.

Table 12. Experiment 1 general information.

| Exp | Replica | Tool | Group | Subgroup | Position | Condition | Tool Condition | Length | Initial diameter | Machined diameter | Cutting time | R measur angle | Run ID |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 021 | 1 | 2 | a | 4 | 0 | 12 | 94 | 93.5 | na | 0 | 1_021_G1_4_a |
| 1 | 1 | 021 | 1 | 2 | a | 4 | 0 | 12 | 94 | 93.5 | na | 60 | 1_021_G1_4_a |
| 1 | 1 | 021 | 1 | 2 | a | 4 | 0 | 12 | 94 | 93.5 | na | 120 | 1_021_G1_4_a |
| 1 | 1 | 021 | 1 | 2 | a | 4 | 0 | 12 | 94 | 93.5 | na | 180 | 1_021_G1_4_a |
| 1 | 1 | 021 | 1 | 2 | a | 4 | 0 | 12 | 94 | 93.5 | na | 240 | 1_021_G1_4_a |
| 1 | 1 | 021 | 1 | 2 | a | 4 | 0 | 12 | 94 | 93.5 | na | 300 | 1_021_G1_4_a |

Subsequently, non-numeric values, missing values, and outliers were searched between the *factors* ($a_p$, $v_c$, f, Tool Condition) and *responses* (Ra, Rsk, Rku, RSm, Rt, $F_x$, $F_y$, $F_z$, F) throughout the data frames. Among 4320 values in Experiment 1, 1908 in the Preparation Phase, and 3888 in Experiment 2, no non-numeric value and only two missing values were discovered (probably due to not saving the measurement in the surface roughness-tester during the experiments). These two values were filled with the mean results of their correspondent experiment condition replica. The analysis of outliers is discussed in the results chapter.

Local preparation, which manipulates data inside the main *Python* program, begins by loading the data as a data frame, then dropping all information values (those depicted in Table 12). Subsequently, it splits the data into training (80% of the samples) and testing (20%) using a stratified strategy that guarantees a balanced division based on one chosen parameter. The feed rate is used because it is the prime influencer on surface roughness (as presented in the

results chapter and discussed in topic 2.8.2). The code lines for this step are shown below. It uses a *scikit-learn* function, which is called *Stratified Shuffle Split*.

```
# ############## TRAIN-TEST SPLIT
# Split train and test data sets (ratio = 80/20)
split = StratifiedShuffleSplit(n_splits=1,
                               test_size=0.2,
                               random_state=42)

for train_index, test_index in split.split(r_df, r_df["f"]):
    stratified_train_set = r_df.loc[train_index]
    stratified_test_set = r_df.loc[test_index]
```

The *Stratified Shuffle Split* uses the argument "*random_state*" to define the *seed* of division. In this case, it was set to an arbitrary value, 42 (this could be any chosen value). With this argument set to the same arbitrary number, the division is always equal: the same values go to the training and testing sets. Keeping the same split sets is helpful because the building phase demands the model to be trained many times for hyper tuning, and a variation in training and test sets would mask the results between the two models.

Subsequently, the *features* (input values, X) and *labels* (output values, y) are defined. Those pairs of features and labels will form the labeled examples given to the neural network in the training phase. The possible features can be:

- Process parameters: depth of cut ($a_p$), feed rate (f), cutting speed ($v_c$), and the artificial feature of the undeformed chip area (A);
- Measured forces: cutting force ($F_x$), passive force ($F_y$), feed force ($F_z$), and the resultant force (F). The machining forces are treated as response variables in the experimental procedure and analysis of results. However, they are used as input values in the machine learning model.

The step of defining the labeled examples uses a simple piece of code. It only drops non-required *labels* keeping the ones of interest in the two sets called *X* and *y*. The possible *labels* are:

- The surface roughness results: Arithmetic mean deviation (Ra), Skewness (Rsk), Kurtosis (Rku), Mean width of profile elements (RSm), and Total height (Rt).

After defining the labels, the code splits the validation and training set so that the model can test generalization during the training phase. This step uses another function of the s*cikit-learn* package, the *train_test_split*. Again, an arbitrary "*random_state*" value is used. The code below shows that four variables (*X_train*, *X_val*, *y_train*, and *y_val*) are defined by splitting two initial variables (*X_train_and_validation* and *y_train*). Each initial variable is split by 80% for training and 20% for validation[9].

```
# ############## SPLIT VALIDATION SETS
# Split train and validation data sets (ratio = 80/20)
X_train, X_val,\
y_train, y_val = train_test_split(
    X_train_and_validation,
    y_train,
    test_size=0.2,
    random_state=42)
```

Subsequently, all *features* (inputs) must be scaled. Scaling the inputs usually leads to a faster convergence (LECUN et al., 2012). Initially, a *scikit-learn* function is used, the *StandardScaler*, that transforms the values into a similar range. This procedure is called standardization. Below, the reader can see the implemented code that standardizes the training values. It first *fits* the function to the given training data, using the mean and standard deviation, then *transforms* the values using the fitted scaler. Subsequently, the scaler fitted to the training data *transforms* the validation and test data sets. The scaler function and parameters are saved in a *pickle* to be available if needed in another dataset.

```
############## SCALE X
# Define the scale method and range
sc = preprocessing.StandardScaler(copy=True,
                                  with_std=True,
                                  with_mean=True)

# Fit scale to the train set
X_train_scale_np_array = sc.fit_transform(X_train)
```

[9] The split of the dataset into training, validation, and testing follows what is presented in Figure 22.

```
# Transform scale to the validation set
X_valid_scale_np_array = sc.transform(X_val)

# Transform scale to the test set
X_test_scale_np_array = sc.transform(X_test)

# Save the scaler
pickle.dump(sc, open('standardize_ANN1.T.L.pkl', 'wb'))
```

### 4.2.2 Artificial neural network implementation

After all preparation steps, the phase of machine learning modeling is reached. The models' hyperparameters were defined by an exploratory approach summed to occasional trial-and-error tests, similar to what is made by many other works (BEATRICE et al., 2014; CHEN et al., 2017; GARG; TAI, 2014; MEDDOUR et al., 2018; MIA et al., 2017).

The hyperparameter tuning process is somewhat complex. In addition, experimentation can quickly become enormous and unfeasible. This work considered seven hyperparameters, and if an experiment with a full factorial design were to be made with their chosen levels (Table 13), 1,200,000 tested combinations would be necessary.

Therefore, a series of hyper tuning tests were performed to improve this work's reproducibility and avoid those enormous experiments. The tests varied a hyperparameter within its levels while keeping the other hyperparameters fixed. One drawback of this solution is that the interactions between hyperparameters are neglected. The seven hyperparameters considered and their levels are depicted in Table 13. Trial-and-error testing existed because of the ease of training a network model after correctly performing the preparation phase.

The tests consisted of training three models for each hyperparameter level available. In the end, the options for each hyperparameter are ranked using the mean results from the three replicas (three trained models), using the mean squared error (MSE) and mean absolute percentage error (MAPE) (calculated in the validation set) as metrics.

The tests were performed in a planned sequence shown in Table 14. The activation function was the first tested hyperparameter. For the activation function tests, the batch size was 50, the optimizer was the SGD, with a learning rate of 0.1 and a momentum of 0.5, the topology had one hidden layer with eight units, and the model was trained for 150 epochs.

When the hyperparameters were selected, they were used in the subsequent tests. For example, in the second test (batch size), the Relu was already adopted as the activation function, as it was the best-ranked in the 1st test.

Table 13. Hyperparameters and the levels tested on the hyper tuning test series. Each hyperparameter was tested through its levels keeping the other hyperparameters fixed.

| *Hyperparameter* | *Levels* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Activation Function | relu | sigmoid | tanh | selu | elu | | | |
| Batch size | 1 | 5 | 10 | 15 | 20 | 50 | 100 | 150 |
| Optimizer algorithm | SGD | RMSprop | Adam | Adagrad | Nadam | | | |
| Optimizer learning rate | 0.1 | 0.01 | 0.001 | 0.0001 | | | | |
| Optmizer momentum | 0.5 | 0.7 | 0.9 | 0.95 | | | | |
| Hidden layers | 1 | 2 | 3 | | | | | |
| Units in the 1st hidden layer | 8 | 16 | 32 | 64 | 128 | | | |
| Units in the 2nd hidden layer | 8 | 16 | 32 | 64 | 128 | | | |
| Units in the 3rd hidden layer | 8 | 16 | 32 | 64 | 128 | | | |

The increase in the epochs' number to 500 at the 3rd test was made to allow that models trained with a slower optimizer could reach the slope of minimum loss (meaning that the model learning has converged). Again, 500 epochs were used to allow slower learning rates (4th test) and slower momentums (5th test) to reach the learning convergence and were adopted to the subsequent tests.

The topology selection used first a broader range of units number in each hidden layer (7th test) and then a smaller range (8th test) to restrict the topology choice. Unavoidably, before the phase of topology selection, non-documented trial-and-error exploration tests were performed with different topologies using even broader ranges of neurons' quantity, but the primary topology configurations were represented in the hyper tuning documented tests.

Moreover, the topology selection phase used seven neurons as inputs (including the cutting forces to the inputs) due to the perceived characteristic that the difference between the topologies was more significant when using more inputs.

Table 14. Hyper tuning tests sequence and tests configuration.

| *Test sequence* | *Hyperparameter* | *Activation Function* | *Batch size* | *Optimizer* | *Learning rate* | *Momentum* | *Topology* | *Epochs* | *Replicas* |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Activation Function | Varied | 50 | SGD | 0.1 | 0.5 | 3-8-1 | 150 | 3 |
| 2 | Batch size | Relu | Varied | SGD | 0.1 | 0.5 | 3-8-1 | 150 | 3 |
| 3 | Optimizer | Relu | 20 | Varied | 0.01 | 0.5 | 3-8-1 | 500 | 3 |
| 4 | Learning rate | Relu | 20 | SGD | Varied | 0.5 | 3-8-1 | 500 | 3 |
| 5 | Momentum | Relu | 20 | SGD | 0.01 | Varied | 3-8-1 | 500 | 3 |
| 6 | Hidden layers | Relu | 20 | SGD | 0.01 | 0.95 | Varied:<br>3-8-1<br>3-8-8-1<br>3-8-8-8-1 | 500 | 3 |
| 7 | Topology (broad) | Relu | 20 | SGD | 0.01 | 0.95 | Varied:<br>7-16-32-8-1<br>7-32-64-16-1<br>7-64-128-32-1 | 500 | 3 |
| 8 | Topology (restrict) | Relu | 20 | SGD | 0.01 | 0.95 | Varied:<br>7-30-60-15-1<br>7-32-64-16-1<br>7-28-56-14-1 | 500 | 3 |

The subsequent paragraphs show the modeling procedure with the hyperparameters chosen after the previously presented hyper tuning phase. Therefore, with the already selected-as-best models with their hyperparameters.

The modeling phase consists of the steps depicted in Figure 40 (blue board on the right side), starting with the model construction. This step does not require algorithmic formulation, thanks to the Keras Sequential API. The *Keras* API provides built-in ANN mathematical implementations, such as the *Feed-Forward Back-Propagation* algorithm used in this work's models. The lines of code relative to model-building are depicted below. A function named "*create_model*" is used comprising the properties of neural topology (shape of the inputs, number of hidden layers and their units, and the size of the output layer), the optimizer properties, and the activation function (*Relu* was used to all models in all hidden layers).

The compile phase sets the loss function: the mean squared error (MSE), which is decreased by the weights' updates by the learning algorithm (SGD with back-prop); The metrics: the *mean absolute percentage* error (MAPE) and the mean absolute error (MAE), which are used to evaluate the model on the validation and test sets; and the optimizer, which is the algorithm used to decrease the loss function—the *Stochastic Gradient Descent* (SGD) was the optimizer used, with a learning rate of 0.01 and momentum of 0.95.

```
# ############## BUILD MODEL
# Clean session
keras.backend.clear_session()

# Define standard architecture features
RegularizedDense = partial(keras.layers.Dense,
                           activation="relu")


# Build model
def create_model():
    model = keras.Sequential()
    model.add(keras.Input(shape=X_train_sc.shape[1:]))
    model.add(RegularizedDense(28))
    model.add(RegularizedDense(56))
    model.add(RegularizedDense(14))
    model.add(layers.Dense(1))

    optimizer = keras.optimizers.SGD(0.01, momentum=0.95)
```

```
    # ############## COMPILE MODEL
    model.compile(loss='mean_absolute_error',
                  optimizer=optimizer,
                  metrics=['mean_squared_error',
                           'mean_absolute_percentage_error'])

    return model
```

Figure 41 shows the *feedforward neural network* of three dense hidden layers built under the code above. This image was made using the helpful tool available on the Internet by Alexander LeNail (2019a, 2019b). The first layer—the input layer—has the number of units equal to the number of inputs, which for models ANN-.H.- is seven. The first, second, and third hidden layers have 28, 56, and 14 units.

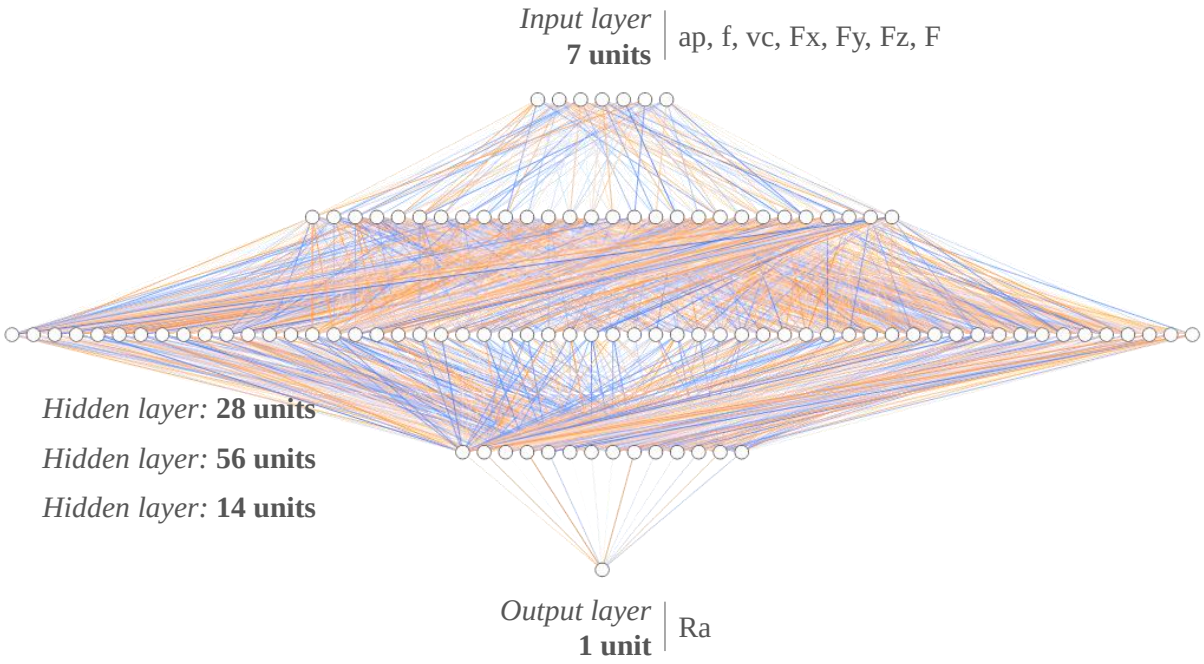



Figure 41. ANN representation. The line colors and opacity degree are only examples and do not represent actual weights' magnitude.

If the model is kept in the learning process after it has learned enough, it can start overfitting the data, harming its generalization capacity. The strategy of stopping the learning process in the middle with a *Keras EarlyStopping* function was used to avoid overfitting and waste of computational effort. After using a checkpoint, this problem was prevented. Another

helpful strategy is to save the model parameters while learning so that the best parameters can be restored after the learning process has ended. This procedure is done with the *Keras ModelCheckpoint*, which saves the weights and biases (in an HDF5 format file) when the minimum mean absolute percentage error (MAPE) occurs. The code below shows how those two strategies have been implemented.

```
checkpoint_path = "ANN1.H.L.hdf5"
cp = ModelCheckpoint(checkpoint_path,
                     monitor='mean_absolute_
                     percentage_error',
                     verbose=1,
                     save_best_only=True,
                     save_weights_only=True,
                     mode='min')

es = EarlyStopping(monitor='loss',
                   min_delta=0.0001,
                   patience=500,
                   verbose=1,
                   mode='min')

callbacks = [cp, es]
```

The next step is to make the model learn (or train the model). The epochs (the number of iterations) are chosen in this step. Despite the epochs being defined as 10,000, the *EarlyStopping* strategy described above will define the correct number of iterations[10]. Code lines below show this step being implemented. One can see that the batch size was chosen as 20.

```
model = create_model()
model.summary()

EPOCH = 10000

history = model.fit(X_train_sc, y_train,
                    epochs=EPOCH,
                    validation_data=(X_val_sc, y_val),
```

[10] At the results section one will see that the longer trainings lasted around 3,000 epochs before the early stopping.

```
                    batch_size=20,
                    callbacks=callbacks,
                    verbose=1)
```

Before the best-observed model is tested in the test set, the model is evaluated by three methods:

- Visually with the loss and metrics curves, searching for overfitting and unrepresentative dataset behavior;
- Computing the errors;
- Visually observing the fitness of plotted predicted values versus actual values (with validation sets).

### 4.2.3 Investigations with the developed models

After developing and selecting the best artificial neural network models, some investigations were performed on them. The effect of dataset size was tested by training the models again using only 1/6 of the available samples.

The models' generalization capacity was tested in a cross-dataset scheme, meaning that the model trained with the train set from Experiment 1 predicted the values from the test set of Experiment 2 and vice versa. This cross experiment intends to observe if the model trained with data produced without tool wear can predict surface roughness values in a scenario where tool wear exists. These tests consisted of the following:

- The best model trained in the train set of Experiment 1—where tool wear does not exist—the model ANN1.H.L predicts the surface roughness values of the test set of Experiment 2 in three replicas;
- The best model trained in the train set of Experiment 2—where tool wear exists—the model ANN2.H.L predicts the surface roughness values of the test set of Experiment 1 in three replicas.

The other investigation made with the models is the comparison of the predictions made with them to their test sets and the predictions made with one of the theoretical equations listed in 2.8.3 in the related test sets.

# 5 Results and Discussion

There are two result categories in this work: the results of machining exploratory experiments—which will focus on statistical analyses of factor-response correlations—and machine learning experimentation—which will focus on performance comparisons and improvement strategies.

## 5.1 Results of machining experiments

The machining investigations were previously divided into Experiment 1, the Preparation Phase for Experiment 2, and Experiment 2. This chapter shows the results observed for these three machining developments. All statistical analyses were implemented in R 4.1.1 and Minitab 21.1. More precisely, the factorial ANOVA, correlations, and related operations were performed in R, and main-effects plots, interaction plots, and descriptive statistics were performed in Minitab[11].

### 5.1.1 Experiment 1

As expected from previous literature research (DAVIM, 2010) (also seen in Figure 14), the feed positively affected the surface roughness mean-parameter Ra, as depicted in Figure 42 (left side). From this visual perspective, the depth of cut and the cutting speed had no pronounced effect. Moreover, the feed presented the same positive effect for the other observed surface roughness parameters, Rt, Rsk, Rku, and RSm. The isolated effect of the parameters on the resultant force is also depicted in Figure 42 (right side). The feed rate and depth of cut had both positive effects, and the cutting speed had no substantial effect. The same effects were observed in the resultant force components: passive force, feed force, and cutting force. The cutting speed having no substantial effect on the surface roughness may be related to the narrow range of levels chosen for this factor.

[11] In graphical results generated at Minitab, the factors' symbols can be slightyly different, without subscripted letters. Please consider ap ≡ (equivalent to) $a_p$, vc ≡ $v_c$, Fx ≡ $F_x$, Fy ≡ $F_y$, Fz ≡ $F_z$.

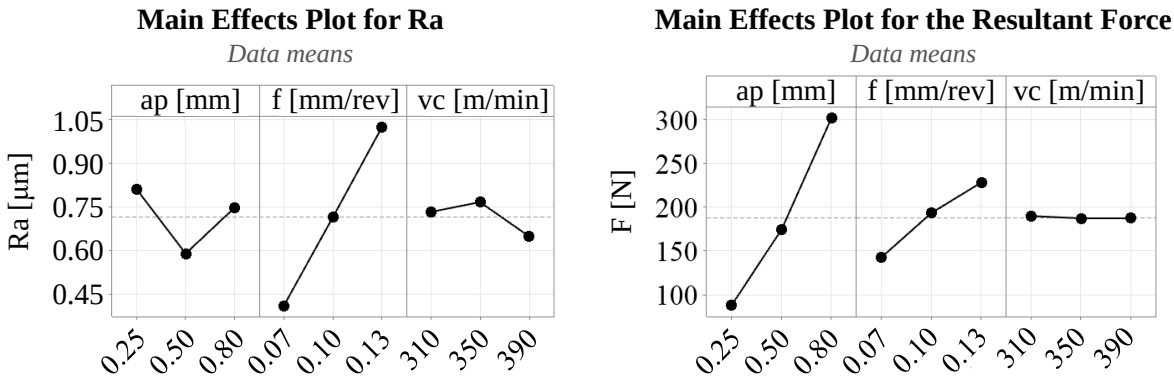



Figure 42. For Experiment 1: Isolated effects on the surface roughness average-parameter Ra (left side) and the resultant force F (right side).

Another helpful step to visualize the data is depicted in Table 15. From left to right, one can see the variables measured in Experiment 1 and their description regarding sample size, mean, standard deviation, and minimum and maximum values. The G corresponds to Grubb's test statistic for outlier search, and the P refers to the p-value of the test, meaning that for a given variable, if the p-value $\leq 0.05$, the null hypothesis of which all the data comes from the same normal distribution is rejected (what will suggest the presence of an outlier). Beforehand, outliers were identified in some of the 27 tested conditions in Ra (two outliers), Rsk (three), Rku (three), RSm (six), and Rt (one). The outliers were then removed by changing each value by the mean of the other five values of its respective condition. Later, looking at the entire set (324 samples), no outlier was identified with a 5% significance level, as depicted in Table 15.

Table 15. Descriptive statistics of Experiment 1 measured variables, including outliers search.

Experiment 1: information of measured variables

| Variable | N | Mean | StDev | Min | Max | G | P |
|---|---|---|---|---|---|---|---|
| Ra | 324 | 0.72 | 0.33 | 0.21 | 1.81 | 3.32 | 0.27 |
| Rz | 324 | 3.04 | 1.26 | 1.07 | 6.55 | 2.78 | 1.00 |
| Rsk | 324 | 0.39 | 0.33 | -0.48 | 1.10 | 2.61 | 1.00 |
| Rku | 324 | 2.18 | 0.37 | 1.49 | 3.20 | 2.71 | 1.00 |
| RSm | 324 | 98.96 | 24.86 | 47.50 | 138.10 | 2.07 | 1.00 |
| Rt | 324 | 3.23 | 1.27 | 1.21 | 6.70 | 2.73 | 1.00 |
| $F_x$ | 324 | 145.11 | 72.08 | 49.23 | 289.60 | 2.00 | 1.00 |
| $F_y$ | 324 | 66.94 | 16.41 | 41.60 | 100.70 | 2.06 | 1.00 |
| $F_z$ | 324 | 92.40 | 72.71 | 18.87 | 227.50 | 1.86 | 1.00 |
| F | 324 | 187.95 | 97.45 | 69.34 | 377.57 | 1.95 | 1.00 |

A factorial ANOVA (Table 16) was conducted to determine the effects of the depth of cut (ap), feed rate (f), and cutting speed (vc) on the surface roughness average-parameter Ra. The ANOVA is a statistical tool commonly used to check the correlation between factors and responses in metal cutting factorial experiments (MIA et al., 2017; MUTHUKRISHNAN; DAVIM, 2009; SENTHILKUMAAR; SELVARANI; ARUNACHALAM, 2012). The best-fitted model was selected using the Akaike Information Criterion test (AIC) (Table 17). The AIC test was conducted between three options, 1: no interaction between independent variables, 2: interaction between the feed rate and the depth of cut, and 3: interaction between all independent variables. The third option presented the most relevant result (verified by the lowest AIC value: -262.08), indicating that the three independent variables and the interactions between them can explain up to 100% of the total variation in the response (Ra) that can be explained by the complete set of models (as shown by the column named AICcWt, which is the AIC weight value). The ANOVA residuals verification is presented in Appendix B – Additional statistical results. After the diagnosis, it is possible to assume that the conclusions drawn from the model presented in Table 16 are reliable.

From the factorial analysis of variance, it is possible to identify the factors with more contribution to the response. The contribution is calculated as the ratio of the individual sum of squares by the total sum of squares. The feed rate was the independent variable with more influence on the surface roughness Ra (57.5%), followed by the depth of cut (8.1%) and the interaction between the depth of cut and cutting speed (5.1%). Even though all variables and interactions were significant, the feed rate is the one with more contribution to the response in terms of the surface roughness average-parameter Ra. Moreover, the cutting speed alone is the variable with less contribution to Ra (2.3%).

In terms of which factor contributes more to the surface roughness formation, a similar result was found by other researchers in which the most influential factor is the feed rate, followed by the depth of cut and cutting speed (ASILTÜRK, 2012; CHEN et al., 2017; MUTHUKRISHNAN; DAVIM, 2009).

Table 16. Factorial ANOVA for the effect of the parameters on Ra in Experiment 1.

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| $a_p$ | 2 | 2.850 | 1.425 | 60.574 | <2E-16 | *** |
| f | 2 | 20.180 | 10.090 | 428.866 | <2E-16 | *** |
| $v_c$ | 2 | 0.792 | 0.396 | 16.824 | 0.000 | *** |
| $a_p$:f | 4 | 1.088 | 0.272 | 11.564 | 0.000 | *** |
| $a_p$:$v_c$ | 4 | 1.797 | 0.449 | 19.100 | 0.000 | *** |
| f:$v_c$ | 4 | 0.317 | 0.079 | 3.369 | 0.010 | * |
| $a_p$:f:$v_c$ | 8 | 1.094 | 0.137 | 5.811 | 0.000 | *** |
| Residuals | 297 | 6.988 | 0.024 | | | |

---

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 17. Experiment 1, surface roughness average-parameters Ra analysis. The AIC test results. The best model on top.

| Model name | Model | K | AICc | Delta_AICc | AICcWt | Cum.Wt | LL |
|---|---|---|---|---|---|---|---|
| Exp1.ANOVA.Ra.3 | Ra ~ $a_p$ * f * $v_c$ | 28 | -262.08 | 0.00 | 1.00 | 1.00 | 161.79 |
| Exp1.ANOVA.Ra.2 | Ra ~ $a_p$ * f + $v_c$ | 12 | -176.16 | 85.92 | 0.00 | 1.00 | 100.58 |
| Exp1.ANOVA.Ra.1 | Ra ~ $a_p$ + f + $v_c$ | 8 | -151.85 | 110.23 | 0.00 | 1.00 | 84.15 |

In the light of exemplification, Figure 43 shows three surface roughness profiles acquired in Experiment 1. All of them are from the first replica and the third measurement in each condition. It is possible to see the difference provoked by the feed rate in the surface roughness parameters, as it is the only highly-influencing parameter that changes from one graph to another. The first graph at the top has a small Ra and Rt consistent with the low feed rate, and the surface roughness grows as the feed grows in both the middle and bottom graphs.

Another noticing point is that the RSm is approximately equal to the feed rate, as envisioned earlier in the experimental planning. Rsk and Rku seem to grow from the first graph at the top to the last graph at the bottom, but further analysis is necessary to conclude whether they positively correlate with the feed rate or correlate more complex with other factors.

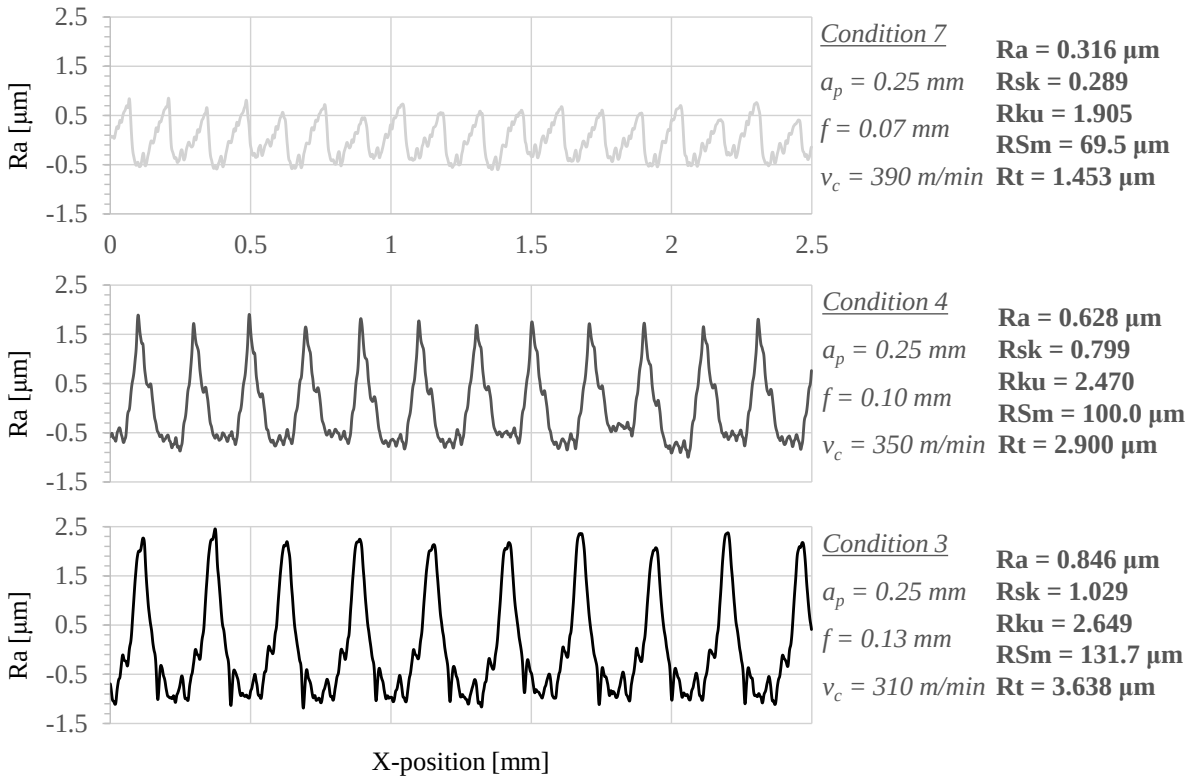



Figure 43. Surface roughness for three different conditions of Experiment 1.

The same statistical approach was adopted to analyze the cutting forces responses. For the resultant force (F), the depth of cut ($a_p$) was the most contributor parameter (Table 18), followed by the feed rate and the interaction between both. The contribution of the cutting speed was not statistically significant (as suggested in Figure 42). Five models were tested through the AIC test (Table 19). 1: Three parameters with isolated effects ($F \sim a_p + f + v_c$); 2: Interaction between the depth of cut and feed rate and cutting speed as an isolated variable ($F \sim a_p * f + v_c$); 3: Interaction between the three parameters ($F \sim a_p * f * v_c$); 4: Interaction between the feed rate and the cutting speed and the depth of cut as an isolated variable ($F \sim a_p + f * v_c$); 5: An artificial variable that corresponds to the theoretical chip area (multiplication of the feed rate by the depth cut values) as the influencing factor (F ~ ChipA). Again, the residuals' diagnosis is depicted in Appendix B – Additional statistical results.

Table 18. Factorial ANOVA for the effect of the parameters on F.

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| $a_p$ | 2 | 417139 | 208569 | 4406.35 | <2E-16 | *** |
| f | 2 | 66573 | 33286 | 703.23 | <2E-16 | *** |
| $v_c$ | 2 | 80 | 40 | 0.84 | 0.437 | |
| $a_p$:f | 4 | 25435 | 6359 | 134.34 | <2E-16 | *** |
| Residuals | 43 | 2035 | 47 | | | |

---

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 19. Experiment 1, resultant force F analysis. The AIC test results. The best model on top.

| Model name | Model | K | AICc | Delta_AICc | AICcWt | Cum.Wt | LL |
|---|---|---|---|---|---|---|---|
| Exp1.ANOVA.F.2 | $F \sim a_p * f + v_c$ | 12 | 380.84 | 0.00 | 1.00 | 1.00 | -174.62 |
| Exp1.ANOVA.F.3 | $F \sim a_p * f * v_c$ | 28 | 463.11 | 82.27 | 0.00 | 1.00 | -171.08 |
| Exp1.ANOVA.F.5 | F ~ ChipA | 3 | 465.51 | 84.67 | 0.00 | 1.00 | -229.52 |
| Exp1.ANOVA.F.1 | $F \sim a_p + f + v_c$ | 8 | 508.97 | 128.12 | 0.00 | 1.00 | -244.88 |
| Exp1.ANOVA.F.4 | $F \sim a_p + f * v_c$ | 12 | 521.30 | 140.46 | 0.00 | 1.00 | -244.85 |

### 5.1.2 The Preparation Phase for Experiment 2

The Preparation Phase, in which the cutting tools were pre-worn to be used in Experiment 2, had results mainly regarding tools condition assessment and surface roughness measurements.

The first result of this phase, depicted in Figure 44, is the relation between tool life, measured by the flank wear value (VB), and the cutting time. From the literature, three wear rate stages along the cutting tool's life are expected: (I) Initial wear, (II) Steady-state wear, and (III) Accelerated wear region (VAUGHN, 1966). However, the position and time in which each stage occurs may vary from one system tool-material-parameters to another. From the results of the Preparation Phase, it was assumed for this set of tool-material-parameters that from a VB $\geq$ 0.1 mm, the tool wear remains in a stage of more slow wear rate until it reaches a VB $\geq$ 0.20 mm when it starts a more accelerated wear rate. The threshold of area (III) is more unclear because the range of 0.15 $\leq$ VB $\leq$ 0.2 mm has not a clear horizontal slope, and it seems that the third area could initiate at VB $\geq$ 0.15 mm in some cases. Therefore, this range of flank wear was avoided.

From a visual perspective, looking at tools 81 and 23, it was decided to select the VB value of 0.1 mm as a target for mid-life tools in Experiment 2. The other tools were worn based on the DoE of Experiment 2: it was needed two tools at the end of life (VB ≥ 0.3 mm, tools 81 and 23), and two tools at their mid-life, desirably, with a slow wear rate (0.1 ≤ VB ≤ 0.15 mm, tools 13 and 33).

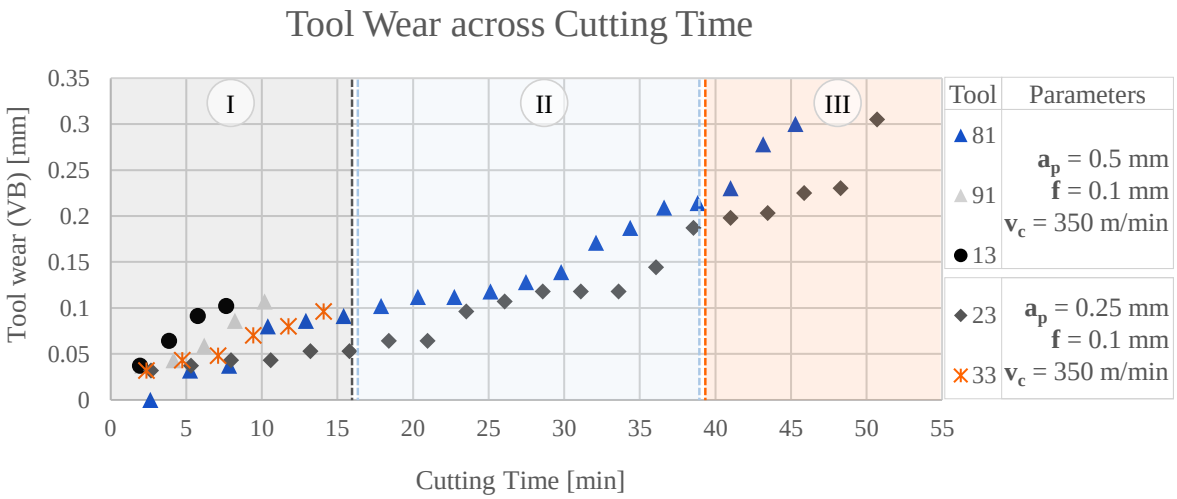



Figure 44. Tool wear across cutting time for five cutting tools.

Figure 45 shows the surface roughness through the development of flank wear VB in the Preparation Phase. Again, from a visual perspective, the tool condition, measured in terms of the flank wear VB, seems not to affect the surface roughness in an obvious way. However, although it was not statistically verified, there may be a slight decrease in Ra while VB increases that can be explained as an effect of a tool's edge rounding due to wear.

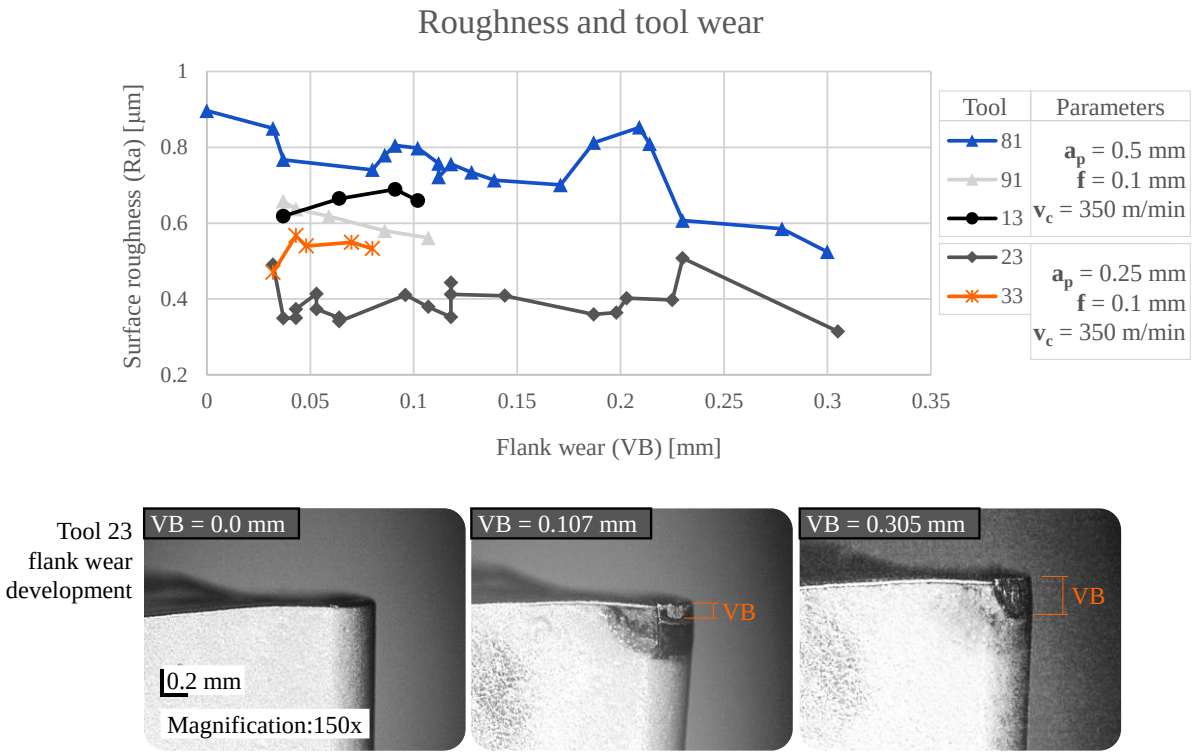



Figure 45. Average surface roughness (Ra) with the development of flank wear VB. Digital microscope pictures of three stages of flank development for tool 23.

A similar non-obvious correlation with the tool condition is seen in Rku and Rt (Figure 46). Rsk was the only surface roughness parameter that seemed to decrease with the increase in the flank wear VB. The tendency line keeps the toward-descend behavior even when removing a possible outlier from the Rsk samples. A Spearman correlation test shows that the surface roughness skewness parameter Rsk correlates rho = -0.49 with the flank wear VB measure (Table 20).

Table 20. Spearman correlation test between surface roughness parameters and the flank wear.

| R parameter | rho | pvalue | |
|---|---|---|---|
| Ra | -0.0366 | 0.7945 | |
| Rsk | -0.4956 | 0.0002 | *** |
| Rku | 0.2435 | 0.0789 | |
| RSm | -0.2187 | 0.1156 | |
| Rt | 0.1309 | 0.3503 | |

---

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’

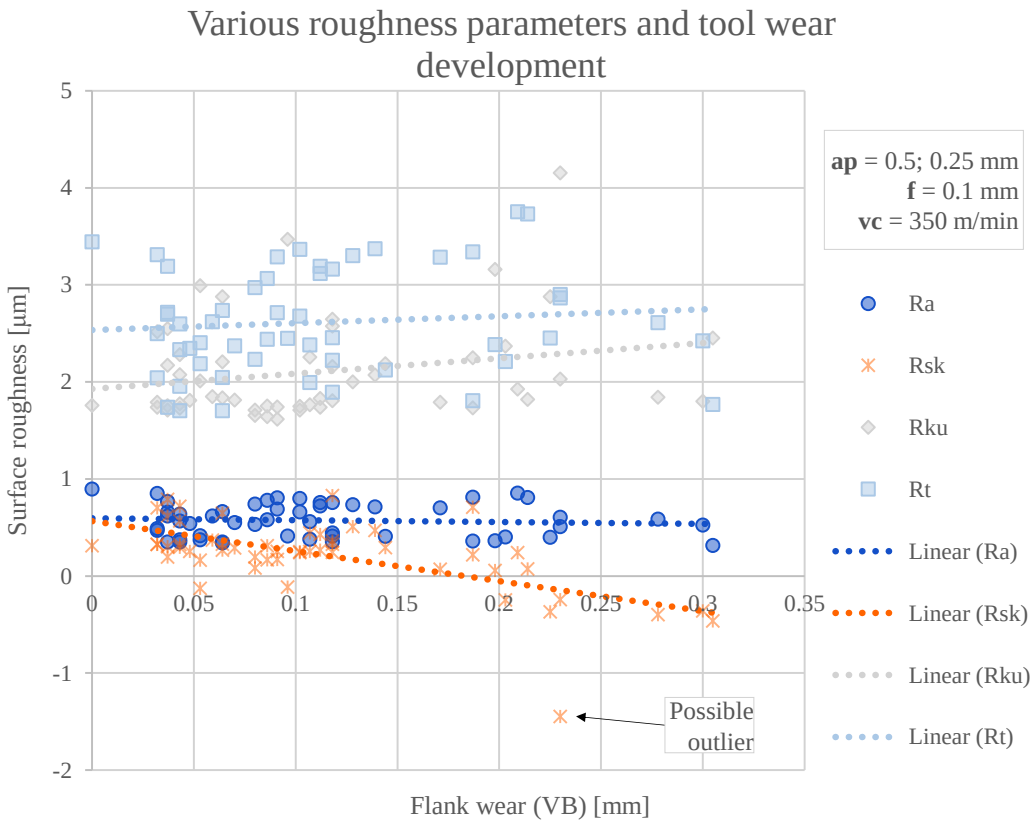



Figure 46. Various surface roughness parameters with the development of flank wear.

### 5.1.3 Experiment 2

The difference between Experiment 1 and Experiment 2 is that the tool was purposely worn to simulate actual production conditions in the second.

Similar to the first experiment, the statistical description of data is depicted in Table 21, containing the variables measured in Experiment 2, their sample size, mean, standard deviation, and minimum and maximum values. Some outliers were identified in Ra (one outlier), Rz (one), Rsk (two), RSm (three), and Rt (one) and removed by the same approach as in Experiment 1 (replaced the value by the mean of the values of the same condition). Table 21 considered the data after outlier correction. No outlier was identified in 288 samples, using Grubb's statistical test with a 5% significance level.

Table 21. Information of measured variables of Experiment 2, including the search for outliers.

Experiment 2: information of measured variables

| Variable | N | Mean | StDev | Min | Max | G | P |
|---|---|---|---|---|---|---|---|
| Ra | 288 | 0.63 | 0.23 | 0.14 | 1.23 | 2.61 | 1.00 |
| Rz | 288 | 2.85 | 0.96 | 0.71 | 4.86 | 2.24 | 1.00 |
| Rsk | 288 | 0.52 | 0.34 | -0.29 | 1.30 | 2.39 | 1.00 |
| Rku | 288 | 2.40 | 0.57 | 1.44 | 4.21 | 3.16 | 0.42 |
| RSm | 288 | 90.59 | 28.04 | 31.40 | 134.40 | 2.11 | 1.00 |
| Rt | 288 | 3.12 | 1.05 | 0.84 | 5.78 | 2.53 | 1.00 |
| $F_x$ | 288 | 127.75 | 47.32 | 51.50 | 230.00 | 2.16 | 1.00 |
| $F_y$ | 288 | 122.49 | 64.13 | 44.64 | 243.70 | 1.89 | 1.00 |
| $F_z$ | 288 | 83.51 | 51.50 | 24.17 | 206.60 | 2.39 | 1.00 |
| F | 288 | 199.54 | 86.50 | 73.73 | 384.26 | 2.14 | 1.00 |

The tool condition affected the average surface roughness parameter (Ra) with statistical significance in the second experiment (Table 22). The feed rate was again the most influential factor in this surface roughness parameter. The depth of cut was the second most influential factor, followed by the interactions between those two process parameters ($a_p$ and f) with the tool condition.

Table 22. Factorial ANOVA for the average surface roughness Ra in Experiment 2.

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| $a_p$ | 1 | 2.101 | 2.1013 | 468.35 | < 2E-16 | *** |
| f | 3 | 6.686 | 2.2286 | 496.73 | < 2E-16 | *** |
| Tcond | 2 | 0.092 | 0.046 | 10.25 | 0.000 | *** |
| $a_p$:f | 3 | 0.841 | 0.2802 | 62.45 | < 2E-16 | *** |
| $a_p$:Tcond | 2 | 1.671 | 0.8355 | 186.24 | < 2E-16 | *** |
| f:Tcond | 6 | 1.892 | 0.3153 | 70.28 | < 2E-16 | *** |
| $a_p$:f:Tcond | 6 | 0.812 | 0.1354 | 30.18 | < 2E-16 | *** |
| Residuals | 264 | 1.184 | 0.0045 | | | |

---

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

The AIC test was used again to select the best model between six models of the surface roughness Ra. The selected-as-best model by the AIC test, between five models, was the one that correlated the surface roughness Ra with the feed rate, depth of cut, tool condition, and the interaction between the three factors (Table 23).

Table 23. Experiment 2, surface roughness average-parameters Ra analysis. The AIC test results. The best model on top.

| Model name | Model | K | AICc | Delta_AICc | AICcWt | Cum.Wt | LL |
|---|---|---|---|---|---|---|---|
| Exp2.ANOVA.Ra.4 | Ra ~ $a_p$ * f * Tcond | 25 | -713.27 | 0.00 | 1.00 | 1.00 | 384.12 |
| Exp2.ANOVA.Ra.5 | Ra ~ f + ($a_p$ * Tcond) + (f * Tcond) | 16 | -480.89 | 232.38 | 0.00 | 1.00 | 257.45 |
| Exp2.ANOVA.Ra.3 | Ra ~ $a_p$ + f * Tcond | 14 | -351.87 | 361.41 | 0.00 | 1.00 | 190.70 |
| Exp2.ANOVA.Ra.2 | Ra ~ f + $a_p$ * Tcond | 10 | -346.94 | 366.33 | 0.00 | 1.00 | 183.87 |
| Exp2.ANOVA.Ra.1 | Ra ~ $a_p$ + f + Tcond | 8 | -263.99 | 449.28 | 0.00 | 1.00 | 140.26 |

For exemplification, three different surface profiles achieved in Experiment 2 are depicted in Figure 47. The graph at the top was achieved with a new tool, the middle graph with a mid-life tool (VB ~ 0.1 mm), and the bottom one with an end-life tool (VB ≥ 0.3 mm). Besides the tool condition, the three profiles were produced with the same configuration. Therefore, they should have had similar (or even theoretically equal) surface roughness parameters and visual aspects. The fact that the surface roughness parameters change from one profile to the other corroborates the influence of the tool wear on the surface finish. Moreover, in a given production, if one looks only at Ra or Rt, depending on the quality acceptance range, the three profiles could be accepted, even with the third one having a significant difference in the kurtosis of the profile—top and middle graphs correspond to Rku < 3 (smoother surface) while the bottom graph has Rku > 3.

The surface roughness profiles of Figure 47 are single examples of Experiment 2 only to represent a possible situation when dealing with tool wear. However, they do not represent a pattern or rule to what happens when the tool wears. The tool wear is the sum of complex phenomena, and the same applies to the surface generation process, which means that this can not be used as a rule. Even in Experiment 2, different behavior and examples could be used to refute this kind of statement. The objective behind Figure 47 is to show that tool wear makes predicting the entire surface roughness profile a complex task. However, given that Ra and Rt have been affected by the tool wear to a lower degree, it can be that predicting these two surface roughness parameters is a more feasible task.

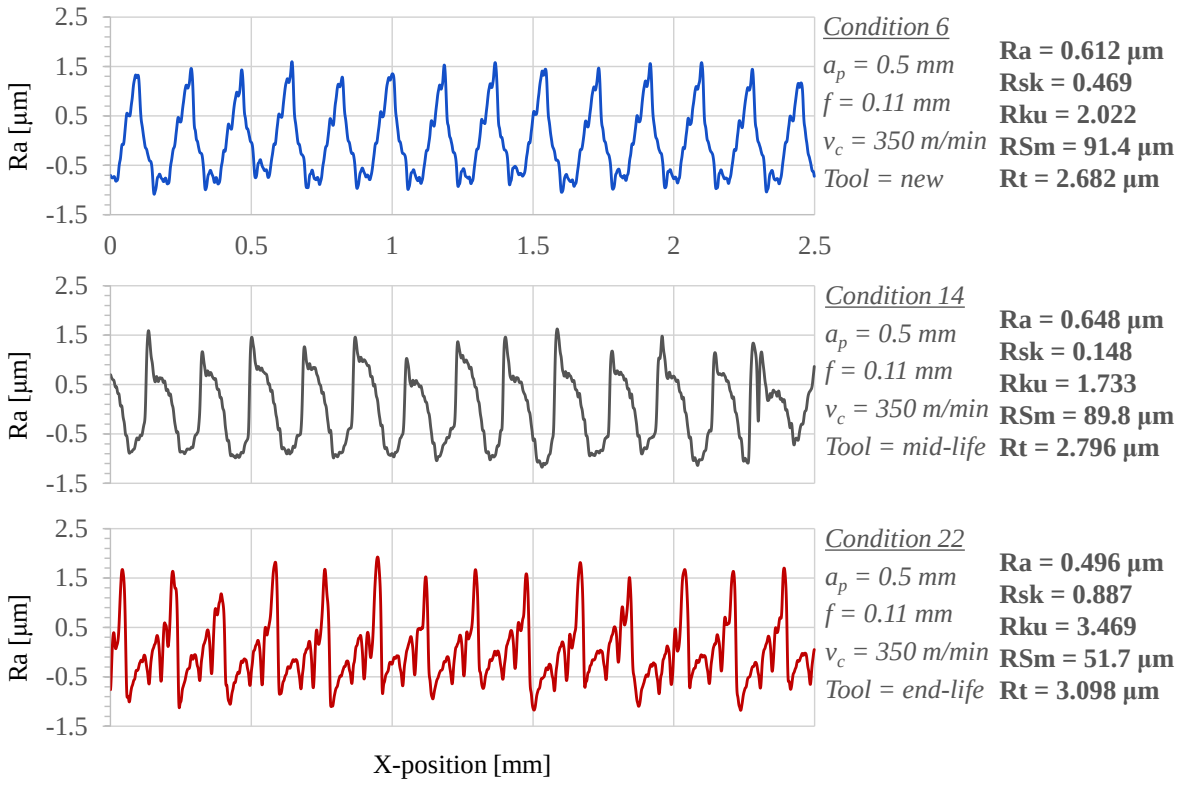



Figure 47. Different surface finish profiles for three different tool conditions in Experiment 2. The profiles were generated with: a new tool (top graph), mid-life tool (middle graph), and end-life tool (bottom graph).

The same statistical approach of Experiment 1 was repeated for the machining forces of Experiment 2. Table 24 and Table 25 show the factorial ANOVA results and the AIC test table for choosing the best model, which was the one that considered the interactions ap:f, Tcond:f, and Tcond:ap.

Table 24. ANOVA results for the machining force in Experiment 2.

| | Df | Sum Sq | Mean Sq | F value | Pr(>F) | |
|---|---|---|---|---|---|---|
| $a_p$ | 1 | 89278 | 89278 | 1219.802 | < 2E-16 | *** |
| f | 3 | 14359 | 4786 | 65.396 | 2.90E-13 | *** |
| Tcond | 2 | 240478 | 120239 | 1642.83 | < 2E-16 | *** |
| $a_p$:f | 3 | 1431 | 477 | 6.517 | 0.001584 | ** |
| f:Tcond | 6 | 2451 | 408 | 5.58 | 0.000554 | *** |
| $a_p$:Tcond | 2 | 7637 | 3818 | 52.172 | 1.71E-10 | *** |
| Residuals | 30 | 2196 | 73 | | | |

---

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 25. Experiment 2, resultant force F analysis. The AIC test results. The best model on top.

| Model name | Model | K | AICc | Delta_AICc | AICcWt | Cum.Wt | LL |
|---|---|---|---|---|---|---|---|
| Exp2.ANOVA.F.5 | $F \sim a_p * f + (Tcond * f) + (Tcond * a_p)$ | 19 | 384.87 | 0.00 | 0.99 | 0.99 | -159.86 |
| Exp2.ANOVA.F.2 | $F \sim f + a_p * Tcond$ | 10 | 394.54 | 9.67 | 0.01 | 1.00 | -184.30 |
| Exp2.ANOVA.F.4 | $F \sim a_p * f * Tcond$ | 25 | 406.70 | 21.83 | 0.00 | 1.00 | -148.80 |
| Exp2.ANOVA.F.1 | $F \sim a_p + f + Tcond$ | 8 | 427.35 | 42.48 | 0.00 | 1.00 | -203.83 |
| Exp2.ANOVA.F.3 | $F \sim a_p + f * Tcond$ | 14 | 438.94 | 54.07 | 0.00 | 1.00 | -199.10 |

Given the influence of the tool wear on the resultant force seen in the ANOVA results, the main effects plot is depicted in Figure 48 for the resultant force and its components. Similar growing (positive correlation) behavior is observed for the cutting, passive, feed, and resultant forces, indicating the tool wear's crucial influence on the machining forces.

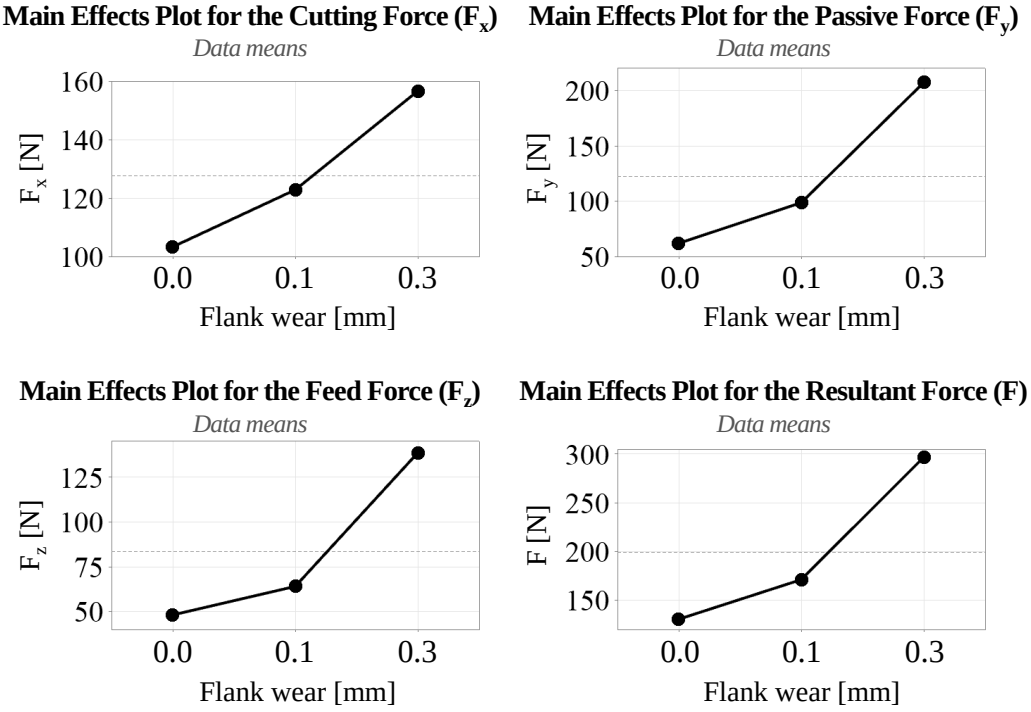



Figure 48. The main effects of flank wear on the forces.

### 5.1.4 Combined results of Experiment 1 and Experiment 2

The results presented in topics 5.1.1 and 5.1.3 for Experiment 1 and Experiment 2, respectively, treated the average surface roughness Ra and the resultant force F. Figure 49 shows the contribution of each factor to the responses of interest, extending the results to the

other surface roughness parameters Rsk, Rku, RSm, and Rt and the force components $F_x$, $F_y$, and $F_z$.

For Experiment 1, the same selection of the best among three models by the AIC test for Ra was conducted for the other chosen surface roughness parameters. Model 3, which considers the interaction of the three parameters ($Ra \sim a_p * f * v_c$), was selected again for all the other surface roughness parameters.

The contribution and statistical significance level for each factor/interaction in the surface roughness responses are depicted in Figure 49. The feed rate primarily influences the average height Ra and the maximum height Rt, respectively, with 57.7% and 66.7%. Moreover, approximately 15% of the responses are explained by the depth of cut, cutting speed, and their interaction.

The skewness is influenced mainly by the feed rate (31.4%), the interaction between the depth of cut and cutting speed (13.6%), and the interaction between the three factors (7.6%). The kurtosis is mainly (and evenly) influenced by the feed rate (15.1%) and the interaction of the three parameters (15.1%). However, given the significant contribution of the residuals, it seems that these two parameters (Rsk and Rku) are not explained only by $a_p$, f, and $v_c$, and the effects of unknown factors are significant.

Given the circumstances of the turning process in which a periodic surface is formed, the RSm, a spacing surface roughness parameter, is mainly influenced by the feed rate (90%). This result follows what is understood in the literature: in turning processes, the RSm is approximately equal to the feed rate and though highly correlated to it (TOMOV; KUZINOVSKI; CICHOSZ, 2016).

Still, in Experiment 1, the same statistical investigation for the resultant force F in topic 5.1.1 was performed with the machining force components $F_x$, $F_y$, and $F_z$. Model 2 ($F \sim a_p * f + v_c$) was selected by the AIC test in all cases. The percentage contribution of each parameter and the statistical significance level is depicted in Figure 49. The depth of cut was the most influential parameter on the four forces, with more than 80% of the contribution for F, $F_x$, and $F_z$. The second most influential parameter was the feed rate. Especially in $F_y$'s responses, the feed rate and depth of cut were almost evenly influential, with 43.3% and 50.9%, respectively. The interaction between depth of cut and feed rate was influential, with less than 10% for the three forces components and the resultant force. The cutting speed was not influential and not statistically significant in all cases.

## ANOVA CONTRIBUTION RESULTS

### EXP 1: ROUGHNESS

| | Ra | | Rsk | | Rku | | RSm | | Rt | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Factor* | *Contr.* | *Signif.* | *Contr.* | *Signif.* | *Contr.* | *Signif.* | *Contr.* | *Signif.* | *Contr.* | *Signif.* |
| $a_p$ | 8.1% | *** | 3.0% | *** | 2.2% | ** | 0.9% | *** | 6.5% | *** |
| f | 57.5% | *** | 31.4% | *** | 15.1% | *** | 90.0% | *** | 66.7% | *** |
| $v_c$ | 2.3% | *** | 1.5% | ** | 0.1% | NS | 0.7% | *** | 2.6% | *** |
| $a_p$:f | 3.1% | *** | 6.1% | *** | 12.9% | *** | 0.1% | NS | 1.5% | *** |
| $a_p$:$v_c$ | 5.1% | *** | 13.6% | *** | 6.1% | *** | 1.2% | *** | 4.8% | *** |
| f:$v_c$ | 0.9% | * | 3.3% | *** | 1.9% | * | 0.1% | NS | 0.7% | ** |
| $a_p$:f:$v_c$ | 3.1% | *** | 7.6% | *** | 15.1% | *** | 0.1% | NS | 2.8% | *** |
| Residuals | 19.9% | | 33.6% | | 46.7% | | 6.9% | | 14.4% | |

### EXP 2: ROUGHNESS

| | Ra | | Rsk | | Rku | | RSm | | Rt | |
|---|---|---|---|---|---|---|---|---|---|---|
| *Factor* | *Contr.* | *Signif.* | *Contr.* | *Signif.* | *Contr.* | *Signif.* | *Contr.* | *Signif.* | *Contr.* | *Signif.* |
| $a_p$ | 13.8% | *** | 4.8% | *** | 1.9% | *** | 7.5% | *** | 12.0% | *** |
| f | 43.8% | *** | 5.0% | *** | 4.5% | *** | 54.3% | *** | 46.5% | *** |
| Tcond | 0.6% | *** | 41.9% | *** | 28.3% | *** | 6.6% | *** | 7.6% | *** |
| $a_p$:f | 5.5% | *** | 5.0% | *** | 5.9% | *** | 1.0% | *** | 2.8% | *** |
| $a_p$:Tcond | 10.9% | *** | 0.0% | NS | 11.0% | *** | 10.0% | *** | 4.3% | *** |
| f:Tcond | 12.4% | *** | 15.1% | *** | 20.7% | *** | 4.6% | *** | 10.3% | *** |
| $a_p$:f:Tcond | 5.3% | *** | 3.4% | *** | 4.8% | *** | 4.0% | *** | 5.4% | *** |
| Residuals | 7.7% | | 24.7% | | 23.0% | | 12.0% | | 11.0% | |

### EXP 1: MACHINING FORCES

| | F | | $F_x$ | | $F_y$ | | $F_z$ | |
|---|---|---|---|---|---|---|---|---|
| *Factor* | *Contr.* | *Signif.* | *Contr.* | *Signif.* | *Contr.* | *Signif.* | *Contr.* | *Signif.* |
| $a_p$ | 81.6% | *** | 80.0% | *** | 50.9% | *** | 86.7% | *** |
| f | 13.0% | *** | 15.7% | *** | 43.3% | *** | 5.4% | *** |
| $v_c$ | 0.0% | NS | 0.0% | NS | 0.0% | NS | 0.0% | NS |
| $a_p$:f | 5.0% | *** | 4.0% | *** | 4.8% | *** | 7.3% | *** |
| Residuals | 0.4% | | 0.4% | | 0.9% | | 0.6% | |

| | | |
|---|---|---|
| Exp1: | High | Low |
| Exp2: | High | Low |

Signif. Codes (Pr(>F)): 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1, NS: Non-significative

### EXP 2: MACHINING FORCES

| | F | | $F_x$ | | $F_y$ | | $F_z$ | |
|---|---|---|---|---|---|---|---|---|
| *Factor* | *Contr.* | *Signif.* | *Contr.* | *Signif.* | *Contr.* | *Signif.* | *Contr.* | *Signif.* |
| $a_p$ | 24.9% | *** | 59.0% | *** | 2.7% | *** | 33.3% | *** |
| f | 4.0% | *** | 15.3% | *** | 1.3% | *** | 0.2% | . |
| Tcond | 67.2% | *** | 21.9% | *** | 93.1% | *** | 59.0% | *** |
| $a_p$:f | 0.4% | ** | 1.2% | *** | 0.1% | NS | 0.1% | NS |
| f:Tcond | 0.7% | *** | 0.2% | ** | 0.8% | ** | 0.4% | . |
| $a_p$:Tcond | 2.1% | *** | 2.1% | *** | 1.2% | *** | 6.1% | *** |
| Residuals | 0.6% | | 0.3% | | 0.8% | | 0.9% | |

Figure 49. ANOVA contributions of each factor to average roughness Ra and the Resultant Force F for Exp1 and Exp2.

For Experiment 2, using the AIC test, the same models that were selected for Ra and F were selected for the other surface roughness parameters and the force components (Ra ~ $a_p$ * f * Tcond and F ~ $a_p$ * f + (Tcond * f) + (Tcond * $a_p$)). Factorial ANOVA models were made for each force component and the other surface roughness parameters, and the results are grouped in Figure 49.

Compared to Experiment 1, the feed rate again affected the surface finish more than the depth of cut for all roughness parameters. However, the presence of tool wear in Experiment 2 changed the magnitude of contributions. It is the case for the feed rate, in which its contribution percentage is smaller for all cases compared to the percentages in Experiment 1. The tool wear was also statistically significant for all the selected roughness parameters. For the average roughness Ra, tool wear has contributed less than 1%, but its contribution is relatively significant and higher than 5% for all the others. In the total height Rt and the mean width of profile RSm, the tool condition contribution was 7.6% and 6.6%, respectively.

Moreover, the interaction between the tool condition and the other factors contributed to the roughness responses. Both interactions between tool condition and depth of cut and tool condition and feed rate contributed around 11% to Ra. In addition, 10% was the contribution of both the interaction between tool condition and depth of cut to RSm and the interaction between tool condition and feed rate to Rt. The triple interaction between tool condition, depth of cut, and feed rate contributed around 4% to all roughness parameters.

The percentage contribution of the factors on the roughness parameters skewness (Rsk) and kurtosis (Rku) was a highlight in Experiment 2 compared to Experiment 1. The tool wear presence affected those two roughness parameters to a relatively great extent. It was found that the tool condition contributed 41.9% and 28.3% to Rsk and Rku, respectively. Moreover, the interaction between tool condition and feed contributed 15.1% and 20.7% to Rsk and Rku, respectively. The interaction between tool condition and depth of cut contributed 11% to Rku but was not influential to Rsk. Tool condition influence over Rsk follows the results from the Preparation Phase (5.1.2), in which a correlation between the tool condition and the skewness was detected.

Given that Rsk and Rku were affected significantly by the tool condition, it can be suggested that the untracked factors that affected these two parameters in the first experiment might be connected to the tool geometry. However, further experiments would have to be done to investigate this assumption.

Another highlighted result was the reduction in the contribution of the feed rate to the mean width of profile elements (RSm) in Experiment 2. This reduction can be caused by the

tool edge geometry changing due to tool wear, which affected the topography's periodical behavior observed in typical turning processes and Experiment 1. With the geometry of the cutting tool changing due to tool wear, the feed starts to influence differently in this sensitive-to-periodicity roughness parameter.

In Experiment 1, the depth of cut was the independent variable that mainly influenced the machining forces, followed by the feed rate. In Experiment 2, this behavior is repeated: the depth of cut is always affecting more than the feed rate. However, the tool condition significantly influences the forces for the second experiment. It is the second contributor to $F_x$ and the first to $F_y$ and $F_z$. Especially for the passive force, the tool condition has a surprising contribution of more than 90%.

## 5.2 Results of the artificial neural network developments

The results of the hyperparameters tuning tests (described in topic 4.2.2) are depicted in Figure 50. In those tests, the hyperparameter options/levels were compared by the results (in terms of the mean squared error (MSE) and mean absolute percentage error (MAPE)) in the validation set during the learning process. The values shown on the graphs are the mean from three replicas. The hyper tuning tests resulted in the following choices:

- Activation function (Figure 50, A): the Relu produced the lowest errors;
- Batch size (Figure 50, B): the batch size of 20 produced the lowest errors, followed by the batch sizes of 5 and 50;
- Optimizer algorithm (Figure 50, C): the stochastic gradient descent (SGD) algorithm produced the lowest errors;
- Optimizer learning rate (Figure 50, D): 0.01 was the chosen learning rate that produced the lowest errors for the SGD optimizer;
- Optimizer momentum (Figure 50, E): 0.95 produced the lowest errors for the SGD optimizer;
- Number of hidden layers (Figure 50, F): 3 layers produced the best results;
- Topology (Figure 50, G and H): the topology 7-28-56-14-1 produced the lowest errors. This topology was also extended to use with the configuration where only the process parameters (ap, f, and vc) are used as inputs (3-28-56-14-1).

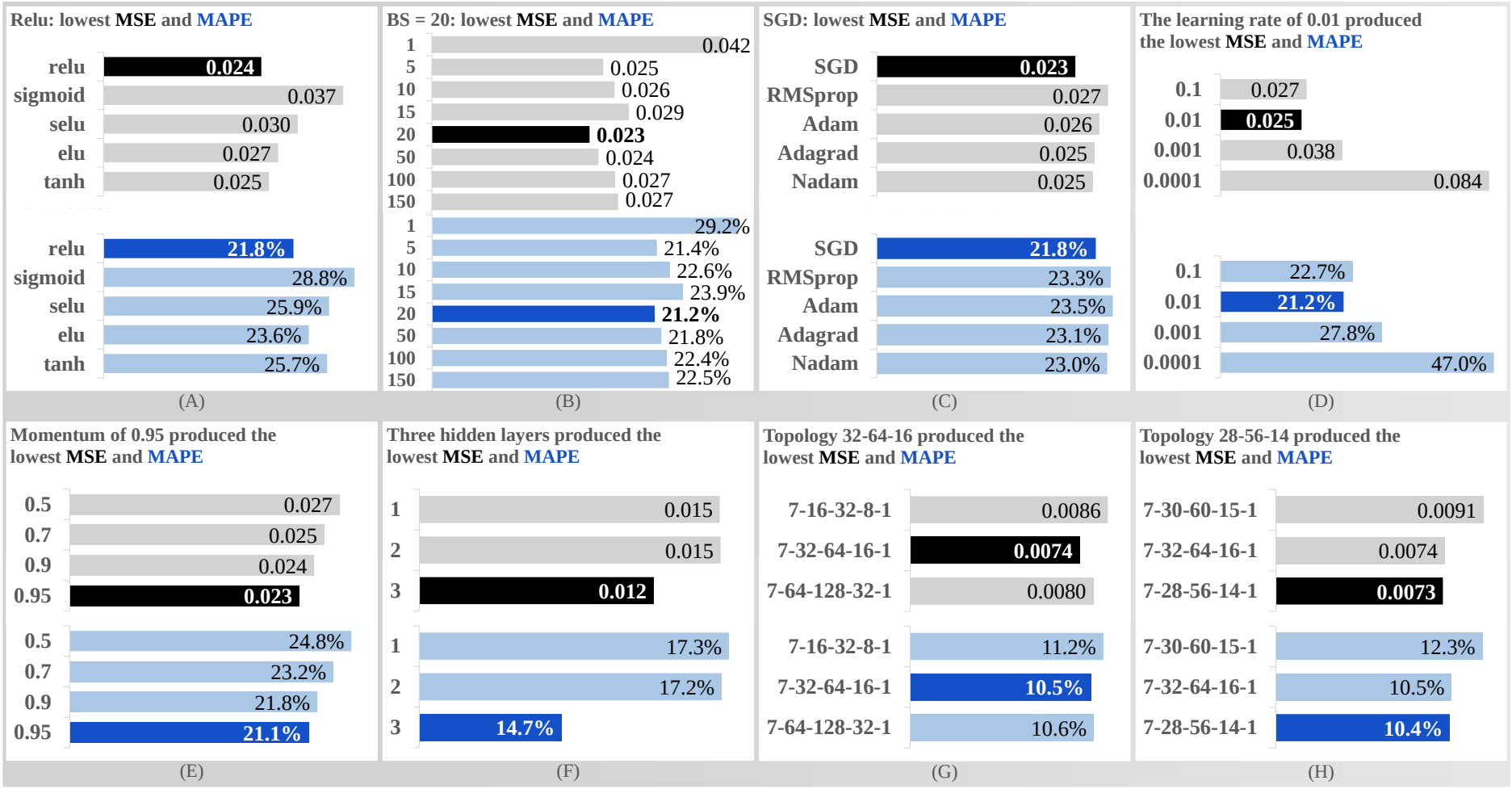



Figure 50. Results of the hyperparameter tuning for the activation function (A), batch size (B), optimizer (C), learning rate (D), momentum (E), number of hidden layers (F), topology (broad range) (G), and topology (small range) (H). All values are the mean of three replicas.

For the momentum and the number of hidden layers, the best result appeared at what was the border of the search range. When this happens, it is recommended to extend the search by testing values beyond the border (MONTAVON; ORR; MÜLLER, 2012), which would mean, in this case, testing a model with more hidden layers and a higher momentum. However, the reasonably low difference between the momentum levels, and as 0.95 is already almost the maximum momentum level, the decision was not to search further. The same happened to the hidden layers: given that the models had only one or two hidden layers in the literature papers, it was not the intention of this work to get too much further away from this pattern by increasing computational cost with more hidden layers—therefore, the decision was not to search further.

### 5.2.1 Models trained with three features: ANN-.T.L

Beforehand, an emphasis must be given to one point: it became helpful to compare the ANNs modeled with Experiment 1 to those modeled with Experiment 2. In this regard, the results will usually be presented in doubles to facilitate comparison. To remember the adopted nomenclature of models, refer again to Figure 39.

The models for experiments 1 and 2 were built under the same hyperparameter. Hence, the explanation of the following paragraph is valid for all models.

The models used a neuronal architecture of one input layer with as many neurons as inputs (3 or 7, depending on the case), three hidden layers (28-56-14), and one output layer with one output neuron. Therefore, the number of neurons and layers is depicted as 3-28-56-14-1, or 7-28-56-14-1. The SGD optimizer that showed the best result in the mean absolute percentage error (MAPE) was adopted for all models. The learning rate was chosen at 0.01 with a momentum of 0.95. The activation function of the three dense hidden layers used was the Relu function. The batch size was chosen at 20.

Figure 51 shows the results and summary of ANN1.T.L and ANN2.T.L. These models were trained with the process parameters as inputs, as shown on the top right side of the figure. The table on the top left side of the figure shows the mean results from three trained replicas. ANN2.T.L achieved mean errors lower than 22%, and ANN2.T.L performed slightly better, reaching mean errors lower than 18%. The $R^2$ metric was similar to both models, 0.74 for ANN1.T.L and 0.64 for ANN2.T.L. The variation of epoch numbers and errors between the three replicas (shown on the bottom of Figure 51, for “epochs,” “MSE,” “MAPE,” “MAE,” and “$R^2$”) was considered low, given the stochastic behavior of the neural networks’ learning process. The table in the middle of the figure shows the hyperparameters summary for models

ANN1.T.L and ANN2.T.L, given a rapid description of some hyperparameters used: batch size, early stopping parameters, optimizer parameters, and activation function.

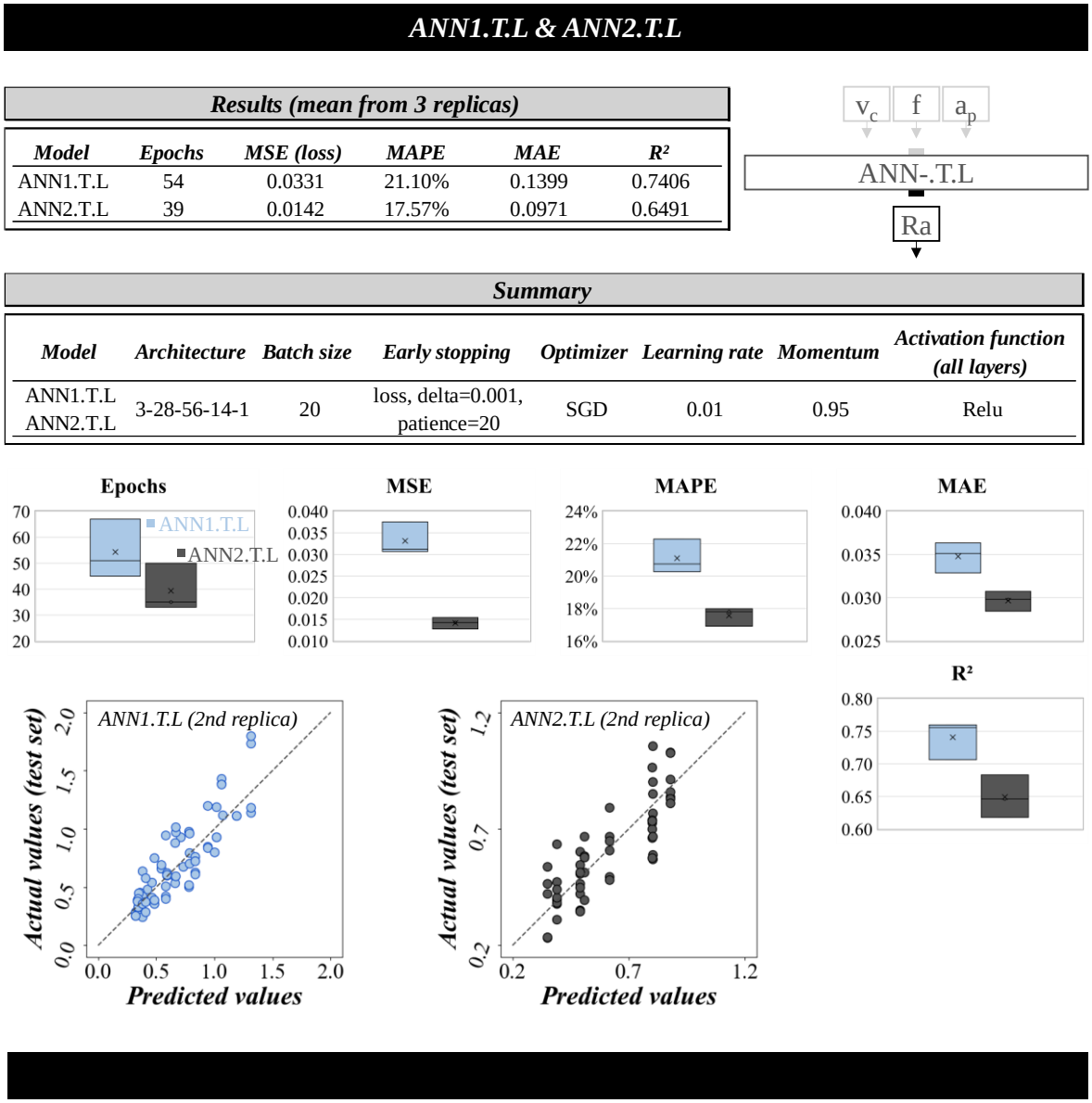



Figure 51. Mean results and summary of ANN1.T.L and ANN2.T.L.

The two graphs at the bottom of Figure 51 show the fitness of the prediction models to the test set. The dashed line represents the equality of actual and predicted values responses. More points falling along this line means that the model can predict with more accuracy the test values and, therefore, have more generalization capacity.

For these first models, it is possible to assume that both perform similarly in a never-seen test set. Furthermore, the relation of surface roughness average-parameter Ra was modeled for both experiment datasets, with a mean percentage error of about 20%. As the tool wear did

not drastically affect this surface roughness parameter (tool condition alone contributed to only 0.6% of Ra responses), it was expected that the surface roughness would be modeled in a similar performance for both experiment datasets, as indeed happened.

Figure 52 displays the learning curves of models ANN1.T.L and ANN2.T.L showing acceptable behavior without overfitting signs.

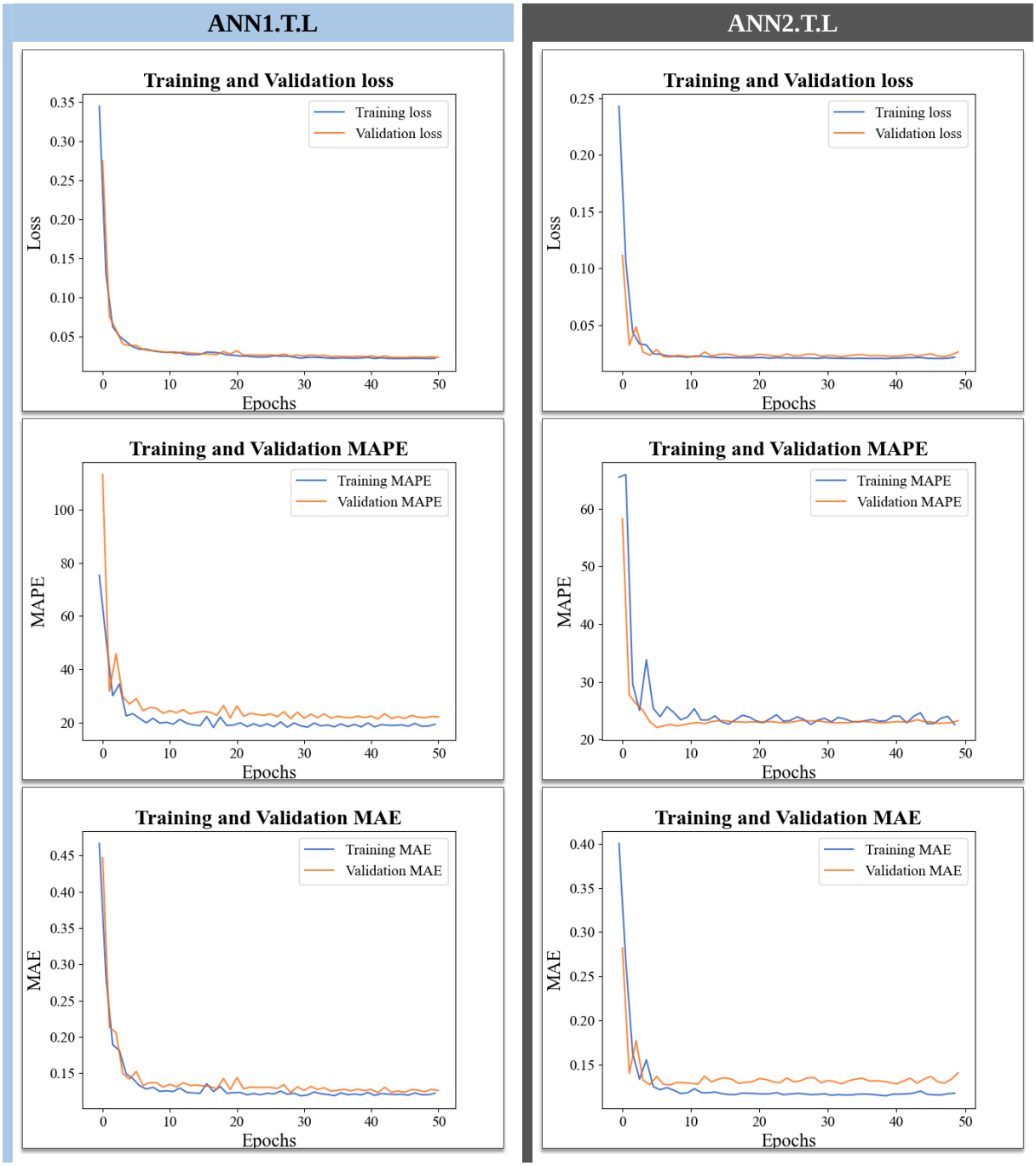



Figure 52. Learning curves for ANN1.T.L and ANN2.T.L.

For both models, the MSE error loss (Figure 52, top graph) starts as high as 0.35 and reaches values lower than 0.05. The MAPE and MAE values also decrease rapidly at the beginning reaching a horizontal slope after 10 to 35 epochs. As expected, the validation errors, represented by the orange curves, remain higher or proximately equal to the training errors, represented by the blue curves—evidence of generalization or overfitting being avoided, as discussed previously in Figure 22.

### 5.2.2 Models trained with seven features: ANN-.H.L

The second round of implemented models (ANN1.H.L and ANN2.H.L) showed better results using the machining forces as inputs besides the process parameters. The top right side of Figure 53 shows the inputs and outputs of the artificial neural network. Three models of each were trained, and the mean results of three replicas are presented in the table on the top left side of the figure. The boxplots at the bottom show the slight variation between the three replicas for the two models (concerning “epochs,” “MSE,” “MAPE,” “MAE,” and “$R^2$”). On average, ANN1.H.L and ANN2.H.L achieved MSE smaller than 0.003, MAPE around 5.5%, and MAE smaller than 0.035. As expected, each training presented a different (but similar) behavior, which can be seen in the number of epochs resulting from the early stopping technique. For models ANN-.H.L, the performance variation (errors MSE, MAPE, and MAE) between replicas was relatively small and thus acceptable.

The two graphs at the bottom of the figure show the prediction capacity for the models. The points falling neatly along the dashed line represent that the model has excellent generalization capacity in the test dataset—what was previously suggested during the model training by the well-behaved learning curves.

Figure 54 displays the learning curves of one of three replicas for the two models. They show an acceptable learning behavior without overfitting signs—the orange curves, representing the validation errors, remain higher than the blue curves, which are the training errors. The same good learning behavior is seen in all replicas.

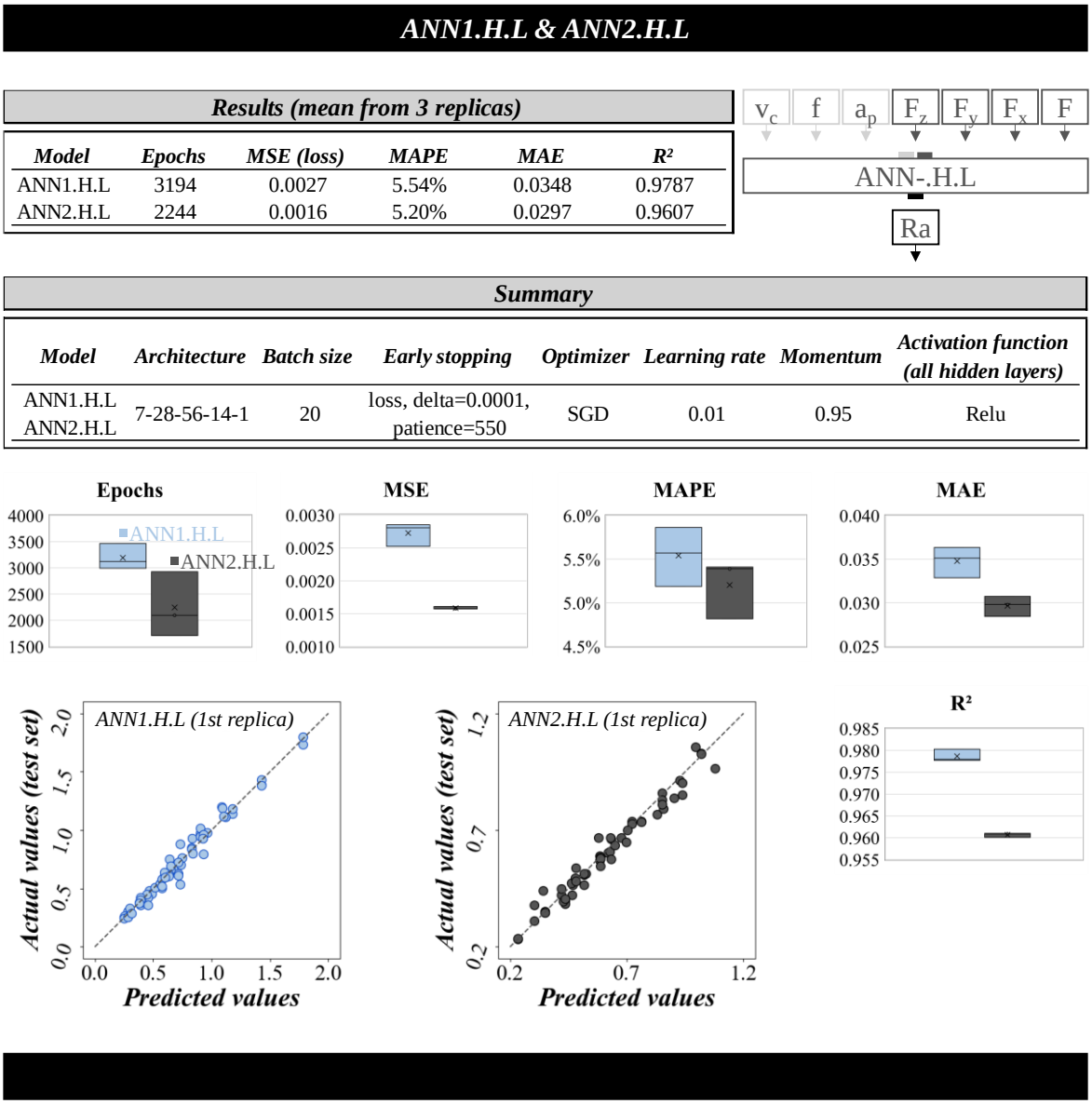



Figure 53. Mean results and summary of ANN1.H.L and ANN2.H.L.

The MAPE curves for ANN1.H.L have a bumpier aspect, which indicates that the changing in neural network parameters (weights and biases) during the learning process provokes high variance in the predictions. This may be caused by the stochastic behavior of the optimizer operating the parameters changes. A checkpoint technique avoided the error peaks by checking the parameters at every epoch's end and saving those that gave the lowest MAPE.

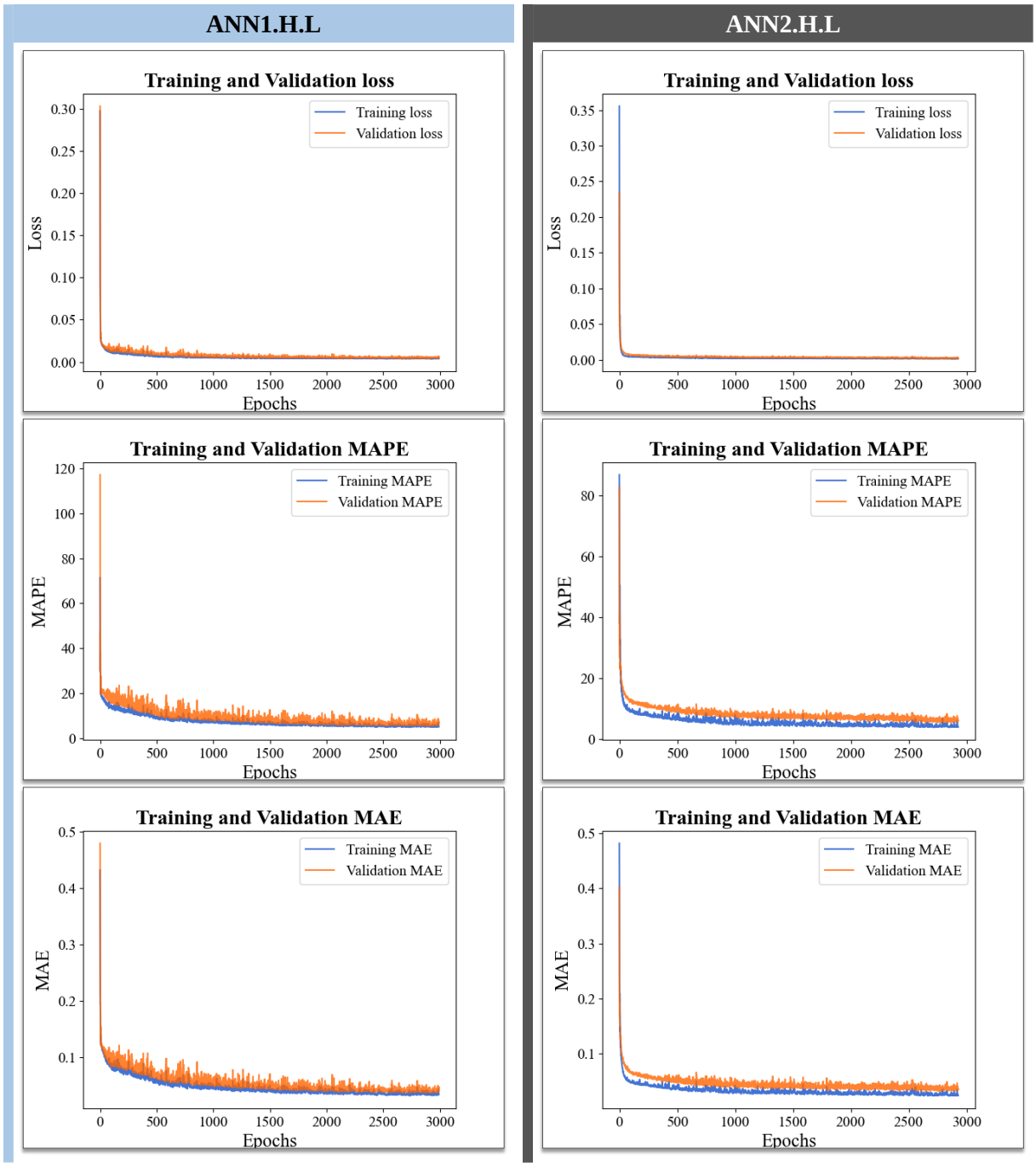



Figure 54. Learning curves for models ANN1.H.L and ANN2.H.L.

The difference in epochs between models T and H is due to different early stopping parameters used in each model. It was observed that a longer learning process with a “more patient” early stopping technique promoted a considerable reduction in the mean absolute percentage error (MAPE) for models ANN1.H.L and ANN2.H.L. While those models have more input values, the learning process may take longer. The reduction gain when using more epochs was as high as 65% (Table 26). For the models ANN1.T.L and ANN2.T.L, as the

reduction did not happen (Table 26), a “less patient” early stopping technique was used to reduce computational effort and avoid overfitting.

Table 26. Using the early stopping technique with different parameters gave different MAPE results for models ANN1.H.L and ANN2.H.L, which used more features.

| ***Early stopping with different parameters*** | | | | | | | |
|---|---|---|---|---|---|---|---|
| ***Model*** | ***Delta*** | ***Patience*** | ***Epochs*** | ***MSE (loss)*** | ***MAPE*** | ***MAE*** | ***MAPE reduction*** |
| ANN1.H.L | 0.001 | 20 | 103 | 0.0146 | 16.2788 | 0.0963 | - |
| ANN2.H.L | 0.0001 | 550 | 3194 | 0.0027 | 5.5404 | 0.0348 | 65.97% |
| ANN1.T.L | 0.001 | 20 | 54 | 0.0331 | 21.1006 | 0.1399 | - |
| ANN2.T.L | 0.0001 | 550 | 805 | 0.0296 | 21.1437 | 0.1377 | -0.20% |

By comparing the T-models (ANN1.T.L and ANN2.T.L) and the H-models (ANN1.H.L and ANN2.H.L), it is possible to see a considerable reduction in the mean absolute percentage error (MAPE), as long as in all the other metric parameters depicted in Figure 51 and Figure 53. The MAPE reduction roughly corresponds to 75%, as depicted in Figure 55.

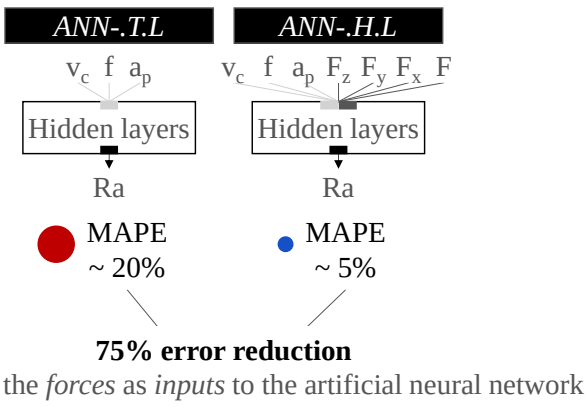



Figure 55. The performance difference between models T (using only the process parameters as inputs) and models H (using the forces and process parameters as inputs).

H-models had better performance than T-models. However, in a scenario where the MAPE of 20% is an acceptable error metric, T-models could be a better choice because they were easier to train, demanding much fewer epochs—around 46 epochs were used in T-models against 2719 on average in H-models. The number of epochs is critical in large-scale learning problems constrained by the maximum computing time (MONTAVON; ORR; MÜLLER, 2012). H-models were better suited than T-models in this study’s context, where the number of examples is the limiting agent.

The next question could be if every surface roughness model described in the literature would be improved when including the machining forces as inputs. If that is the case, it is a piece of important information to those designing cyber-physical monitoring systems for machining processes.

### 5.2.3 Models trained with smaller datasets: ANN-.-.S

Additional models were trained with small datasets to investigate the influence of the dataset size on prediction performance. The learning curves showed overfitting aspects when using the same hyperparameters as the large-dataset models—ANN-.-.L. Because of that overfitting signs, some techniques were employed: dropout, weights initialization, and regularization. The models trained with the overfitting-diminishing techniques produced good learning curves, as depicted in Figure 56 (right). Figure 56 (left) also shows the model's learning curves trained without the techniques presenting overfitting signs—the validation error increases while the training error decreases.

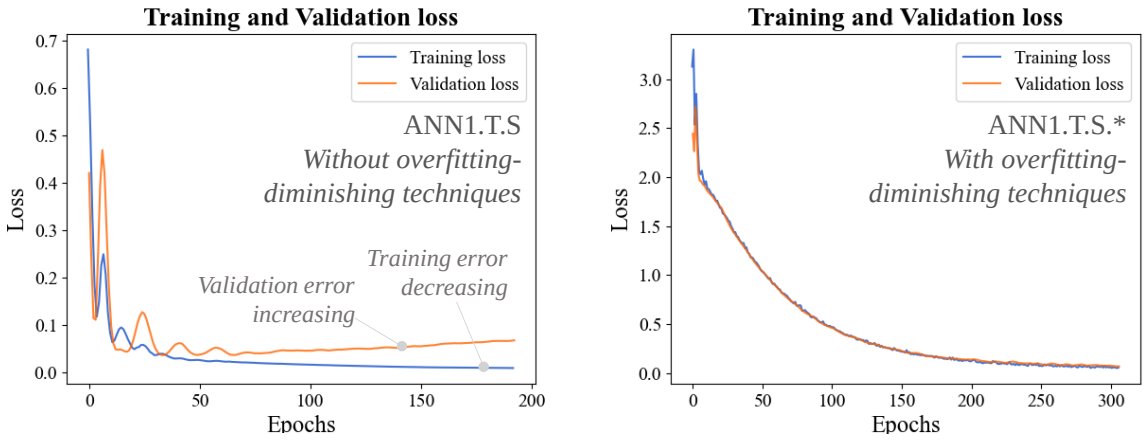



Figure 56. Learning curves of models trained with (left) and without (right) overfitting-diminishing techniques (left).

The overfitting-diminishing techniques were attached to the model-building code lines as represented below. The weights initialization, regularization, and dropout are simple to implement: each dense hidden layer receives two additional arguments—*kernel_initializer*, *kernel_regularizer*; and a dropout layer is added after each hidden layer—*model.add(layers.Dropout(0.1))*:

```
# Define standard architecture features
RegularizedDense = partial(keras.layers.Dense,
```

```
                            activation="relu",
                            kernel_initializer="he_normal",
                            kernel_regularizer=
                            keras.regularizers.L2(0.01))

# Build model
def create_model():
    model = keras.Sequential()
    model.add(keras.Input(shape=X_train_sc.shape[1:]))
    model.add(RegularizedDense(28))
    model.add(layers.Dropout(0.1))
    model.add(RegularizedDense(56))
    model.add(layers.Dropout(0.1))
    model.add(RegularizedDense(14))
    model.add(layers.Dropout(0.1))
    model.add(layers.Dense(1))
    optimizer = keras.optimizers.SGD(0.01, momentum=0.95)

    # ############## COMPILE MODEL
    model.compile(loss='mean_squared_error',
                  optimizer=optimizer,
                  metrics=['mean_absolute_error',
                           'mean_absolute_percentage_error'])

    return model
```

Comparing the models trained with large data sets and those with small datasets suggests that size matters when training neural networks. The H-models—trained with seven inputs (process parameters and forces)—showed better performance when trained with larger training sets, as depicted in Figure 57. This figure also shows the standard deviation of the MAPE error between the three replicas of the trained models (top right side) and the mean of MAPE from each model replica (bottom right side). A t-test statistic was performed between each model MAPE to verify if their means were statistically different. The result is presented in the table on the bottom right side of Figure 57—for results with the same letter, the difference between the means is not statistically significant, while for results with different letters, the means are significantly different.

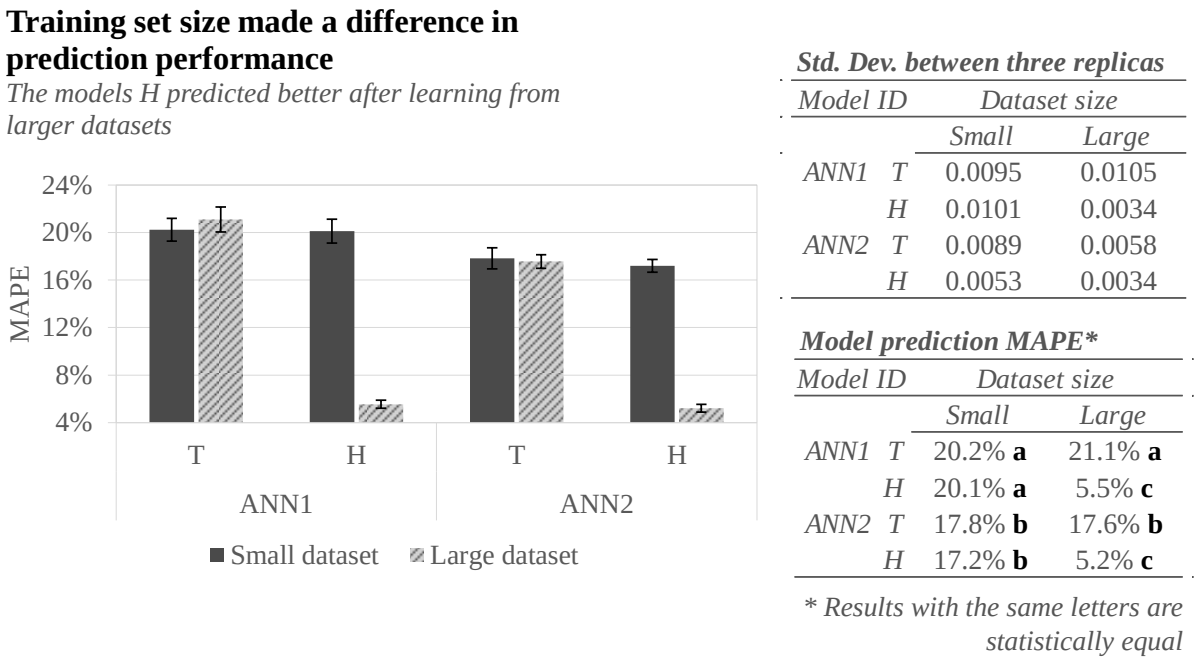



Figure 57. Comparison results from learning with small or large datasets.

The same hyperparameters of the early stopping technique as those used in models of topics 0 and 5.2.2 were used to train the models with the smaller training sets (ANN-.-.S).

### 5.2.4 Cross-dataset testing

An investigation intended to test the generalization capacity of the models outside of their test set ranges was made, as presented in topic 4.2.3. The best model trained in the Experiment 1 training dataset was used to predict the values of the Experiment 2 testing dataset. And vice versa.

Figure 58 shows that both models ANN1.H.L and ANN2.H.L generalize well, predicting never seen test values with excellent accuracy (5.5% and 5.2%, respectively) when tested in the test sets generated in the same situation as the data from which they learned. However, when both models try to predict the values of each other test sets, the model that has learned in a wear-free scenario shows much worse performance than the other one that has learned in a wear-existing scenario.

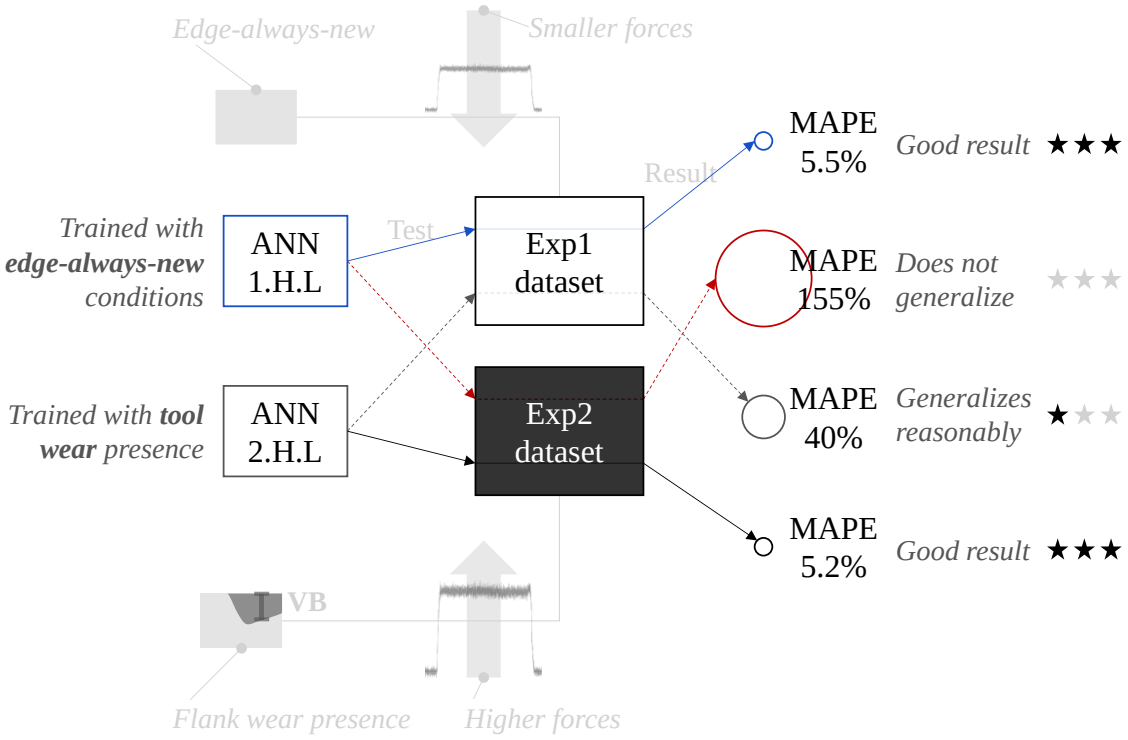



Figure 58. Prediction results of models ANN1.H.L and ANN2.H.L when tested in their own test sets and on each other test sets (cross-dataset test).

The difference in factors' levels from one experiment to the other was expected to cause a not-so-good prediction performance in the cross-dataset tests. However, the much better prediction performance observed for model ANN2.H.L is highlighted. Model ANN1.H.L has learned from a wear-free situation with smaller forces and slightly different surface roughness values (Figure 58). ANN1.H.L cannot predict the surface roughness values of the test set from Experiment 2, where the presence of tool wear exists, showing a significant error of 155% (MAPE), while ANN2.H.L can reasonably predict the surface roughness values of the test set from Experiment 1 showing an error of 40% (MAPE).

MAPE was selected to appear in Figure 58 because it was the most used metric to evaluate the models in this work. However, the other selected error metrics were also large enough to question whether models trained in wear-free scenarios can be used in actual production, where tool wear happens. Figure 59 shows the mean absolute error (MAE), mean absolute percentage error (MAPE), and mean squared error (MSE) for the cross-dataset tests. In all cases, the model ANN1.H.L performed much worse than the ANN2.H.L. Figure 59 also shows the differences between ANN2.H.L and ANN1.H.L performance metrics: 71% for MAE, 74% for MAPE, and the more considerable difference, 93% for MSE.

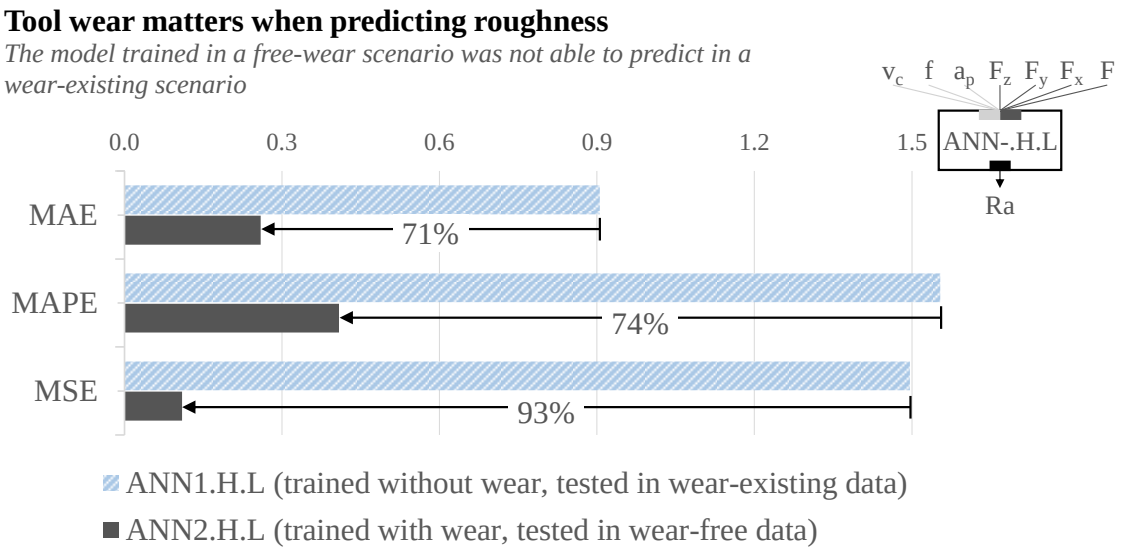



Figure 59. Results of the cross-dataset test. The error values represented by the bars are the mean from three replicas.

The next question appears to be whether this awful prediction performance of model ANN1.H.L comes from the fact that there are significant differences in the machining forces from one experiment to the other. The forces were used in models ANN1.H.L and ANN2.H.L as inputs and are essential to the models' prediction process. Maybe, the models ANN1.T.L and ANN2.T.L would perform with more negligible differences in wear-free and wear-existing scenarios because they do not depend on the forces' magnitudes to make their predictions.

Figure 60 shows the results of this hypothesis. Indeed, the performance difference was lower, but the difference is still there, which raises the proposition that the tool wear matters when building ANN models to predict surface roughness in turning processes.

Using models ANN1.T.L and ANN2.T.L, the prediction performance in the cross-dataset tests was: for ANN1.T.L, MSE = 0.0885, MAPE = 43.55%, MAE = 0.2385; for ANN2.T.L, MSE = 0.0794, MAPE = 36.23%, MAE = 0.2167.

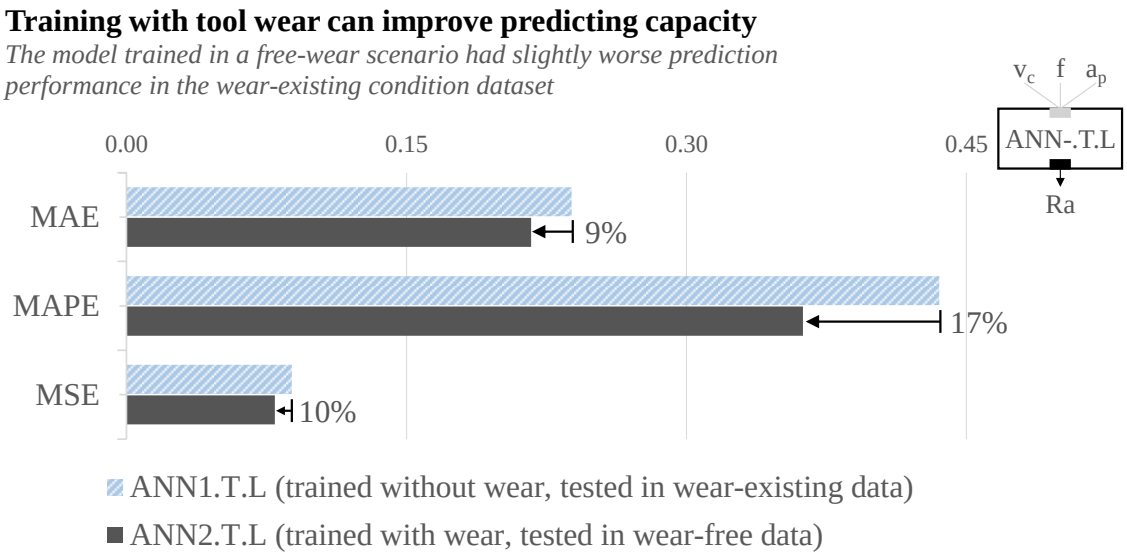



Figure 60. Results of the cross-dataset test. The error values represented by the bars are the mean from three replicas.

### 5.2.5 Comparison of theoretical and ANN models

The theoretical equation of Ra presented in topic 2.8.3 was used in the test sets of Experiment 1 and Experiment 2 (those test sets were separated when building models ANN1.H.L and ANN2.H.L, as described in topic 4.2.1). Then the prediction performance of the theoretical equation was compared to the prediction performance of the ANN models.

Figure 61 shows the predicted and actual Ra values for Experiment 1 and Experiment 2. The prediction performance using artificial neural networks exceeded those made with the theoretical equation in both experiments. For the first experiment, the theoretical equation produced even a reasonable model, falling inside the actual values range, resulting in errors (Figure 62) of reasonable magnitude (MSE = 0.084, MAPE = 35.6%, MAE = 0.217). However, for Experiment 2, the predictions made with the theoretical equation did not follow the same behavior, falling outside the range of actual values and producing mean errors of much larger magnitudes (MSE = 0.116, MAPE = 46.8%, MAE = 0.276).

In both graphs, each horizontal slope of the curve of Ra predicted by the equation corresponds to a given feed rate. In the case of the first graph, the first of those regions corresponds to the feed rate of 0.07 mm/rev, the second to 0.1 mm/rev, and the third to 0.13 mm/rev. In the second graph, as Experiment 2 explored different feed rates, the horizontal slopes correspond to the values of Ra resulting from the feeds of 0.07, 0.09, 0.11, and 0.13 mm/rev.

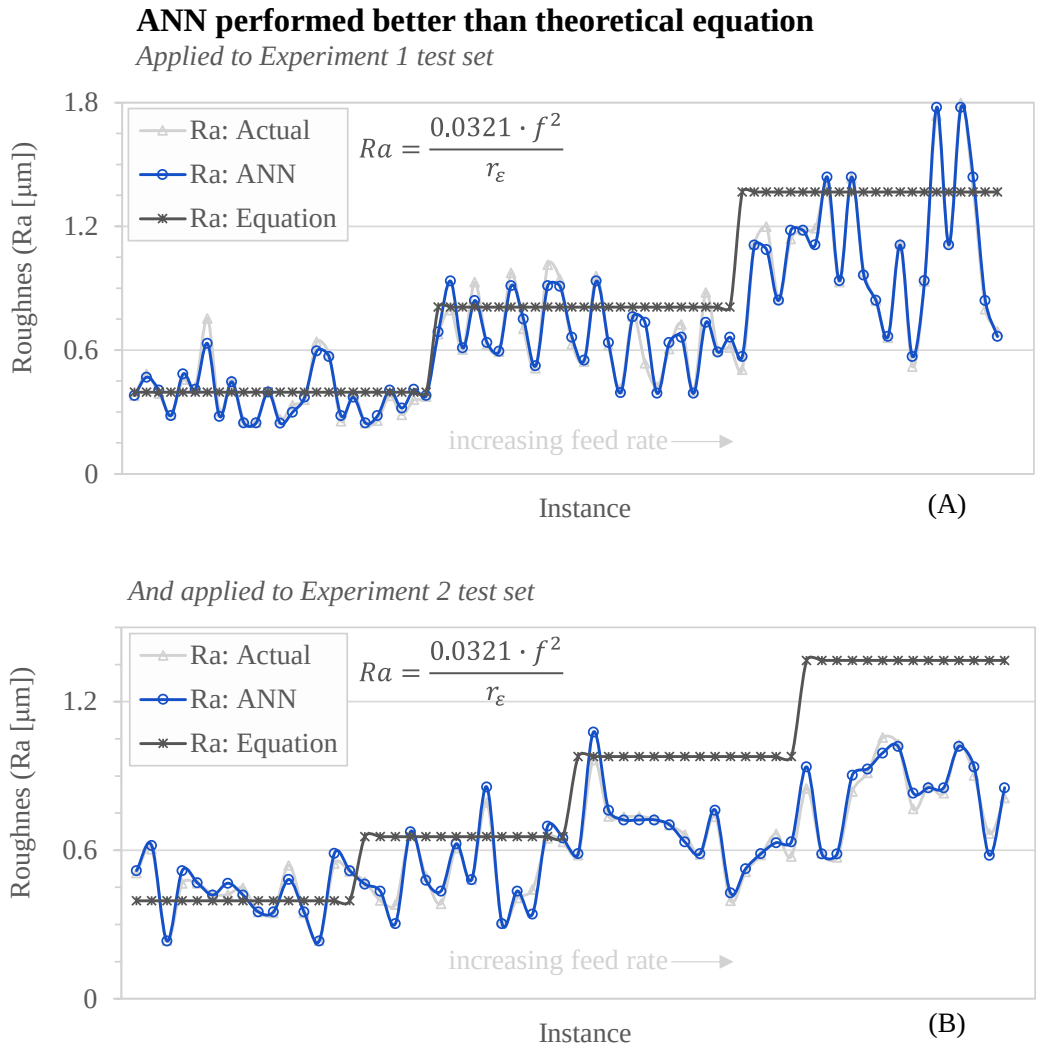



Figure 61. Comparison of Ra prediction made with artificial neural networks and theoretical equation in (A) Experiment 1 and (B) Experiment 2.

The falling-outside-range values predicted by the equation in Figure 61 (B) must be discussed in the light of tool wear presence. Experiment 2 reproduced the actual production situation where tool wear occurs, and such disturbance has made the theoretical equation much worse in predicting surface roughness values.

Apart from the mean errors, the R-squared metric was also calculated for the four models (ANN1.H.L, ANN2.H.L, equation applied in Experiment 1, and equation applied to Experiment 2). The R-squared for the theoretical equation applied to Experiment 1 test values is 0.34, while the R-squared for the equation applied to Experiment 2 is negative, concluding that the equation cannot model the surface roughness values in the test set of Experiment 2 at all.

Figure 62 shows the performance metrics of the four models. Two highlights must be given: ANN models have much better performance metrics, and the equation applied to Experiment 2 wear-existent values is much worse than that applied to Experiment 1 wear-free values.

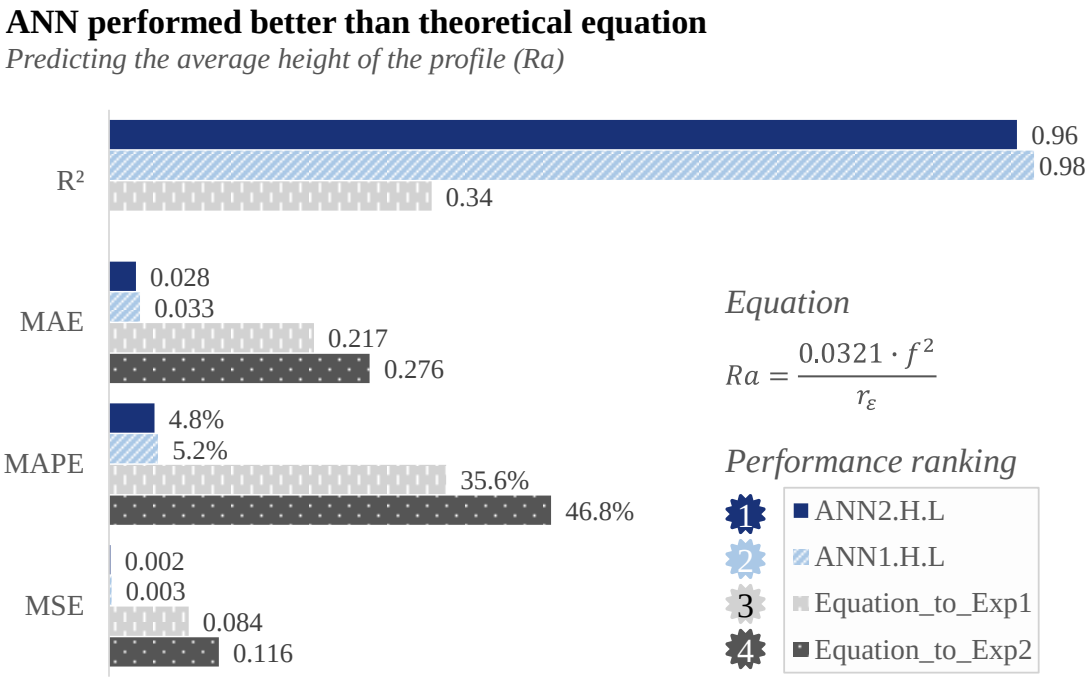



Figure 62. Performance metrics comparison between theoretical equations and ANN models.

## 5.3 Summarized results

Reviewing, in summary, the results presented in the previous topics before heading to the conclusions can be helpful. This topic does not present any new ideas, therefore serving only as a summary for easy reference of results.

### 5.3.1 Empirical machining results (answering Q1 and Q2)

The summarized results of machining experiments are as follows:

- In a scenario that does not consider tool wear—Experiment 1:
  - The surface roughness Ra was affected more significantly by the feed rate and the depth of cut, respectively;

  - The feed rate was the prime influencing factor for four of the five studied surface roughness parameters: Ra, Rsk, RSm, and Rt;
  - The machining forces were affected to a higher degree by the depth of cut and the feed rate, respectively.
- In a scenario that considers tool wear, represented by the flank wear width—Experiment 2:
  - The tool condition's main effect was reasonably low to Ra (0.6%), RSm (6.6%), and Rt (7.6%);
  - Tool condition main effect was high for Rsk (41.9%) and Rku (28.3%);
  - Tool condition interaction with the depth of cut and feed rate had a medium effect (up to 20%) on all the studied surface roughness parameters;
  - Tool edge wear has caused different surface roughness profiles with similar Ra and Rt;
  - Tool flank wear has decreased the influence of all factors on the surface roughness, especially the influence of the feed rate on RSm—the surface roughness parameter that was more dependent on the feed rate in Experiment 1;
  - Tool wear greatly influenced the machining forces, increasing all force components.

### 5.3.2 Computational (answering Q3–Q11)

The machine learning summarized results are as follows:

- Best ANN models predicted on the test set (never seen samples) with an error (MAPE) of around 5%;
- A strategy to multiply the dataset was proposed: surface roughness was measured six times in each test part. The strategy showed valuable results in training the neural networks contributing to avoiding overfitting and improving the prediction performance;
- On the one hand, the ANN models developed with theoretically cutting-tool–wear-free training data performed poorly in cutting-tool–wear-existent data. While on the other hand, the ANN developed with wear-existent training data

performed reasonably better in the theoretically wear-free data, indicating the importance of considering cutting tool wear to develop surface roughness' prediction models;

- The addition of machining forces as inputs improved the prediction performance of ANN models by 75% compared to using only the depth of cut, feed rate, and cutting speed
- ANNs predicted surface roughness much better than theoretical models. This result is even more evident when the cutting tool wear is considered

## 5.4 Considerations over ANN modeling

Here the author presents his point of view on machine learning model implementation.

A practical machine learning implementation would be more straightforward with the aid of specialized data scientists and data engineers than it would be for individuals willing to venture into machine learning for the first time—even when those individuals have some statistics background. In addition, the big picture of machine learning applications is that there are many more tasks than just modeling, as discussed in Figure 17. However, despite the numerous challenges and effort spent on all-around tasks, there seems to be an inflection point when the machine learning model is built and works for the first time. Few interactions and small changes on hyperparameters can be made from then on, but the main program would be ready to make predictions (not considering deployment).

In the case of the current dissertation, the author had no previous experience with *Python* programming—and very few previous experiences with programming at all—or machine learning methods, thus demanding much reading and trial-and-error practice. Despite coding not being where the most profound knowledge demand lies, much effort is needed from people with few or no programming skills. One cannot copy and paste the line codes from a dedicated book or article and expect it to perform the same for its problem with different datasets, software versions, and so on. For those fond of programming, the learning curve can be smoother (this is generally the case for every activity; when the learning journey is enjoyed, it tends to be faster).

A good point with *Python* is that it is a very popular programming language, which means that many people are struggling with similar problems, and one can find out that his very brand-new question was answered a couple of years ago on an online blog. There is indeed a

considerable movement of helping each other in those online communities. Moreover, the more popular a stack (programming language) is, the more is the number of people discussing it—hence, more helping topics.

In what matters the specific machine learning challenges, one newcomer will find it particularly tricky to select the best hyperparameters. To a more experienced practitioner, automated machine learning techniques are recommended—the reasons one should use them and how to use them can be further explored in the Automated Machine Learning open-access book (HUTTER; KOTTHOFF; VANSCHOREN, 2019). Moreover, the all-around tasks, such as preparing the data and scaling the features, demand much energy. Those tasks can be performed in actual production conditions with automated machine learning pipelines described in dedicated books such as the practical guide entitled Hands-On Machine Learning with scikit-learn & TensorFlow (GÉRON, 2017).

Therefore, apart from the challenges of learning how to program the software and understand and use a machine learning tool, the preciseness of how the methods can model such complex, non-linear factor-response relations makes it worth it. Also, in a scenario where the individuals already have the demanded skills, machine learning is even more advantageous and can be a powerful and efficient modeling tool.

# 6 Conclusion

This work had the objective of understanding the feasibility of predicting surface roughness—Ra—using artificial neural networks. Two machining experiments were planned to generate valuable data to train machine learning models.

The effect of the process parameters depth of cut, feed rate, and cutting speed was statistically assessed by analysis of variance, and the results follow what is presented in the literature. Feed rate was the factor with more influence on the surface roughness, followed by the depth of cut and the cutting speed. The little influence of the cutting speed was understood to be a consequence of the slight variation in this factor's levels. The tool wear effect on the surface roughness drew additional interest, given that it was expected that when increasing tool wear, the surface roughness Ra was increased or that the surface finish became worse in general. However, the tool wear increase contributed to the surface roughness Ra of only 0.6%. On the other hand, the increase in tool wear state positively correlated with all the machining forces components. This result indicates that it may be enough to control the cutting forces to avoid exceeding surface roughness limits, given that the forces start to rise before the surface roughness do.

The soft computing method for modeling the surface roughness parameter Ra showed excellent prediction capacity, with errors (MAPE) ranging from 4.8% to 22%, better than a classical theoretical equation. The best models were achieved using three process parameters and four machining forces as inputs. Moreover, the tests showed that the tool wear state significantly influenced the prediction performance of the ANN models. Also, the training set size influenced prediction performance. The soft computing modeling findings are pointed out as follows:

1. The strategy of collecting the surface roughness more times for the same test part can be made to increase the data samples, given that the use of larger training sets resulted in the best prediction performance and less tendency for overfitting;
2. Using the machining force components as inputs may improve the ANN model;
3. The tool wear state should be considered when modeling surface roughness (Ra) in turning processes with ANN;
4. ANN predicts better than a classical theoretical equation.

# 7 Suggestions for Future Works

The current dissertation is essentially a work dedicated to opening a research line inside a research group. The artificial neural network may sometimes be seen as a too complex tool to model surface roughness in machining processes—however, ANNs present advantages in future works where the exploration of online monitoring systems may occur.

The subsequent actions for future works are divided into:

- Further exploration of the ANN modeling—using the dataset generated in this work:
    1. Modeling more complex surface roughness parameters such as the Skewness and Kurtosis parameters—that were significantly affected by the cutting tool wear;
    2. Exploring the use of the cutting tool wear as a feature (input);
    3. Utilizing the available data to map different responses, such as the machining forces or the cutting tool wear.
    4. Replicate the models using cross-validation methods and investigate data multiplication's efficacy regarding data leakage—more on this in Appendix C.

- Exploration of different features, responses, or manufacturing processes—generating new experimental data:
    1. Collecting different features. Accelerometers tend to be more affordable and easier to install in production systems, and software to monitor the machine signals is already commercially available, making acceleration, vibration, and electric current attractive features to use;
    2. Modeling different responses, such as the cutting tool wear, by collecting data from more complete tool-life tests;
    3. Expand the methodology to other:
        3.1. Manufacturing processes, such as milling or grinding;
        3.2. Machines to explore their difference in terms of rigidity; or
        3.3. Materials, especially regarding their hardness.

# BIBLIOGRAPHY

ABBAS, A. *et al*. ANN Surface roughness optimization of AZ61 magnesium alloy finish turning: minimum machining times at prime machining costs. **Materials**, v. 11, n. 5, p. 808, May 2018.

ABUAF, N.; BUNKER, R. S.; LEE, C. P. Effects of surface roughness on heat transfer and aerodynamic performance of turbine airfoils. *In*: INTERNATIONAL GAS TURBINE AND AEROENGINE CONGRESS AND EXHIBITION, 1997, Orlando, FL. **Proceedings** [...]. New York: American Society of Mechanical Engineers, 1998.

AGGARWAL, C. C. **Neural networks and deep learning**. Cham: Springer International Publishing, 2018. 512 p.

ALBANESI, S.; VAMOSSY, D. **Predicting consumer default**: a deep learning approach. Cambridge, MA: National Bureau of Economic Research, 2019. 72 p. (NBER Working Paper Series).

ALPAYDIN, E. **Introduction to machine learning**. 3rd ed. Massachusetts, MA: MIT Press, 2014. 640 p. (Adaptative Computation and Machine Learning).

ALTINTAS, Y. **Manufacturing automation**. 2nd ed. Cambridge, UK: Cambridge University Press, 2012. (Robotics and Automation Handbook).

ANDREONI, A.; UPADHYAYA, S. **Growth and distribution pattern of the world manufacturing output**: a statistical profile. Vienna: UNIDO, 2014. 101 p. (Research, Statistics and Industrial Policy Branch).

ASILTÜRK, İ. Predicting surface roughness of hardened AISI 1040 based on cutting parameters using neural networks and multiple regression. **The International Journal of Advanced Manufacturing Technology**, v. 63, n. 1–4, p. 249–257, Nov. 2012.

BAYER, R. G.; SIRICO, J. L. The influence of surface roughness on wear. **Wear**, v. 35, n. 2, p. 251–260, Dec. 1975.

BEATRICE, B. A. *et al*. Surface roughness prediction using artificial neural network in hard turning of AISI H13 steel with minimal cutting fluid application. **Procedia Engineering**, v. 97, p. 205–211, 2014.

BEDDOES, J.; BIBBY, M. J. **Principles of metal manufacturing processes**. Carleton: Elsevier, 1999.

BENARDOS, P. G.; VOSNIAKOS, G. C. Predicting surface roughness in machining: a review. **International Journal of Machine Tools and Manufacture**, v. 43, n. 8, p. 833–844, 2003.

BHADESHIA, H. K. D. H. Steels for bearings. **Progress in Materials Science**, v. 57, n. 2, p. 268–435, Feb. 2012.

BOOTHROYD, G.; KNIGHT, W. A. **Fundamentals of machining and Machine Tools**. 3rd ed. Boca Raton, FL: Taylor & Francis, 2006.

BOUZAKIS, K.-D. *et al*. Coated tools' wear description in down and up milling based on the cutting edge entry impact duration. **CIRP Annals**, v. 61, n. 1, p. 115–118, 2012a.

BOUZAKIS, K. D. *et al*. Effect of tool diameter and cutting edge entry impact duration on coated tool wear in milling of various kinematics. **Journal of the Balkan Tribological Association**, v. 18, n. 2, p. 238–249, 2012b.

BRINKSMEIER, E. *et al*. Process signatures: a new approach to solve the inverse surface integrity problem in machining processes. **Procedia CIRP**, v. 13, p. 429–434, 2014.

ÇAYDAŞ, U.; EKICI, S. Support vector machines models for surface roughness prediction in CNC turning of AISI 304 austenitic stainless steel. **Journal of Intelligent Manufacturing**, v. 23, n. 3, p. 639–650, June 2012.

CHEN, Y. *et al*. A nested-ANN prediction model for surface roughness considering the effects of cutting forces and tool vibrations. **Measurement**, v. 98, p. 25–34, Feb. 2017.

DAVIM, J. P. (ed.). **Surface integrity in machining**. London: Springer, 2010.

DAVIM, J. P. (ed.). **Traditional machining processes**. Berlin: Springer-Verlag, 2015.

DEKEL, O. From online to batch learning with cutoff-averaging. *In*: INTERNATIONAL CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 21., 2008, Vancouver. **Proceedings** […]. Red Hook, NY: Curran Associates, 2008. p. 1769–1776.

DEUTSCHES INSTITUT FÜR NORMUNG. **DIN 4760**: form deviations, concepts, classification system. Englewood, CO: IHS, 1982.

DOMENECH, T.; FOKEER, S. **Why innovative manufacturing and circularity are key for a resilient manufacturing industry post-COVID-19**. Vienna: UNIDO, 2020. Available at: https://www.unido.org/news/why-innovative-manufacturing-and-circularity-are-key-resilient-manufacturing-industry-post-covid-19. Accessed on: 19 Apr. 2022.

DONOHO, D. 50 years of data science. **Journal of Computational and Graphical Statistics**, v. 26, n. 4, p. 745–766, Oct. 2017.

EBERHART, R. C.; DOBBINS, R. W. (ed.). **Neural network PC tools**: a practical guide. San Diego: Academic Press, 1990. 414 p.

ERTEL, W. **Introduction to artificial intelligence**. 2nd ed. Cham: Springer International Publishing, 2017. 365 p. (Undergraduate Topics in Computer Science).

FANG, F. Z. *et al*. Manufacturing and measurement of freeform optics. **CIRP Annals**, v. 62, n. 2, p. 823–846, 2013.

GADELMAWLA, E. S. *et al*. Roughness parameters. **Journal of Materials Processing**

**Technology**, v. 123, p. 133–145, 2002.

GARG, A.; TAI, K. Stepwise approach for the evolution of generalized genetic programming model in prediction of surface finish of the turning process. **Advances in Engineering Software**, v. 78, p. 16–27, Dec. 2014.

GARTNER. **Emerging technology roadmap for large enterprises**. Stamford, CT: Gartner, 2021.

GÉRON, A. **Hands-on machine learning with scikit-learn & tensorflow**. Sebastopol: O'Reilly Media, 2017. 564 p.

GÉRON, A. **Mãos à obra:** aprendizado de máquina com scikit-learn, keras & tensorflow. 2nd ed. Rio de Janeiro: Alta Books, 2021. 577 p.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. Massachusetts, MA: MIT Press, 2016. 800 p.

HAYKIN, S. **Neural networks**: a comprehensive foundation. 2nd ed. Upper Saddle River: Prentice Hall, 1999. 823 p.

HE, C. L.; ZONG, W. J.; ZHANG, J. J. Influencing factors and theoretical modeling methods of surface roughness in turning process: state-of-the-art. **International Journal of Machine Tools and Manufacture**, v. 129, p. 15–26, Feb. 2018.

HUTTER, F.; KOTTHOFF, L.; VANSCHOREN, J. (ed.). **Automated machine learning**. Cham: Springer International Publishing, 2019. 222 p. (The Springer Series on Challenges in Machine Learning).

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 3685(E):1993**: tool life testing with single point turning tools. Vernier, Geneva: ISO, 1993.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 4288:1996**: rules and procedures for assessment of surface texture. Vernier, Geneva: ISO, 1996.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 4287:1997**: terms, definition and surface texture parameters. Vernier, Geneva: ISO, 1997.

JAFARIAN, F.; TAGHIPOUR, M.; AMIRABADI, H. Application of artificial neural network and optimization algorithms for optimizing surface roughness, tool life and cutting forces in turning operation. **Journal of Mechanical Science and Technology**, v. 27, n. 5, p. 1469–1477, May 2013.

JAIN, V. Everything you need to know about "activation functions" in deep learning models. **Towards Data Science**, 30 Dec. 2019. Available at: https://towardsdatascience.com/everything-you-need-to-know-about-activation-functions-in-deep-learning-models-84ba9f82c253. Accessed on: 09 Apr. 2022.

JAMWAL, A. *et al*. Industry 4.0: an indian perspective. *In*: DOLGUI, A. *et al*. (ed.). **Advances in production management systems**. Artificial intelligence for sustainable and

resilient production systems. Cham: Springer International, 2021. v. 630.
JURKOVIC, Z. *et al*. A comparison of machine learning methods for cutting parameters prediction in high speed turning process. **Journal of Intelligent Manufacturing**, v. 29, n. 8, p. 1683–1693, Dec. 2018.

KARA, F. *et al*. Effect of machinability, microstructure and hardness of deep cryogenic treatment in hard turning of AISI D2 steel with ceramic cutting. **Journal of Materials Research and Technology**, v. 9, n. 1, p. 969–983, Jan. 2020.

KARPATHY, A. **Yes you should understand backprop**. [*S.l.*]: Medium, 2016. Available at: https://karpathy.medium.com/yes-you-should-understand-backprop-e2f06eab496b. Accessed on: 10 Apr. 2022.

KASABOV, N. K. **Foundations of neural networks, fuzzy systems, and knowledge engineering**. Cambridge, MA: MIT Press, 1996. 581 p.

KLOCKE, F. **Manufacturing processes 1**. Berlin, Heidelberg: Springer, 2011. 524 p.

KOSTADINOV, S. Understanding backpropagation algorithm. **Towards Data Science**, 8 Aug. 2019. Available at: https://towardsdatascience.com/understanding-backpropagation-algorithm-7bb3aa2f95fd. Accessed on: 10 Apr. 2022.

KUBAT, M. **An introduction to machine learning**. 2nd ed. Cham: Springer International Publishing, 2017. 348 p.

LAOUISSI, A. *et al*. Investigation, modeling, and optimization of cutting parameters in turning of gray cast iron using coated and uncoated silicon nitride ceramic tools. Based on ANN, RSM, and GA optimization. **The International Journal of Advanced Manufacturing Technology**, v. 101, n. 1–4, p. 523–548, Mar. 2019.

LECUN, Y. A. *et al*. Efficient BackProp. *In*: MONTAVON, G.; ORR, G. B.; MÜLLER, K.-R. (ed.). **Neural networks**: tricks of the trade. 2nd ed. Berlin, Heidelberg: Springer, 2012. p. 9–48.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, v. 521, n. 7553, p. 436–444, May 2015.

LENAIL, A. **NN SVG:** publication-ready NN-architecture schematics. [*S.l.: s.n.*], [2019]. Available at: http://alexlenail.me/NN-SVG/index.html. Accessed on: 17 Mar 2022a.

LENAIL, A. NN-SVG: Publication-ready neural network architecture schematics. **Journal of Open Source Software**, v. 4, n. 33, p. 747, Jan. 2019b.

LEUNG, K. The dying ReLU problem, clearly explained. **Towards Data Science**, 30 Mar. 2021. Available at: https://towardsdatascience.com/the-dying-relu-problem-clearly-explained-42d0c54e0d24#0ebe. Accessed on: 10 Apr. 2022.

LU, L. *et al*. Dying ReLU and initialization: theory and numerical examples. **Communications in Computational Physics**, v. 28, n. 5, p. 1671–1706, Nov. 2020.

LUND, H. B.; VILDÅSEN, S. S. The influence of Industry 4.0 narratives on regional path development. **Regional Studies, Regional Science**, v. 9, n. 1, p. 82–92, Dec. 2022.

MACHADO, A. R. *et al*. **Teoria da usinagem dos materiais**. São Paulo: Blucher, 2009. 371 p.

MEDDOUR, I. *et al*. Prediction of surface roughness and cutting forces using RSM, ANN, and NSGA-II in finish turning of AISI 4140 hardened steel with mixed ceramic tool. **The International Journal of Advanced Manufacturing Technology**, v. 97, n. 5–8, p. 1931–1949, July 2018.

MIA, M. *et al*. Effect of time-controlled MQL pulsing on surface roughness in hard turning by statistical analysis and artificial neural network. **The International Journal of Advanced Manufacturing Technology**, v. 91, n. 9–12, p. 3211–3223, Aug. 2017.

MIA, M. *et al*. Prediction and optimization of surface roughness in minimum quantity coolant lubrication applied turning of high hardness steel. **Measurement**, v. 118, p. 43–51, Mar. 2018.

MIA, M.; DHAR, N. R. Prediction of surface roughness in hard turning under high pressure coolant using Artificial Neural Network. **Measurement**, v. 92, p. 464–474, Oct. 2016.

MITUTOYO. **Quick guide to surface roughness measurement**: reference guide for laboratory and workshop Mitutoyo America Corporation. Aurora, IL: Mitutoyo America Corporation. Available at: http://www.mitutoyo.com/wp-content/uploads/2012/11/1984_Surf_Roughness_PG.pdf. Accessed on: 08 Aug. 2021.

MIWA, M. *et al*. Effects of the surface roughness on sliding angles of water droplets on superhydrophobic surfaces. **Langmuir**, v. 16, n. 13, p. 5754–5760, June 2000.

MODA, L. B.; MUNDIM, R. B.; BORILLE, A. V. Projeto e validação de dispositivo de fixação para dinamômetro piezelétrico em torno CNC de torre rotativa. *In*: CONGRESSO BRASILEIRO DE ENGENHARIA DE FABRICAÇÃO, 9., 2017, Joinville, SC. **Proceedings** [...]. Rio de Janeiro, RJ: ABCM, 2017.

MOLL, I. The myth of the Fourth Industrial Revolution. **Theoria**, v. 68, n. 2, p. 167, June 2021.

MONTAVON, G.; ORR, G. B.; MÜLLER, K.-R. (ed.). **Neural networks**: tricks of the trade. 2nd ed. Berlin, Heidelberg: Springer, 2012. 753 p. (Lecture Notes in Computer Science, v. 7700).

MONTGOMERY, D. C. **Design and analysis of experiments**. 8th ed. New York: John Wiley & Sons, 2013.

MUNDIM, R. B. **Evaluation of strategies for milling thin-walled aluminium components**. 2014. Dissertation (Master of Science Dissertation in Manufacturing Technology) – Instituto Tecnológico de Aeronáutica. São José dos Campos, 2014.

MUNDIM, R. B. **Cutting tool and process parameters selection in milling of thin-walled**

**aluminum structures**. 2018. Thesis (PhD in Materials, Manufacturing and Automation) – Instituto Tecnológico de Aeronáutica, São José dos Campos, 2018.

MUNDIM, R. B.; LUCAS, E. O. Effect of tool wear on surface topography in finish turning of ABNT 1045 steel. *In*: BRAZILIAN CONGRESS OF MECHANICAL ENGINEERING, 21, 2011, Natal, RN. **Proceedings** [...]. Natal, RN: ABCM, 2011.

MUTHUKRISHNAN, N.; DAVIM, J. P. Optimization of machining parameters of Al/SiC-MMC with ANOVA and ANN analysis. **Journal of Materials Processing Technology**, v. 209, n. 1, p. 225–232, 2009.

NAIR, V.; HINTON, G. E. Rectified linear units improve restricted boltzmann machines. *In*: INTERNATIONAL CONFERENCE ON MACHINE LEARNING, 27., 2010, Haifa, Israel. **Proceedings** […]. Madison, WI: Omnipress, 2010. p. 807–814.

NASCIMENTO JÚNIOR, C. L.; YONEYAMA, T. **Inteligência artificial em controle e automação**. São Paulo: Blucher, 2000.

NIELSEN, M. A. **Neural networks and deep learning**. [*S.l.*]: Determination Press, 2015. 224 p.

NWANKPA, C. *et al*. Activation functions: comparison of trends in practice and research for deep learning. *In*: INTERNATIONAL CONFERENCE ON COMPUTATIONAL SCIENCES AND TECHNOLOGIES, 2., 2020, Jamshoro, Pakistan. **Proceedings** […]. Glasgow: University of Strathclyde, 2020.

PLUMMER, T. Flaked stones and old bones: biological and cultural evolution at the dawn of technology. **American Journal of Physical Anthropology**, v. 125, n. S39, p. 118–164, 2004.

PONTES, F. J. *et al*. Optimization of radial basis function neural network employed for prediction of surface roughness in hard turning process using Taguchi's orthogonal arrays. **Expert Systems with Applications**, v. 39, n. 9, p. 7776–7787, July 2012.

REGO, R. **Topography**. São José dos Campos: Instituto Tecnológico de Aeronáutica, 2020. 111 p. Material from the course MT-204: Integridade de Superfícies.

REHORN, A. G.; JIANG, J.; ORBAN, P. E. State-of-the-art methods and results in tool condition monitoring: a review. **The International Journal of Advanced Manufacturing Technology**, v. 26, n. 7–8, p. 693–710, Oct. 2005.

RICH, E. **Artificial Intelligence**. New York: McGraw-Hill, 1983.

ROSER, C. **Faster, better, cheaper in the history of manufacturing**. Boca Raton, FL: CRC Press, 2016.

ROZENBERG, G.; BÄCK, T.; KOK, J. N. (ed.). **Handbook of Natural Computing**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. 2105 p. 4 v.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **Nature**, v. 323, n. 6088, p. 533–536, Oct. 1986.

RUSSEL, S. J.; NORVIG, P. **Artificial Intelligence**. 3rd ed. New Jersey: Prentice Hall, 2010. 1132 p.

SARIC, T.; SIMUNOVIC, G.; SIMUNOVIC, K. Use of Neural Networks in prediction and simulation of steel surface roughness. **International Journal of Simulation Modelling**, v. 12, n. 4, p. 225–236, Dec. 2013.

SCHUH, G. *et al*. (ed.). **Industry 4.0 maturity index**. Aachen: Acatech, 2017. 60 p. (Acatech Study).

SCULLEY, D. *et al*. Hidden technical debt in machine learning systems. *In*: CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 29., 2015, Montréal, Canada. **Proceedings** […]. New York: ACM, 2015. p. 2494–2502.

SENTHILKUMAAR, J. S.; SELVARANI, P.; ARUNACHALAM, R. M. Intelligent optimization and selection of machining parameters in finish turning and facing of Inconel 718. **The International Journal of Advanced Manufacturing Technology**, v. 58, n. 9–12, p. 885–894, Feb. 2012.

SHALEV-SHWARTZ, S.; BEN-DAVID, S. **Understanding machine learning**: from Theory to Algorithms. Cambridge, MA: Cambridge University Press, 2014. 449 p.

SHANMUGANATHAN, S.; SAMARASINGHE, S. (ed.). **Artificial Neural Network Modelling**. Cham: Springer International Publishing, 2016. 468 p. (Studies in Computational Intelligence, v. 628).

SHARMA, A. Understanding activation functions in Neural Networks. **The Theory of Everything**, [*S.l.*], 30 Mar. 2017. Available at: https://medium.com/the-theory-of-everything/understanding-activation-functions-in-neural-networks-9491262884e0. Accessed on: 09 Apr 2022.

SHARMA, S.; SHARMA, S.; ATHAIYA, A. Activtion functions in neural networks. **International Journal of Engineering Applied Sciences and Technology**, v. 4, n. 12, p. 310–316, 2020.

SHI, D.; GINDY, N. N. Tool wear predictive model based on least squares support vector machines. **Mechanical Systems and Signal Processing**, v. 21, n. 4, p. 1799–1814, May 2007.

SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes neurais artificiais para Engenharia e Ciências Aplicadas**. 2nd ed. São Paulo: Artliber, 2016. 431 p.

SKANSI, S. **Introduction to deep learning**: from logical calculus to Artificial Intelligence. Cham: Springer International Publishing, 2018. 196 p. (Undergraduate Topics in Computer Science).

SMITH, G. T. **Cutting tool technology**: industrial handbook. London: Springer, 2008. 605 p.

SMITH, L. N. **A disciplined approach to neural network hyper-parameters**: part 1:

learning rate, batch size, momentum, and weight decay. Washington, DC: US Naval Research Laboratory, 2018. 21 p. (Technical Report 5510-026).

SONY, M.; MEKOTH, N. Employee adaptability skills for Industry 4.0 success: a road map. **Production & Manufacturing Research**, v. 10, n. 1, p. 24–41, Dec. 2022.

STEMMER, C. E. **Ferramentas de corte I**. 3rd ed. Florianópolis, SC: Editora da UFSC, 1993. 249 p.

SWIFT, K. G.; BOOKER, J. D. **Manufacturing process selection handbook**. Oxford, MA: Elsevier, 2013.

TAYLOR, M. **Neural Networks**: a visual introduction for begginers. Canada: Blue Windmill Media, 2017.

TOMOV, M.; KUZINOVSKI, M.; CICHOSZ, P. Development of mathematical models for surface roughness parameter prediction in turning depending on the process condition. **International Journal of Mechanical Sciences**, v. 113, p. 120–132, July 2016.

TRENT, E.; WRIGHT, P. **Metal cutting**. 4th ed. Woburn, MA: Butterworth-Heinemann, 2000. 464 p.

TSCHÄTSCH, H. **Applied machining technology**. Berlin: Springer, 2009. 375 p.

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANIZATION. **Industrial statistical data and metadada**. Wien: UNIDO, [2020]. Available at: https://stat.unido.org/database/National%20Accounts%20Database. Accessed on: 15 Apr. 2022a.

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANIZATION. **What is manufacturing value added?** Wien: UNIDO, [2020?]. Available at: https://stat.unido.org/content/learning-center/what-is-manufacturing-value-added%253f. Accessed on: 15 Apr. 2022b.

UNITED STATES. The White House. **The Biden-Harris Plan to Revitalize American manufacturing and secure critical supply chains in 2022**. Washington, DC, 24 Feb. 2022. Available at: https://www.whitehouse.gov/briefing-room/statements-releases/2022/02/24/the-biden-harris-plan-to-revitalize-american-manufacturing-and-secure-critical-supply-chains-in-2022/. Accessed on: 19 Apr. 2022.

URBAN, S. **Neural Network architectures and activation functions**: a Gaussian Process Approach. 2017. Thesis (PhD in Naturwissenschaften) – Fakultät für Informatik, Technische Universtät München, Munich, 2018.

VILLARES METALS. **VH13ISO**. São Paulo: Villares Metals, 2006. 2 p. (Aços para trabalho a quente).

WORLD BANK GROUP. **Resilient industries**: competitiveness in the face of disasters. Washington, DC: World Bank, 2020.

WORLD MANUFACTURING FOUNDATION. **Manufacturing beyond COVID-19**: key findings for a resilient manufacturing sector in the new normal. Cernobbio, Lake Como: World Manufacturing Foundation, 2020. 20 p. (Back to the Future).

XU, L. DA; XU, E. L.; LI, L. Industry 4.0: state of the art and future trends. **International Journal of Production Research**, v. 56, n. 8, p. 2941–2962, Apr. 2018.

YANG, C.; TARTAGLINO, U.; PERSSON, B. N. J. Influence of surface roughness on superhydrophobicity. **Physical Review Letters**, v. 97, n. 11, p. 116103, Sept. 2006.

YOUSSEF, H. A.; EL-HOFY, H. **Machining technology**: machine tools and operations. Boca Raton, FL: CRC Press, 2008.

# Appendix A – Additional Information

This section comprises further information about the workpiece's material and equipment used in this work.

## A.1 Machine tool

The machine tool used was the *ROMI* CNC turning center model E 280 (Figure 63). It has a maximum nominal power of 18.5 kW and a maximum speed of 4000 rpm.

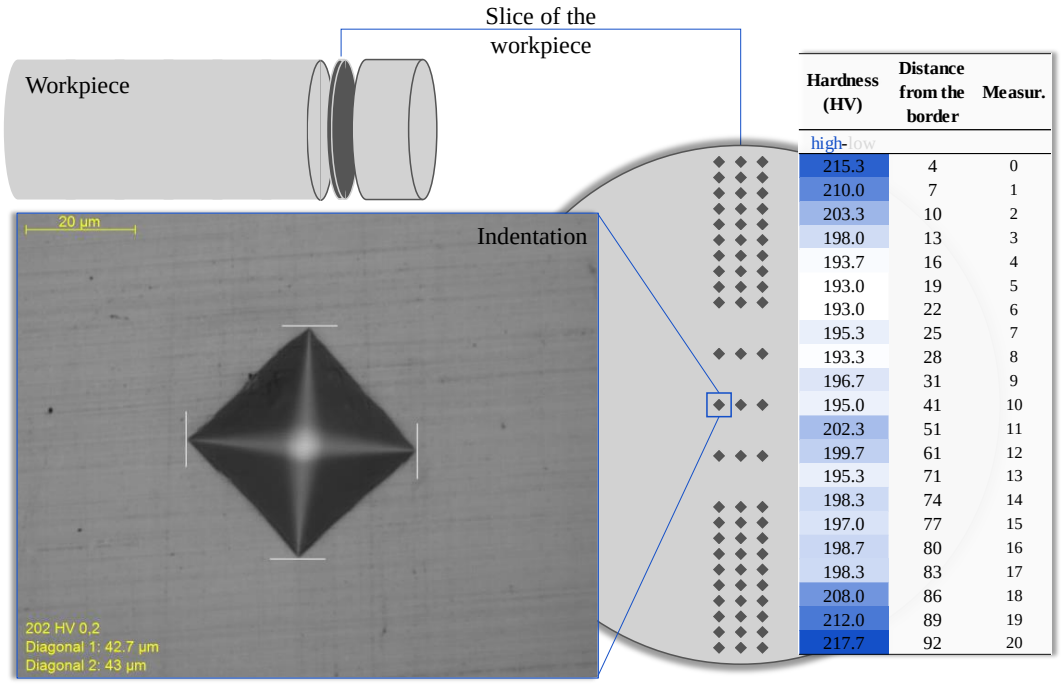

Figure 63. *ROMI* CNC turning center model E 280.

## A.2 Cutting fluid

The machining experiments: Exp1, Exp2, and the Preparation Phase, all used the same cutting fluid: a mixture of *Blaser Swisslube* Vasco 7000 with water in 8%. The acidity (pH) of the mixture was around 8.

## A.3 Workpiece's hardness

The roughness can be correlated to the part's hardness (HE; ZONG; ZHANG, 2018). Given this correlation, the hardness was measured as depicted in Figure 64. The table on the figure's right side indicates the measured hardness values. Each line is the mean of three consecutive measurements. The mean roughness measured was 200.7 HV, with a standard deviation of circa 7 HV and amplitude of 24.7 HV, being the greater values on the border of the part. As the machining happened inside a range of 15 mm, usually 4 mm away from the border, the considered hardness for the machining experiments had the following statistical values (using the measurement samples {1 2 3 4 5 6 14 15 16 17 18 19} from column "Measur."): Mean = 200.3 HV, Amplitude = 19.0 HV, and Standard deviation ≈ 6.3 HV.




| Hardness (HV) | Distance from the border | Measur. |
|---|---|---|
| high- | | -low |
| 215.3 | 4 | 0 |
| 210.0 | 7 | 1 |
| 203.3 | 10 | 2 |
| 198.0 | 13 | 3 |
| 193.7 | 16 | 4 |
| 193.0 | 19 | 5 |
| 193.0 | 22 | 6 |
| 195.3 | 25 | 7 |
| 193.3 | 28 | 8 |
| 196.7 | 31 | 9 |
| 195.0 | 41 | 10 |
| 202.3 | 51 | 11 |
| 199.7 | 61 | 12 |
| 195.3 | 71 | 13 |
| 198.3 | 74 | 14 |
| 197.0 | 77 | 15 |
| 198.7 | 80 | 16 |
| 198.3 | 83 | 17 |
| 208.0 | 86 | 18 |
| 212.0 | 89 | 19 |
| 217.7 | 92 | 20 |

Figure 64. Hardness tests on the workpiece material (AISI H13).

## A.4 Dynamometer adapted fixture plate

The dynamometer used in this work was built initially to work on horizontal platforms. However, a fixture plate (Figure 65) was developed by Moda, Mundim, and Borille (2017) explicitly to enable its use in the CNC turning center model E 280, taking into consideration critical requirements to force measurement tests and developments in turning investigations, such as surface finish, vibrations, rigidity, weight, and dislocation, considered sufficient to cutting forces higher as 2500 N.

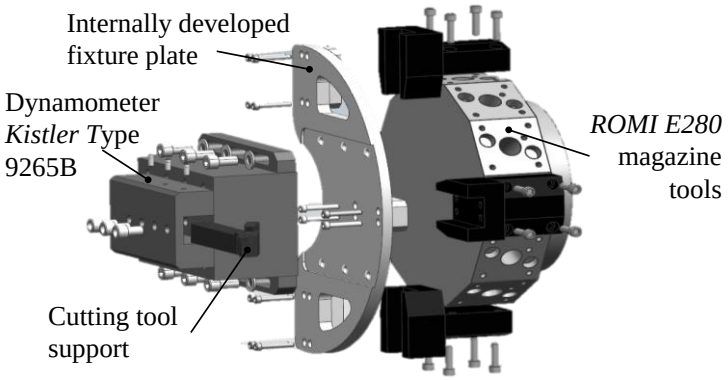



Figure 65. Fixture plate scheme (MODA; MUNDIM; BORILLE, 2017).

## A.5 Workpiece circular run-out state

A large workpiece run-out can significantly change the actual depth of cut during the turning process, which would affect the machining area, chip's formation, cutting forces, and quantity of material removed, among others. Therefore, the run-out is significant to the experiment's correctness and was assessed before the machining.

The workpiece's circular run-out was measured in relation to the tailstock. The mean value ranged around 17 µm (Figure 66), which is far below the smallest depth of cut—250 µm—and though acceptable with virtually no interference.

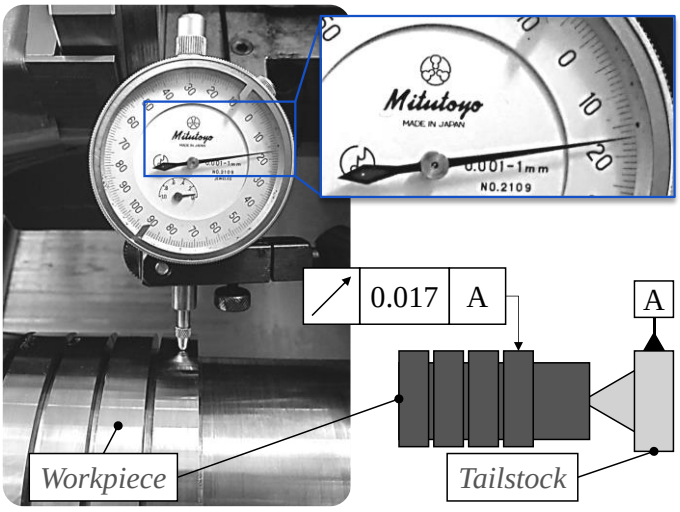



Figure 66. Run-out measurement.

# Appendix B – Additional Statistical Results

## B.1 Diagnosis for the factorial ANOVAs

The diagnostic graphs are presented in Figure 67, Figure 68, Figure 69, and Figure 70, with whom it is possible to analyze if the ANOVA model fits the data well. From them, one could conclude that:

- Residuals do not have non-linear patterns (Residuals vs. Fitted shows roughly horizontal red lines);
- Residuals are normally distributed (Normal Q-Q shows the residuals lined well on the straight line);
- The model fits the assumption of homoscedasticity (Scale-Location shows roughly horizontal red lines and the residuals equally randomly spread);
- There is no systematic contribution of the independent factors (Residuals vs. Factor Levels shows a random scatter);
- Lastly (but not visually represented): there is no influential outlier (Residuals vs. Leverage shows the points spread inside the cook's distance limits lines).

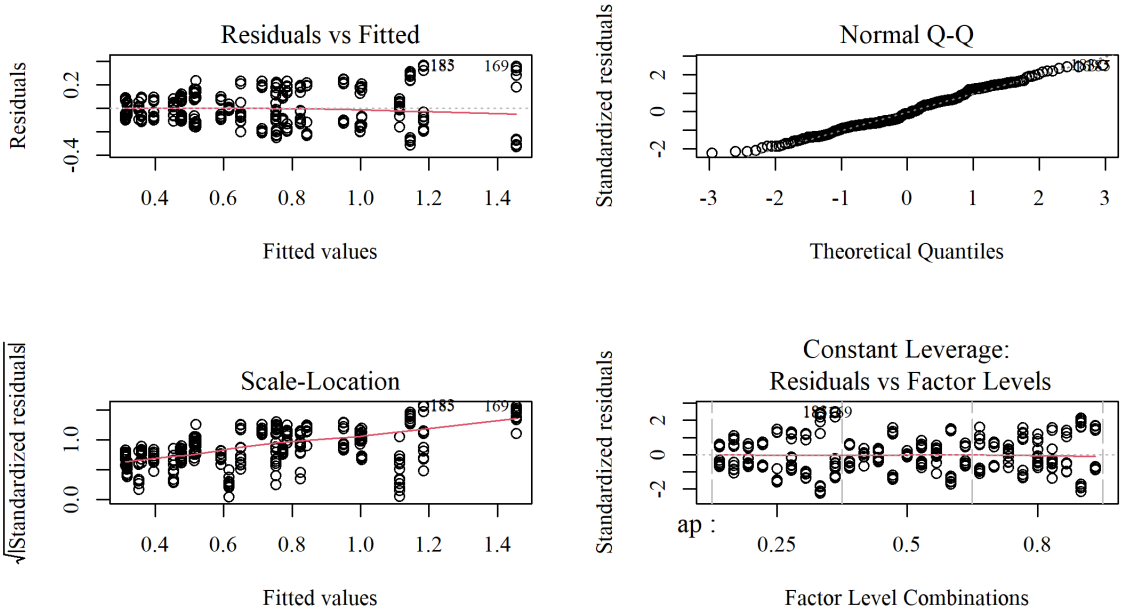



Figure 67. Residuals' diagnosis for the factorial ANOVA investigation of the factors influencing the average roughness Ra in Experiment 1.

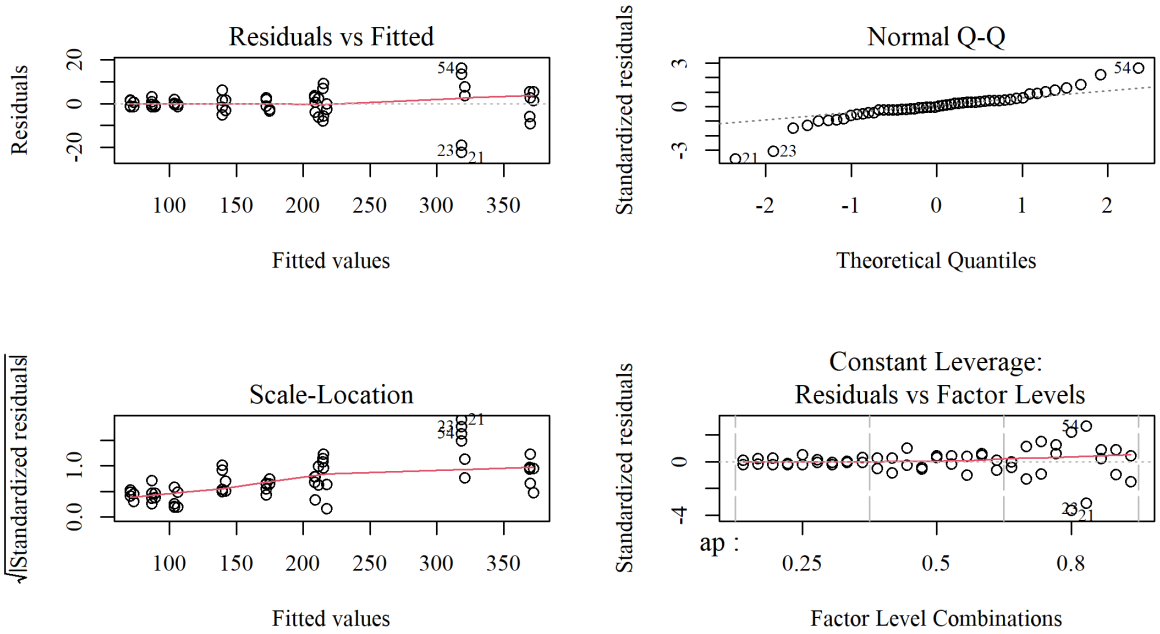



Figure 68. Residuals' diagnosis for the factorial ANOVA investigation of the factors influencing the resultant force in Experiment 1.

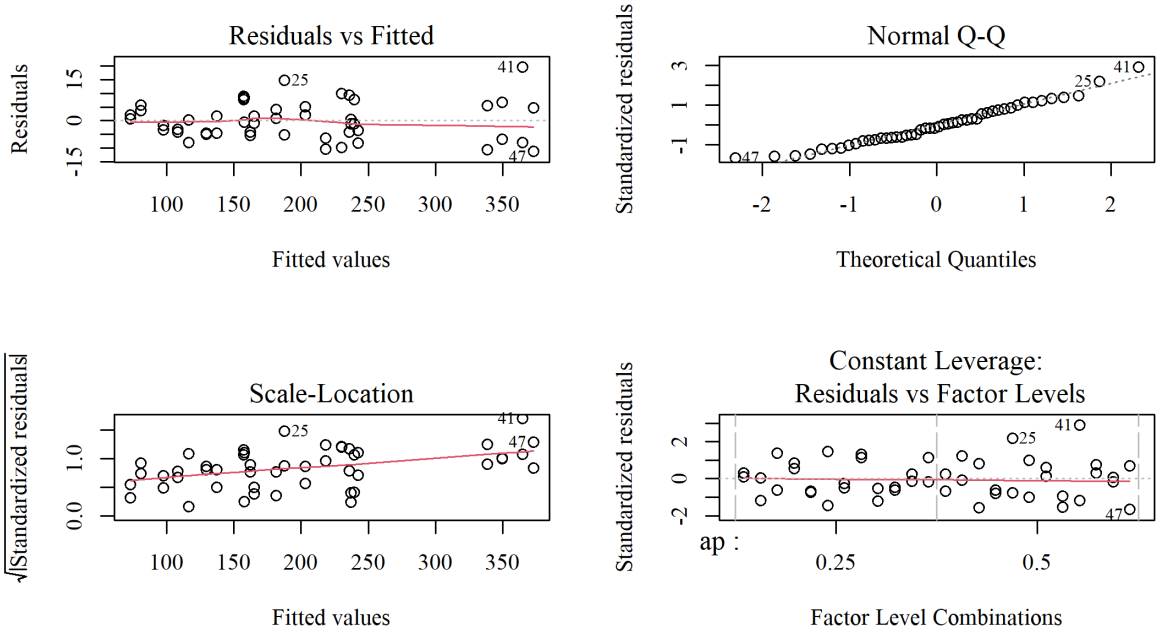



Figure 69. Residuals' diagnosis for the factorial ANOVA investigation of the factors influencing the average roughness Ra in Experiment 2.

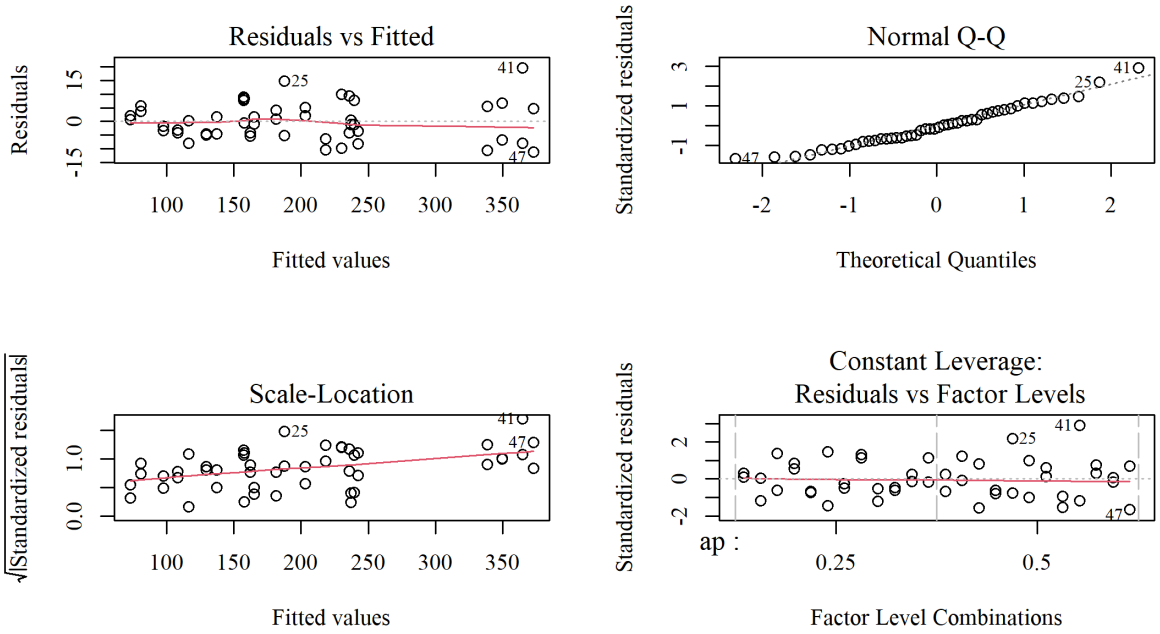



Figure 70. Residuals' diagnosis for the factorial ANOVA investigation of the factors influencing the resultant force in Experiment 2.

## B.2 Additional main effects plots and interaction plots for Experiment 1

The influence of the factors on the average roughness Ra and resultant force F was already discussed in the results section. However, additional information is depicted[12] in the main effects plots of Figure 71 and Figure 72. The only factor that has a diverging effect within the different roughness parameters is the depth of cut ($a_p$). It had a negative correlation with Rsk.

The force components were affected by the factors in the same direction, and the depth of cut and the feed rate had slight influencing differences. The change in $F_z$ is steeper when changing from the second to the third level of ap, while for $F_y$, it is steeper when changing from the first to the second level of $a_p$.

[12] The style of the graphs generated in Minitab are more complex to adjust. In the figures of appendix B.2 and B.3, comma is used as the decimal separator.

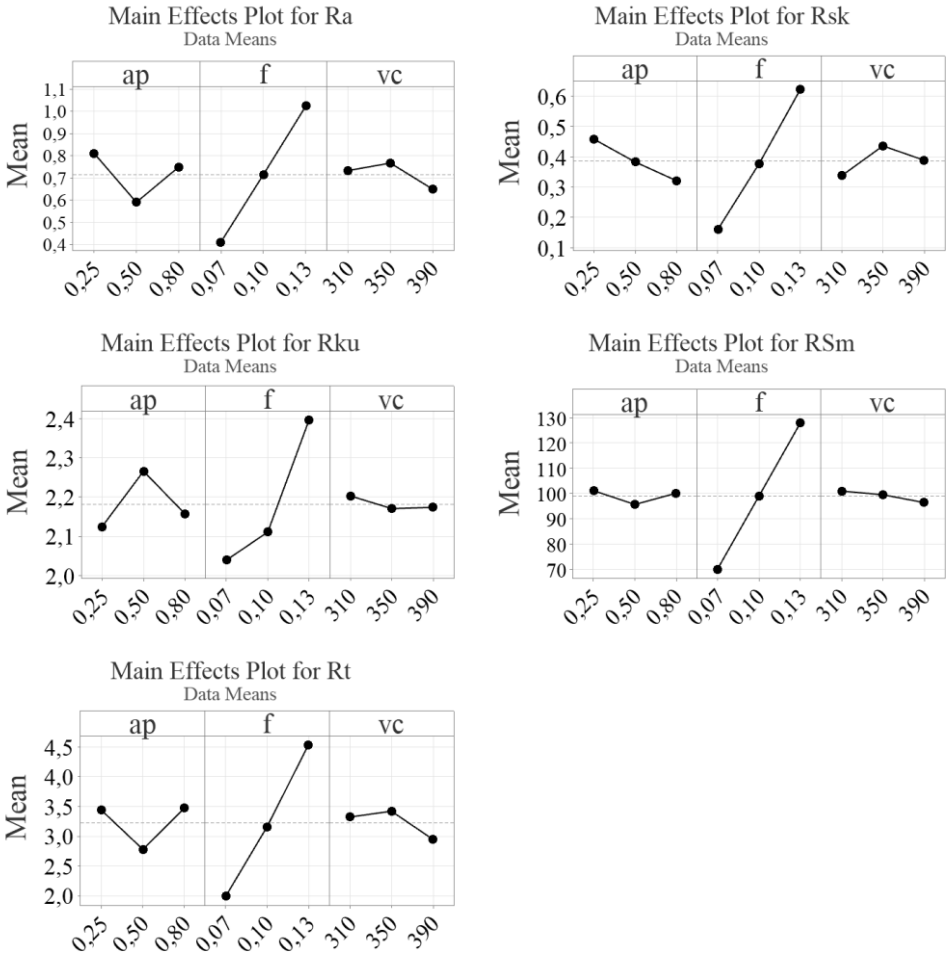



Figure 71. Main effects plot of the roughness components in Experiment 1.

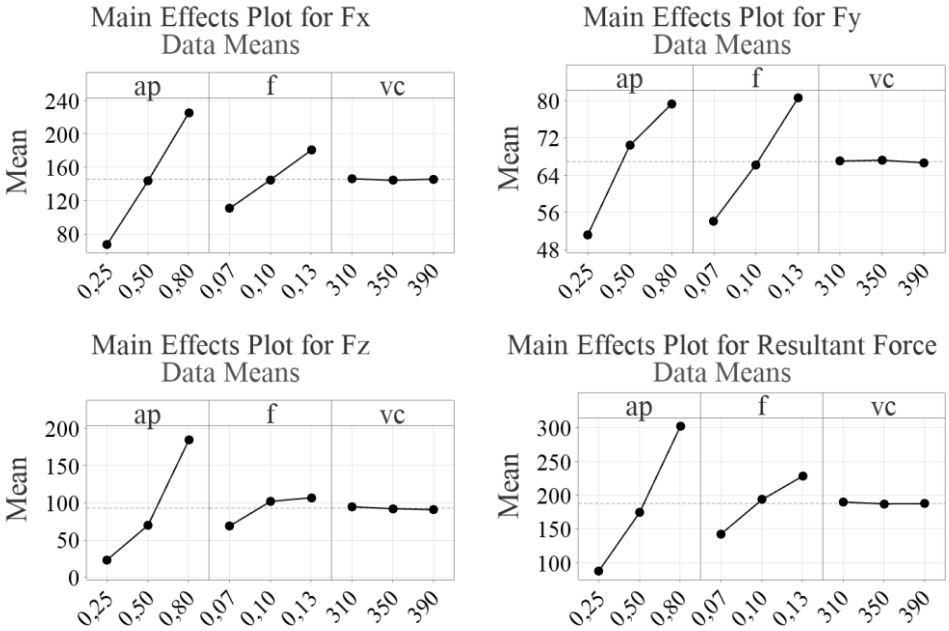



Figure 72. Main effects plot for the force components in Experiment 1.

The interaction plots for the roughness parameters, force components, and the processes parameters in Experiment 1 are depicted in Figure 73 and Figure 74.

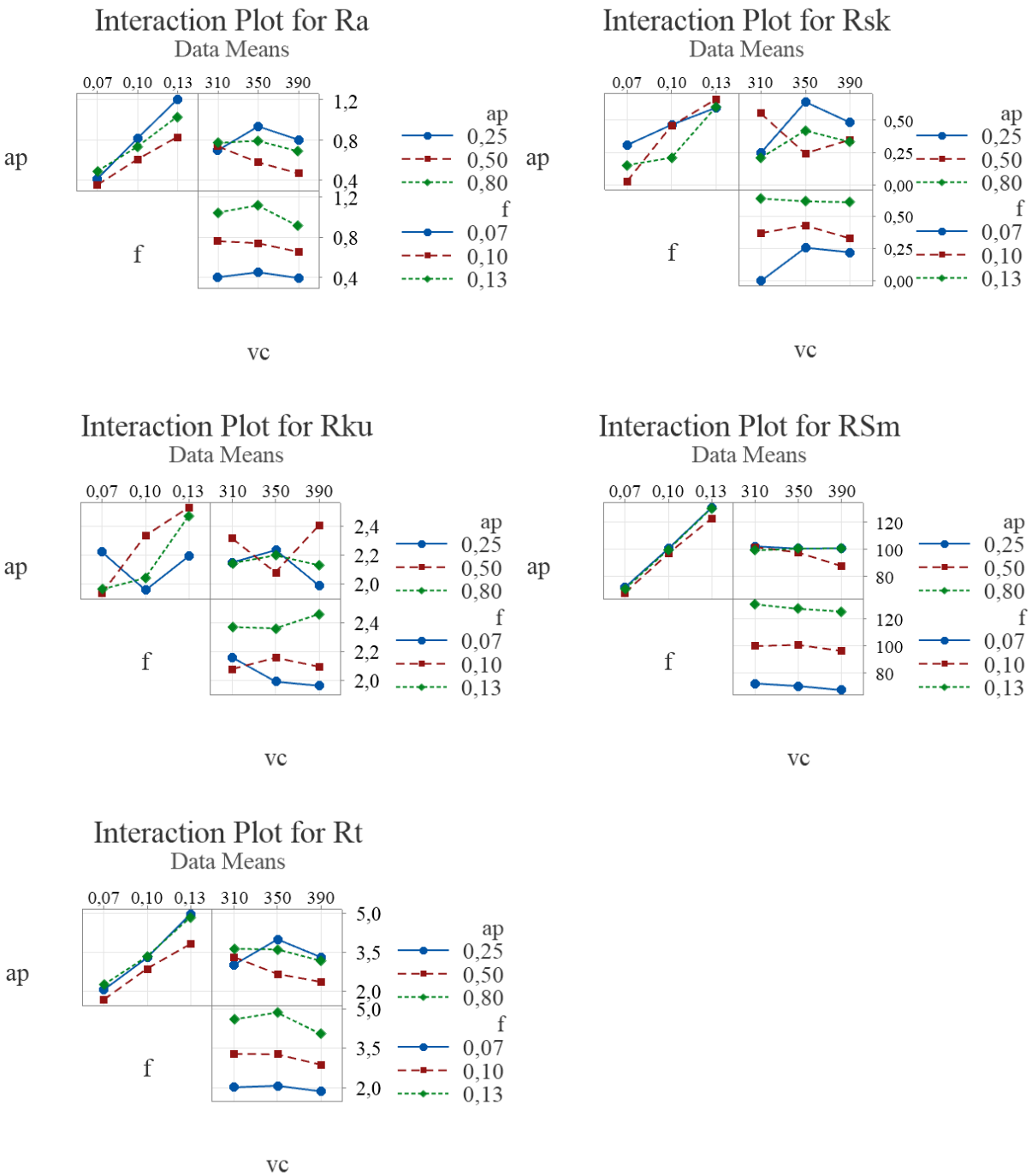



Figure 73. Interaction plots for the roughness parameters and processes parameters in Experiment 1.

The following observations stand out:

- The interaction between the depth of cut of 0.5 mm with the cutting speed provoked a reduction in Ra and Rt when increasing the cutting speed. This conclusion cannot be taken to the other depth of cut levels because there is an increase in roughness (Ra) when changing the cutting speed from the first to the second level.
- When increasing the feed using the second and third depth of cut levels, the kurtosis also is increased.
- The feed force ($F_z$) has a steeper increase with the feed rate when the depth of cut is at its highest level (0.8 mm) in comparison to other depth of cut levels. In this depth of cut level, the feed force becomes higher than the passive force for the first time.

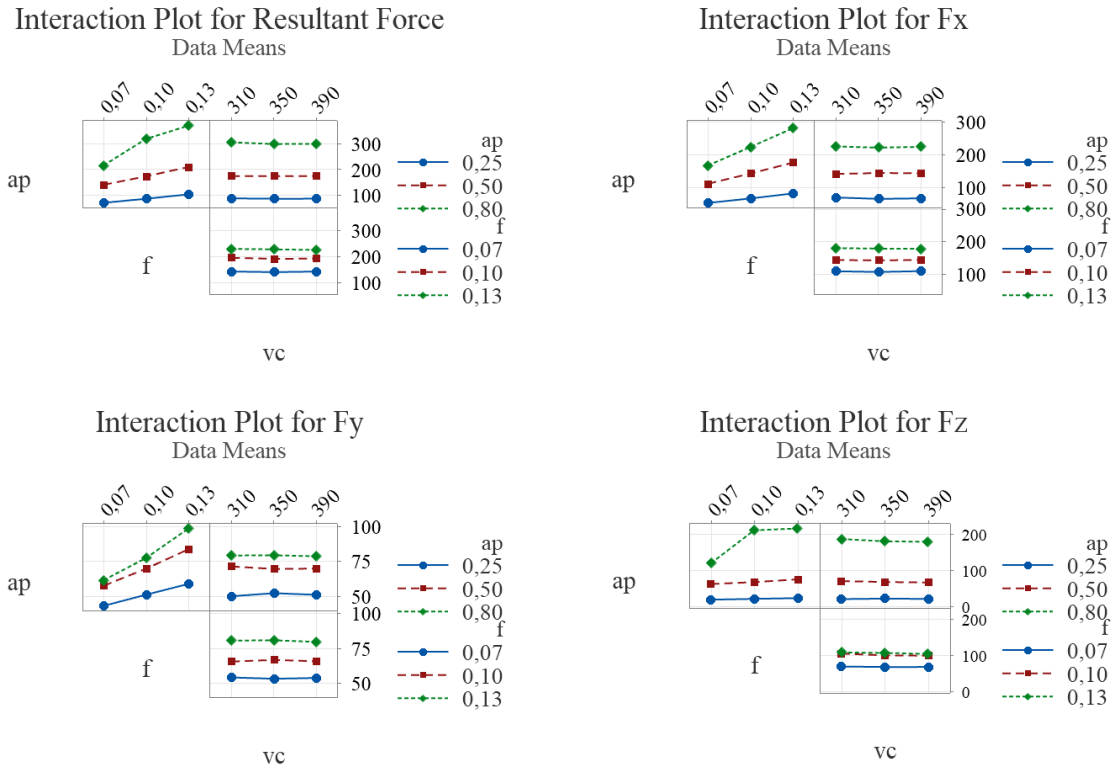



Figure 74. Interaction plots for the force components and processes parameters in Experiment 1.

## B.3 Additional main effects plots and interaction plots for Experiment 2

The influences' highlights of the depth of cut, feed rate, and tool condition on the roughness and machining forces were already presented and discussed in the results section. However, further information can be extracted from the main effects and interaction plots of Figure 75, Figure 76, Figure 77, and Figure 78.

The depth of cut ($a_p$) has a diverging influence on different roughness parameters. Ra, RSm, and Rt show a positive correlation, while a negative correlation exists for Rsk and Rku. The feed rate positively affects all roughness parameters but the Rku.

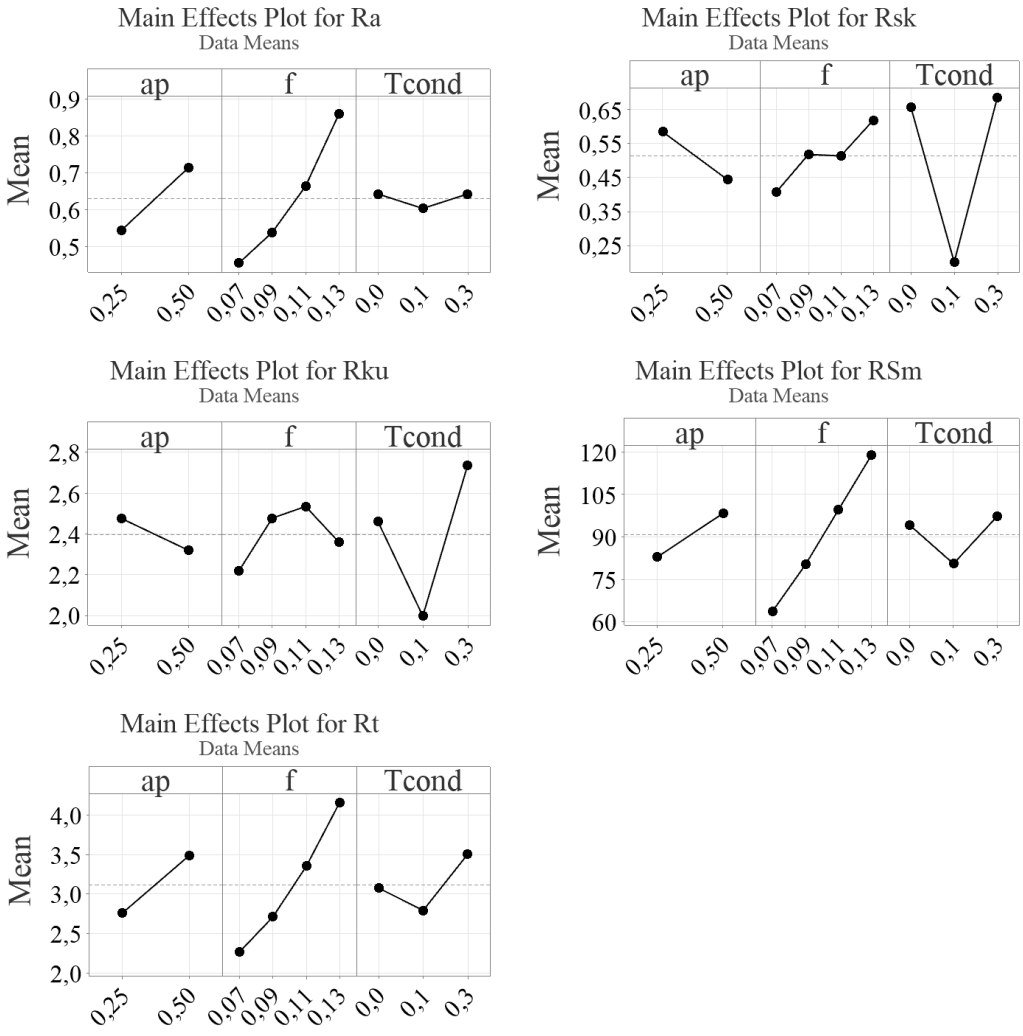



Figure 75. Main effects plots for the roughness in Experiment 2.

The tool condition (represented as the flank wear width VB) shows the same effect for all roughness parameters: the roughness decreases when the flank wear is 0.1 mm (mid-life

tool) and increases again when the flank wear is 0.3 mm (end life tool). This influence was more accentuated on Rsk, Rku, and RSm.

The factors influenced the force components similarly. All the effects were positive, meaning that there was an increase in the forces for an increase in the factor. Also, for $F_y$ and $F_z$, the feed rate showed little influence.

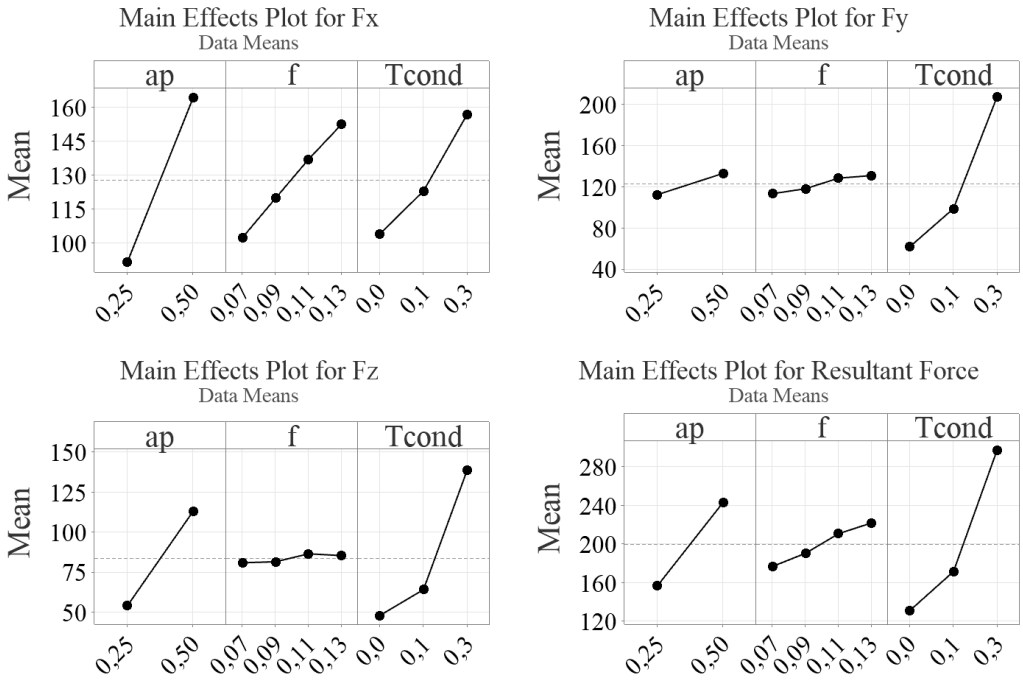



Figure 76. Main effects plots for the force components in Experiment 2.

From the interaction between the factors, it can be highlighted that:

- Ra tended to decrease with the increase in tool wear for the depth of cut of 0.5 mm. For the depth of cut of 0.25 mm, the behavior was the opposite: the Ra increased with tool wear.
- The increase on $F_z$ was steeper with tool wear when the depth of cut ($a_p$) was 0.5 mm than with the depth of cut of 0.25 mm.

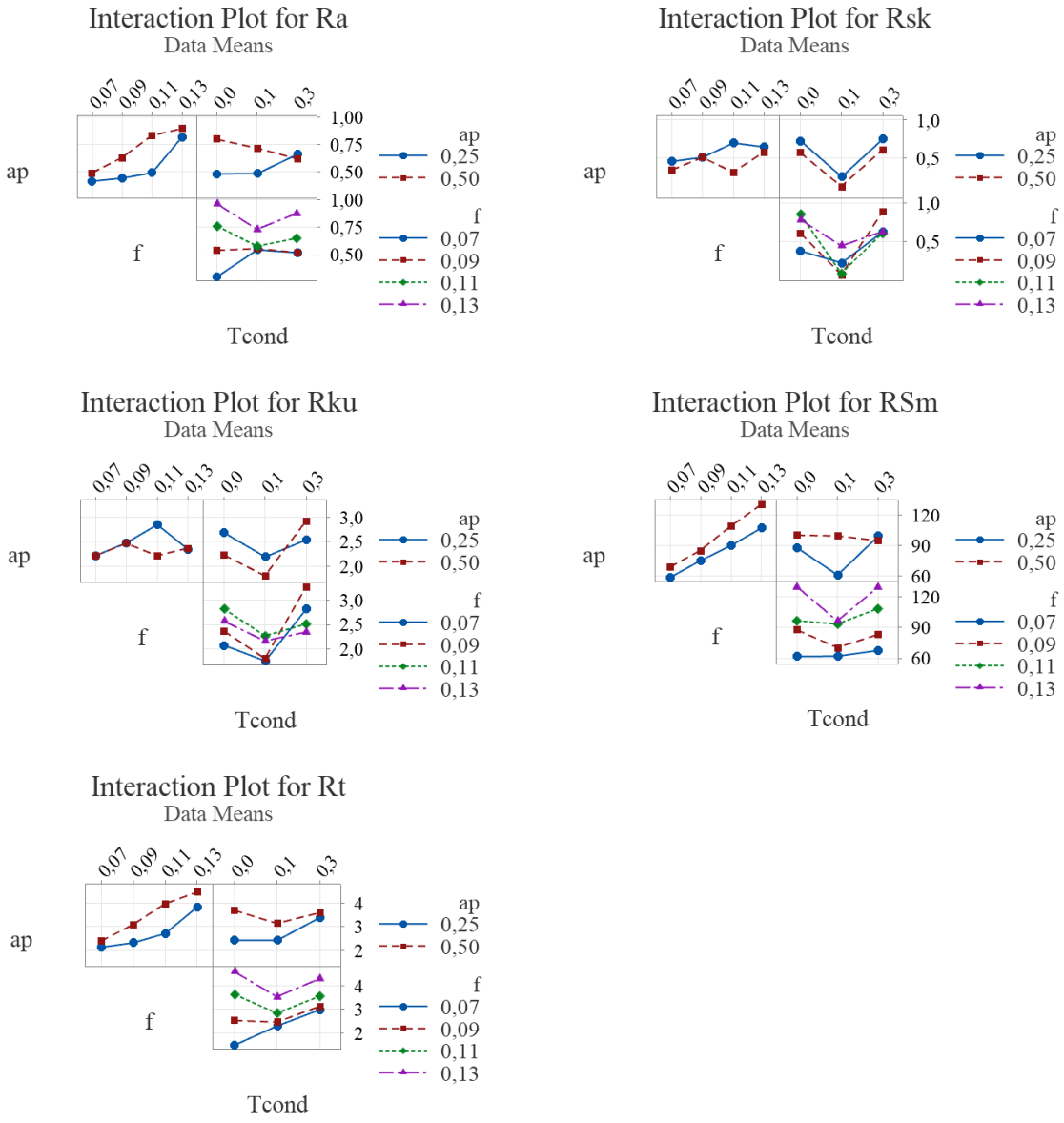



Figure 77. Interaction plots for the roughness in Experiment 2.

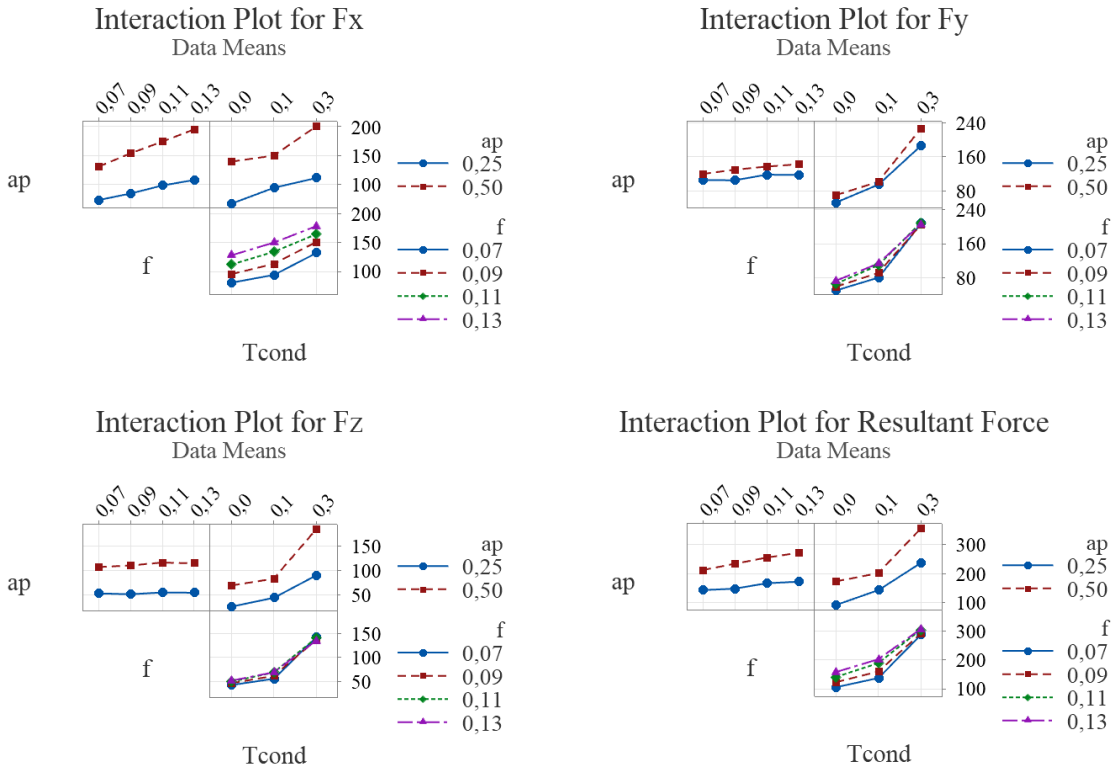



Figure 78. Interaction plots for the force components in Experiment 2.

# Appendix C – Considerations Over the Machine Learning Experiments

Some aspects must be registered regarding the topics 4.2.3 Investigations with the developed models, 5.2.3 Models trained with smaller datasets: ANN-.-.S, and 5.2.4 Cross-dataset testing.

## C.1 Avoiding data leakage

This dissertation's data comes from two experiments containing two replicas, meaning that one repetition was performed for every tested condition. For Experiment 1, there were 27 unique conditions of process parameter arrangements and 54 experiments when the replica is counted. Measuring the surface roughness six times for each condition has led to more data available—the 54 examples of Experiment 1 became 324 examples. The "large set" of 324 examples comprises 54 groups, with six examples each. The train and test sets were divided so that from the same group—meaning the same inputs—some of the six labels remained in the training set and others in the test set.

Another train-test split was tested, guaranteeing that each of the 54 groups would be only at the train or the test set. With this configuration, the MAPE was raised to circa 25%. In the former version, the model was able to train with more outputs from the same given inputs, which may have also configured a type of data leakage phenomenon, that, in this newer configuration, was avoided—what probably caused the error to increase. This result might be further improved with more hyper tuning experimentation. To the records, the surface roughness (Ra) in Experiment 1 presented a maximum coefficient of variation of 16.8%.

The code lines below show how to split train and test sets keeping the six examples of each of the 54 groups together. The difference from the original train-test split is the split function used: a *GroupShuffleSplit*—from scikit-learn—that introduces an argument "groups=" to which the division will follow. Moreover, keeping the column "Run ID" in the main table was necessary.

In addition, another validation method could be used to evaluate the model's generalizing efficacy, such as the many variations of cross-validation.

```
# ############## TRAIN-TEST SPLIT
# Split train and test data sets (ratio = 80/20)
```

```
split = GroupShuffleSplit(n_splits=1,
                                test_size=0.2,
                                random_state=42)

for train_index,\
    test_index in split.split(r_df, r_df["f"],
                              groups=r_df["Run ID"]):

    stratified_train_set = r_df.loc[train_index]
    stratified_test_set = r_df.loc[test_index]
```

## C.2 Improvement attempt for the cross-dataset test

The cross-dataset test results suggested that the models trained with values acquired with tool wear presence performed better than when the tool wear was not considered in data generation acquisition. The attempt of improvement referred to here is that, after overfitting was discarded, the models were trained again, using the whole dataset—all 324 values for Exp1 and 288 for Exp2—using the same setup of hyperparameters to produce more powerful models with improved generalization capacity. This attempt, however, did not produce better prediction responses. The model equivalent to the ANN1.H.L—but trained with all 288 values—had a mean MAPE of 40% when predicting the Exp1 dataset responses, a similar error to the original methodology.

# Appendix D – Program and Datasets

The *Python* packages used in this work are *keras 2.6.0*, *matplotlib 3.4.3*, *numpy 1.19.5*, *pandas 1.3.3*, *pickleshare 0.7.5*, *pip 21.1.2*, *scikit-learn 1.0*, *tensorflow 2,6.0*. Moreover, the *R* package used is the *AICcmodavg 2.3-1*.

The code lines and datasets used to train the ANN models are based on the Internet under the Digital Object Identifier (DOI): https://doi.org/10.34740/kaggle/ds/2205074.

# FOLHA DE REGISTRO DO DOCUMENTO

| 1. CLASSIFICAÇÃO/TIPO | 2. DATA | 3. REGISTRO N° | 4. N° DE PÁGINAS |
|---|---|---|---|
| DM | 02 de agosto de 2022 | DCTA/ITA/DM 066/2022 | 158 |

5. TÍTULO E SUBTÍTULO:

Surface roughness analysis in turning processes using ANN.

6. AUTOR(ES):

**André Dorigueto Canal**

7. INSTITUIÇÃO(ÕES)/ÓRGÃO(S) INTERNO(S)/DIVISÃO(ÕES):

Instituto Tecnológico de Aeronáutica – ITA

8. PALAVRAS-CHAVE SUGERIDAS PELO AUTOR:

1. Turning process, 2. Surface roughness, 3. Machine learning, 4. Artificial Neural Networks, 5. Cutting tool wear, 6. Machining forces.

9. PALAVRAS-CHAVE RESULTANTES DE INDEXAÇÃO:

Rugosidade de superfície; Usinagem; Metais; Ferramentas de corte; Máquinas aprendizes; Redes neurais; Engenharia de materiais

10. APRESENTAÇÃO: **X Nacional** **Internacional**

ITA, São José dos Campos. Curso de Mestrado. Programa de Pós-Graduação em Engenharia Aeronáutica e Mecânica. Área de Materiais, Manufatura e Automação. Orientador: Prof. Dr. Anderson Vicente Borille. Defesa em 03/06/2022. Publicada em 2022

11. RESUMO:

Surface roughness is one crucial quality element in machining production, and its prediction has been the case of study for many years. Generally, three approaches are performed to model it: empirical methods, theoretical/simulation methods, and soft computing methods. In this work, artificial neural network models (ANN) were trained with two training datasets generated by machining experiments—turning of an AISI H13 steel with cutting fluid. The first experiment with theoretically new-tool conditions produced 324 samples, and the second with cutting tool flank wear varying in three levels produced 288 samples. The ANN models were formed with three hidden layers (28-56-14) and Ra as the predicted output. One model used the depth of cut, feed rate, and cutting speed as inputs, and another model used those features and the machining force components—cutting force ($F_c$), feed force ($F_f$), passive force ($F_p$), and the resultant force. A strategy was proposed to augment the available data six times by increasing the measurements without increasing the experiments. The models trained with a smaller dataset were prone to overfitting and achieved less performance than those with larger datasets. The model trained with the machined forces performed 75% better than the model without them, producing a prediction error (MAPE) of less than 5%. Moreover, the models trained with the first dataset (without tool wear) could not generalize to the second dataset (with tool wear), indicating that tool wear is an essential factor when modeling surface roughness in turning processes using ANNs. Also, the ANN models performed better than a classical theoretical surface roughness equation.

12. GRAU DE SIGILO:

**( X ) OSTENSIVO** **( ) RESERVADO** **( ) SECRETO**