technische universität dortmund | lehrstuhl kunststoff-technologie | BOSCH | TÜRK-ALMAN ÜNİVERSİTESİ TÜRKISCH-DEUTSCHE UNIVERSITÄT

Masterarbeit

Manufacturing Technology

Investigating the Effects of Post-mold Cure on Warpage in Electronic Packaging Processes

vorgelegt von

cand. Arda Kadoglu
Matr.-Nr. 230794

Betreuer:

Prof. Dr. Ulrich A. Handge
Assist. Prof. Dr. Mehmet Ipekoglu
M.Sc. Corinna Niegisch
M.Sc. Daniel Hesse

Dortmund, 09.01.2023

BOSCH

Master Thesis:

Investigating the Effects of Post-mold Cure on Warpage in Electronic Packaging Processes

Background:

- Epoxy molding compounds (EMC) are commonly used thermosetting materials for producing molded encapsulants in electronic packaging industry.
- Post-mold cure (PMC) is a process in which the molded thermosetting parts are exposed to elevated temperatures again after their molding processes to reach higher cross-linking densities to ensure improved material properties. In addition to altering thermal and mechanical properties, part warpage is also affected by PMC process. Despite the alleviating effects on warpage, the PMC operations in industry are mostly conducted with the same parameters (~175 °C & ~4 h) with the main goal of reaching increased cross-linking densities without major considerations on warpage. Therefore, the effects of varying PMC parameters on resultant part warpage are not well-known.
- Due to reactive EMC nature and property deviations in between different batch deliveries from the manufacturers, it is always a challenging task to adjust process parameters for EMC materials via simulations and analytical methods. Therefore, there is a need for alternative ways to adjust and optimize process parameters.

Goals of the Thesis:

- Conducting literature research regarding electronic packaging reliability, warpage, and cure mechanisms that can be related to warpage, as well as the state of the art for machine learning concepts.
- Experimentally investigating the warpage reduction mechanisms that take place during PMC processes, acquiring process data from the experiments and correlating the resultant warpage with the PMC parameters used.
- Creating a data driven model with the experimental data to optimize the PMC parameters considering the resultant warpage as the main quality criterion.

Contact person:

Corinna Niegisch, M.Sc.

Robert Bosch GmbH

E-Mail: corinna.niegisch@de.bosch.com

Eidesstattliche Versicherung

(Affidavit)

Kadoglu, Arda

Name, Vorname
(surname, first name)

230794

Matrikelnummer
(student ID number)

☐ Bachelorarbeit
(Bachelor's thesis)

■ Masterarbeit
(Master's thesis)

Titel
(Title)

Investigating the Effects of Post-mold Cure on Warpage in Electronic Packaging Processes

Ich versichere hiermit an Eides statt, dass ich die vorliegende Abschlussarbeit mit dem oben genannten Titel selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

I declare in lieu of oath that I have completed the present thesis with the above-mentioned title independently and without any unauthorized assistance. I have not used any other sources or aids than the ones listed and have documented quotations and paraphrases as such. The thesis in its current or similar version has not been submitted to an auditing institution before.

Renningen, 09.01.2023

Ort, Datum
(place, date)

Unterschrift
(signature)

Belehrung:
Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden. (§ 63 Abs. 5 Hochschulgesetz - HG -).

Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu 3 Jahren oder mit Geldstrafe bestraft.

Die Technische Universität Dortmund wird ggf. elektronische Vergleichswerkzeuge (wie z.B. die Software „turnitin") zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:
Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offense can be punished with a fine of up to EUR 50,000.00. The competent administrative authority for the pursuit and prosecution of offenses of this type is the Chancellor of TU Dortmund University. In the case of multiple or other serious attempts at deception, the examinee can also be unenrolled, Section 63 (5) North Rhine-Westphalia Higher Education Act (*Hochschulgesetz, HG*).

The submission of a false affidavit will be punished with a prison sentence of up to three years or a fine.

As may be necessary, TU Dortmund University will make use of electronic plagiarism-prevention tools (e.g. the "turnitin" service) in order to monitor violations during the examination procedures.

I have taken note of the above official notification:*

Renningen, 09.01.2023

Ort, Datum
(place, date)

Unterschrift
(signature)

***Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the Bachelor's/ Master's thesis is the official and legally binding version.**

Abstract

In electronic packaging industry, post-mold cure is a commonly applied process in which the molded thermosetting encapsulants are exposed to elevated temperatures again for several hours after their molding operations to ensure improved material properties via reaching higher cross-linking densities. Some studies in literature and industrial applications showed that it is also a useful method to alleviate the package warpage, which is a major quality criterion. Despite these findings, the effects of different post-mold cure parameters on resultant warpage are still not well-known since it is mostly conducted with the same parameters in industry only to reach higher cross-linking densities without the considerations on warpage.

In this study, the effects of post-mold cure process on warpage of an epoxy molding compound encapsulant are investigated. Material characterization tests are performed to investigate the thermal and cure behaviors of encapsulant components. Experiments with varying post-mold cure temperatures and times are conducted to correlate these parameters with the resultant warpage to create a process knowledge. Part warpages are evaluated via 3D optical profilometer scans, and infrared spectra of the samples are measured to non-destructively detect the alterations in chemical composition as an indicator of changing cross-linking density. In the light of the experiments, the warpage reduction mechanism during post-mold cure is explained, which is related to thermal and cure shrinkages, as well as the glass transition behaviors of the materials. Moreover, utilizing the acquired experimental data, a machine learning algorithm is proposed which predicts the optimal post-mold cure parameters in accordance with the desired warpage and cross-linking density values after the operation. The algorithm was able to accurately predict the resultant warpage for test samples and the predicted optimal post-mold cure parameters showed an accordance with the experiments.

Both experiment results and algorithm predictions showed that the post-mold cure time can be reduced to ~15 minutes from several hours when warpage is considered as the main quality criterion to optimize for the materials and the geometry tested in the experiments.

Zusammenfassung

Post-mold Cure ist ein häufig angewandtes Verfahren im Bereich Electronic Packaging, bei dem duromere Teile nach der Formgebung erneut erhöhten Temperaturen ausgesetzt werden, um eine höhere Vernetzungsdichte und damit bessere Materialeigenschaften zu erreichen. Bisherige Studien und industrielle Anwendungen haben gezeigt, dass der Post-mold Cure auch eine nützliche Methode ist, um den Verwölbung der Verkapselung, der ein wichtiges Qualitätskriterium ist, zu verringern. Trotz dieser Erkenntnisse sind die Auswirkungen verschiedener Post-mold Cure Parameter auf die resultierende Verwölbung noch immer nicht genau bekannt, da der Post-mold Cure meist mit denselben Parametern durchgeführt wird. Das Ziel ist hierbei eine möglichst hohe Vernetzungsdichte zu erreichen, ohne die Verwölbung genau zu berücksichtigen.

In dieser Arbeit werden die Auswirkungen des Post-mold Cures auf die Verwölbung der verwendeten Prüfkörper untersucht. Materialcharakterisierungstests werden durchgeführt, um das Wärme- und Aushärtungsverhalten der Materialen zu untersuchen. Experimente mit unterschiedlichen Post-mold Cure Temperaturen und Zeiten werden durchgeführt, um diese Parameter mit der resultierenden Verwölbung zu korrelieren und ein Prozesswissen zu generieren. Die Verwölbung der Prüfkörper wird durch optische 3D-Profilometerscans bewertet. Es werden außerdem Infrarotspektren der Materialien gemessen, um die Veränderungen der chemischen Zusammensetzung des Epoxidharzes während des Post-Mold Cures als Indikator für Änderungen in der Vernetzungsdichte zu verwenden. Anhand der Experimente wird der Mechanismus zur Verringerung der Verwölbung während des Post-mold Cures erklärt und mit der thermischen und chemischen Schrumpfung, sowie dem Glasübergangsverhalten der Materialien in Verbindung gebracht. Darüber hinaus werden die gewonnenen experimentellen Daten verwendet, um einen Machine Learning Algorithmus aufzusetzen, der die optimalen Post-mold Cure Parameter vorhersagt, die am Ende die gewünschte Verwölbung und Vernetzungsdichte ergeben. Mit Hilfe des Machine Learning Algorithmus ist es möglich, die resultierende Verwölbung und die möglichst optimalen Post-mold Cure Parameter für die Testproben vorherzusagen.

Die experimentellen Ergebnisse und die Vorhersagen des Machine Learning Modells zeigen, dass die Post-mold Cure Zeit für die getesteten Materialien und Geometrien auf ~15 Minuten reduziert werden kann, wenn der Verzug als wichtigstes Qualitätskriterium für die Optimierung betrachtet wird.

Contents

1 Introduction 1
2 State of the Art 3
2.1 Electronic Packaging 3
2.1.1 Encapsulant Materials Used in Electronic Packaging 5
2.1.2 Epoxy Molding Compounds 6
2.1.3 Encapsulation Methods 8
2.1.4 Quality Criteria of EMC Encapsulants 11
2.1.5 Warpage in EMC Encapsulants 13
2.1.6 Post-mold Cure and Its Effects on Warpage 15
2.2 Machine Learning 19
2.2.1 Types of Machine Learning Algorithms According to Data Availability ... 21
2.2.2 Model Validation 25
2.2.3 Partial Least Squares Regression 26
2.2.4 Artificial Neural Networks 26
2.3 Summary & Motivation 31
3 Materials, Instrumentation and Methods 33
3.1 Materials 34
3.2 Material Characterization 35
3.2.1 Differential Scanning Calorimetry 35
3.2.2 Thermomechanical Analysis 37
3.3 Post-mold Cure Experiments 38
3.3.1 Sample Production 38
3.3.2 Measurements 41
3.3.3 Post-mold Cure Experiments 45
3.4 Optimization Algorithm 45
3.4.1 Data Preparation 47
3.4.2 Warpage Regression 48
3.4.3 Enthalpy Regression 48

3.4.4 Optimization Cycle ... 49

4 Results ... 51

4.1 Results from Material Characterization Tests ... 51

4.1.1 Differential Scanning Calorimetry Measurements ... 51
4.1.2 Thermomechanical Analysis ... 54
4.1.3 Overview and Interpretation of Material Characterization Tests ... 55

4.2 Results from Post-mold Cure Experiments ... 56

4.2.1 Results on Warpage ... 56
4.2.2 Interpretation of Results on Warpage ... 59
4.2.3 Results on Infrared Spectra ... 64

4.3 Results from Machine Learning Algorithm ... 65

4.3.1 Results from Warpage Regression ... 65
4.3.2 Interpretation of Results from Warpage Regression ... 67
4.3.3 Results from Enthalpy Regression ... 67
4.3.4 Interpretation of Results from Enthalpy Regression ... 70
4.3.5 Results from Optimization Model ... 71
4.3.6 Interpretation of Results from Optimization Model ... 73

5 Conclusion & Outlook ... 77

References ... 80

A1 Methodology for Updating Efficiency Terms ... I

List of Figures

Figure 2.1: Cross-section of an encapsulated IC chip 4
Figure 2.2: Cutting and encapsulation processes of an IC chip................................ 4
Figure 2.3: Different levels of electronic packages 5
Figure 2.4: EMC composition with simplified list of ingredients 7
Figure 2.5: Encapsulation process technologies.. 8
Figure 2.6: Compression molding of a multichip module package............................ 9
Figure 2.7: Schematic diagram of a transfer molding press (a) initial state and (b) during molding 10
Figure 2.8: Schematic illustration for defect types in an encapsulant 11
Figure 2.9: Void formation during transfer molding a) pores and surface defects on the pellet b) void propagation during transfer 12
Figure 2.10: Warpage formation in an IC chip encapsulant.................................... 13
Figure 2.11: Volumetric shrinkage due to cooling and curing during molding.......... 14
Figure 2.12: Schematic of warpage driven failures in an IC package 15
Figure 2.13: 2D representation of thermoset curing a) unreacted monomers b) formation of small-branched molecules c) gelled but incomplete cross-linked network d) cured network with almost no free monomers 16
Figure 2.14: Time-temperature-transformation diagram for a thermosetting system 17
Figure 2.15: Specific volume change during PMC... 18
Figure 2.16: Simplified workflow of a machine learning model 19
Figure 2.17: Overview of different types of machine learning algorithms 21
Figure 2.18: Classification algorithm a) train data with labels b) external test data without labels c) predicted labels 22
Figure 2.19: Regression through curve fitting 22
Figure 2.20: A two-dimensional domain with a) unlabeled clustered data points b) extracted clusters and anomalies 23
Figure 2.21: Dimensionality reduction via merging a) clustering data points in a 2D feature space b) point projections on regression line c) reduced feature space with projected features 24
Figure 2.22: Schematic illustration of a) underfitting b) overfitting c) good fit........... 25
Figure 2.23: Schematic of a single artificial neuron with m inputs 27
Figure 2.24: Commonly used activation functions a) sigmoid b) ReLU c) tanh........ 28
Figure 2.25: A multi-layered artificial neural network 28

Figure 2.26: Basic idea of back propagation a) initial weight with negative derivative b) first iteration in positive direction c) second iteration in negative direction 30
Figure 3.1: Overall steps and workflow of the thesis... 33
Figure 3.2: Main materials used in samples a) EMC granules pressed in sheet form b) PCB substrates without mounted electronics 34
Figure 3.3: An example DSC curve 35
Figure 3.4: Schematic of a Thermomechanical Analyzer 37
Figure 3.5: Molded sample 38
Figure 3.6: Press machine and the mold cavity used to produce samples 39
Figure 3.7: Schematic process steps in local pressure molding a) placing the EMC sheets and PCB substrate b) heating plates close c) compression stamps transfer and compress molten EMC d) solidified part is ejected 39
Figure 3.8: Fourier transformation of an interferogram to a spectrum 42
Figure 3.9: Attenuated Total Reflectance 42
Figure 3.10: 3D Optical Profilometer a) schematic CGIS setup b) grid illuminated image c) 3D height map of the sample 44
Figure 3.11: Principle of python script for warpage evaluation a) scanned image and area of interest b) 3D height data c) circle fitting in cross-section d) calculating the deflection via the curvature 44
Figure 3.12: Flowchart for process optimization with the machine learning model .. 46
Figure 3.13: Flowchart for warpage regression function... 48
Figure 3.14: Flowchart for enthalpy regression function 49
Figure 3.15: Flowchart for optimization tool 49
Figure 4.1: Dynamic heating section of DSC curves acquired from EMC samples molded at 170 °C with different cycle times 52
Figure 4.2: Measured reaction enthalpies of EMC samples molded at 170 °C with different cycle times 52
Figure 4.3: Glass transition temperatures acquired from DSC curves a) non-cured raw EMC b) EMC molded for 55 s at 170 °C 53
Figure 4.4: DSC curves of PCB substrate before and after 2 hours PMC at 185 °C 54
Figure 4.5: TMA measurements for highly cross-linked EMC................................. 55
Figure 4.6: Measured warpage values of sample groups A and B 57
Figure 4.7: Average warpage reduction rates of samples from Group A 58
Figure 4.8: Average warpage reduction rates of samples from Group B 59

Figure 4.9: Warpage formation after molding a) part geometry and the reference length l b) shrinkage graph for EMC and PCB c) warpage modes just after ejection (0) and when cooled down to room temperature (1) ... 60
Figure 4.10: Warpage states during post-mold curing ... 61
Figure 4.11: Intervals with different warpage reduction patterns ... 62
Figure 4.12: Color change in PCB substrates after PMC operations with different temperatures and times ... 63
Figure 4.13: IR Spectra of EMC before and after PMC operations with a) 170 °C and 30 min b) 170 °C and 120 min ... 64
Figure 4.14: Data points used for training the warpage regression tool ... 65
Figure 4.15: Comparison of measured and predicted final warpage values ... 66
Figure 4.16: Predicted and measured reaction enthalpy values for EMC encapsulants of the molded parts ... 68
Figure 4.17: Average predicted enthalpy values of molded parts before and after PMC ... 69
Figure 4.18: Comparison of predicted optimal point with experimental results ... 73
Figure 4.19: Iteration steps from a) optimization test (2) without dummy data and b) optimization test (4) with dummy data ... 74
Figure A1.1: Extended flowchart for optimization tool ... I
Figure A1.2: Flowchart for updating iterated efficiency terms ... II

List of Tables

Table 3.1: PMC experiments - molding parameters used in different sample groups 45
Table 3.2: PMC experiments - parameters used in the experiments for different sample groups 45
Table 3.3: Data Set 1 – Data acquired from PMC experiments 47
Table 3.4: Data Set 2 – External data acquired from DSC specimens 48
Table 4.1: Tested ANN structures and their resultant mean absolute errors 66
Table 4.2: Parameter set used for optimization model 72
Table 4.3: Predictions for optimal PMC parameters of a test sample 72

List of Abbreviations and Symbols

Abbreviations

3D	Three dimensional
ANN	Artificial neural network
CGSI	Confocal grid structured illumination
CTE	Coefficient of thermal expansion
EMC	Epoxy molding compound
FTIR	Fourier transform infrared
IC	Integrated circuit
LVDT	Linear variable displacement transducer
PCA	Principal component analysis
PLSR	Partial least squares regression
PMC	Post-mold cure
ReLU	Rectified linear unit
SVM	Support vector machine

Symbols

Symbol	Unit	Description
α	%	Degree of cure
h	J/g	Reaction enthalpy
h_0	J/g	Initial reaction enthalpy
h'	J/g	Final reaction enthalpy
T_g	°C	Glass-transition temperature
$T_{g(m)}$	°C	Midpoint temperature
$T_{g(r)}$	°C	Return-to-baseline temperature
T_m	°C	Melting temperature
δ	µm	Warpage
δ_0	µm	Initial warpage
δ'	µm	Final warpage
$\Delta\delta$	%	Warpage reduction rate
v	m^3/kg	Specific volume

1 Introduction

Semiconductors are employed in a very broad range of industrial applications varying from consumer electronics and home appliances with low costs and high production volumes; to automotive, military and aerospace industries with longer life cycles and harsh working conditions [1,2]. Semiconductor encapsulation is an important step in electronics production which directly affects the final quality of the electronic system in terms of integrity and safety against environmental damages [2,3]. Molding has been a dominant method for high volume polymer encapsulated electronic package production [4]. Epoxy molding compounds (EMCs) are the most commonly used polymer encapsulation materials today with their good thermal stabilities and matching properties for molding processes [5]. Since a high cross-linking density in thermosetting EMC is required for good thermal stability of an encapsulant, post-mold cure (PMC) is a common process that is applied to electronic packages, in which they are exposed to elevated temperatures again for elongated times after their molding operations since reaching high chemical conversion levels is not always possible during molding itself due to long mold times needed and economic constraints [6,7].

During molding, several defects might occur in the package such as warpage, delamination, voids and polymer cracks [8]. Warpage is a major defect in molded encapsulants, in which the part loses its planarity mostly due to thermal mismatch in between the encapsulant material and other components in the package [8,9]. Even though the selected materials and molding parameters are the key factors to avoid warpage formation, post-mold cure is also a common way to alleviate the manufacturing induced warpage in electronic packages [10,11]. Despite being a common process, the optimal PMC parameters are still a field of study that is open to improvements. Since its main purpose is reaching higher cross-linking densities to ensure improved mechanical properties of the encapsulant, the common practice in industry is to post-mold cure the parts for long times up to 16 hours at elevated temperatures close to their mold temperatures [12]. Considering warpage as the main quality criterion, the effects of different PMC parameters on resultant warpage are not well-known.

Due to time dependent properties of EMCs, it is always a challenge to analytically model or simulate their behavior. Despite the huge annual production volume, the

parameters for EMC encapsulation processes are still mostly arranged via trial and error manners [13]. Due to alterations in raw material properties during storage, every batch of products might have different quality levels even though they are produced with the same parameters. In addition to these storage related alterations, every material batch delivered in different times from the same manufacturer might have slightly differing material properties which makes the parameter adjustment step even more challenging [5]. These difficulties make the process simulations and analytical studies very costly and time consuming for industrial uses since they need to be repeated every time. Therefore, there is a search for alternative methods to optimize mass production of electronic packages. For such purposes, data driven methods are proven to be useful such as the machine learning method, which is the study of statistical algorithms that has been a mainstay in information technologies and successfully employed for industrial optimization purposes as well [14,15]. Due to its capabilities on creating data driven prediction models which are applicable to product and production optimization, there is an increasing trend for machine learning usage in industry [15]. Such methods are successfully applied to injection molding processes with polymers and can be utilized for other processes with different approaches [16].

In this master thesis, the effects of different post-mold cure parameters on warpage of an EMC encapsulated package are investigated. Molded parts are post-mold cured with different parameters and their warpage values before and after PMC operations are compared. Similarly, infrared spectra of the samples are measured to correlate the alterations in the chemical composition of EMC during PMC with the changes in cross-linking density. With the goal of optimizing the PMC parameters considering warpage as the main quality criterion, a machine learning based optimization algorithm is built via the results acquired from the experiments.

Starting with an overview of the state of the art in Chapter 2, the materials and the methods used for the experimental work are introduced in Chapter 3, as well as the description of the developed machine learning algorithm. The acquired results are then presented in Chapter 4, followed by the conclusion and outlook of the work in Chapter 5.

2 State of the Art

This chapter introduces two important concepts regarding the thesis, namely electronic packaging, and machine learning. Starting with the state of the art for electronic packaging and summarizing its challenges in Chapter 2.1, the encapsulant materials used in electronic packaging are introduced in Chapter 2.1.1. Being the material used in the experiments, detailed composition and property details of epoxy molding compounds are given afterwards in Chapter 2.1.2. Common encapsulation methods are introduced in Chapter 2.1.3, followed by the description of overall quality criteria in polymer encapsulated electronic packages. The main focuses of the thesis, the warpage issue in electronic packages and the post-mold cure process are described in Chapters 2.1.5 and 2.1.6 respectively. Following these, in Chapter 2.2, the machine learning concept is explained which is one of the most trending topics in a large variety of industrial applications including electronic packaging [17,18]. After the concept introduction and historical overview, different algorithm types are presented in Chapter 2.2.1. Being a main concern, the model validation phase is described in 2.2.2. After these introductory parts to machine learning, two algorithms Partial Least Squares Regression and Artificial Neural Networks are introduced in Chapters 2.2.3 and 2.2.4 respectively. Finally, these two main sub-chapters are summarized, and the motivation of the thesis is presented in Chapter 2.3.

2.1 Electronic Packaging

With the most simple approach, an electronic package is an encapsulant which houses electronic elements such as transistors, chips, resistors and capacitors to protect them from environmental damages and provide their interconnections in between each other and other electronic systems [19]. The encapsulant protects the package content against corrosives, mechanical and thermal shocks, vibration and damages during transportation, storage and usage [3]. As electronics are highly involved in daily life and all branches of industry, electronic packages are used in a very broad product spectrum varying from the low cost and high volume products in consumer electronics such as cell phones, computers and home appliances that work in relatively benign environments; to the products working under harsh conditions with high life cycles such as the products used in automotive and aerospace industries [1]. **Figure 2.1** shows the cross-section of an encapsulated Integrated Circuit (IC) chip which is a basis

element used in many electronic products. The functional elements such as the silicon IC chip, the wires that connect the encapsulated elements, and the lead-frame for outer connections of the package are assembled and protected via the encapsulant [19].

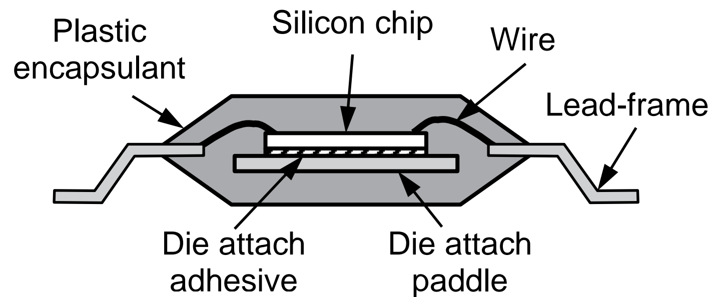

Figure 2.1: Cross-section of an encapsulated IC chip [19]

Before encapsulation, the initial step is generally the chip or the wafer level component [20]. An example regarding the encapsulation of an IC chip is shown in **Figure 2.2**. The silicon IC chip is cut out from the wafer using a diamond blade and then attached to a lead-frame using an adhesive. In some applications, printed circuit board (PCB) substrates can also be used instead of lead-frames to mount chips and sensors [21]. After the connection of conductive wires to the lead-frame, the components are encapsulated with a polymer material [20]. After the encapsulation, the excessive lead-frame parts are trimmed and formed into their final geometries that is required for further integration to other electronic assemblies [20,22].

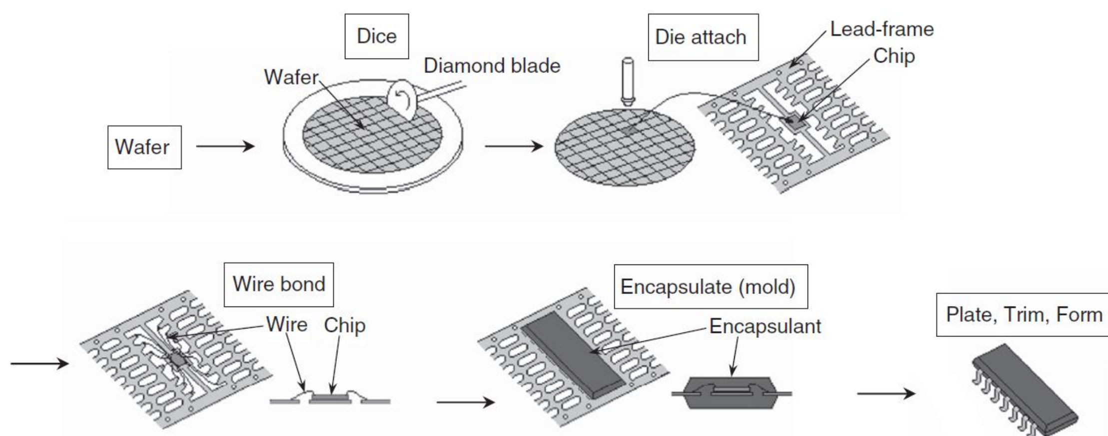

Figure 2.2: Cutting and encapsulation processes of an IC chip [20]

Integrating multiple electronic packages, electronic systems are built as shown in **Figure 2.3**. Starting from the IC chip as the zero level, the first level consists of the encapsulation and interconnection of the chip and other conductive elements. The

assembly in which multiple first level packages are connected to a printed circuit board (PCB) is called the second level package. Further connections of multiple packages create more complex assemblies that form electronic systems in the end. [19]

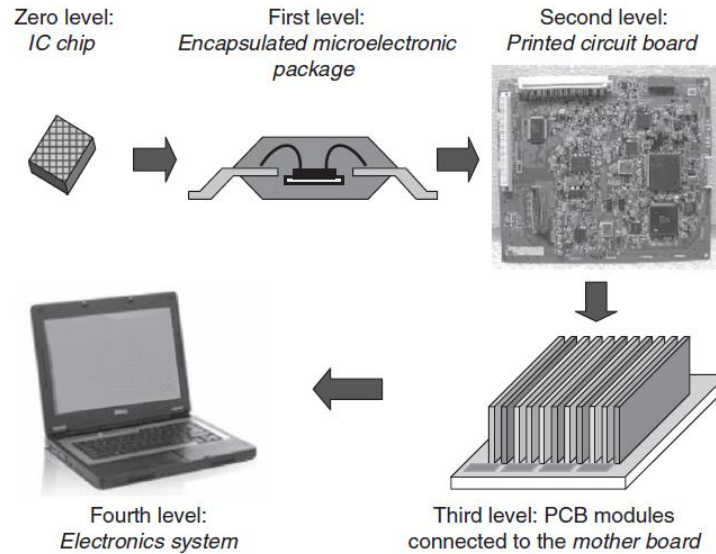

Figure 2.3: Different levels of electronic packages [19]

Depending on the product use cases, other materials such as glass, ceramics or metals can also be used with different encapsulation methods in accordance with the required properties for the package working conditions [23]. The selection of proper materials and manufacturing methods is a major aspect that directly affects the final product reliability [24].

2.1.1 Encapsulant Materials Used in Electronic Packaging

Apart from the major concern of housing and protecting the elements from the environmental effects, functionality and safety of an electronic system are also amongst the main duties of an encapsulant [19]. The package must withstand certain physical, thermal, and electrical loads. Especially in power electronics applications, the package is exposed to extreme thermal and electrical load conditions which makes it even more important to have a sound encapsulant against these loads [25]. With regard to these operational conditions, it is needed to satisfy some limits in certain material properties of encapsulant such as coefficient of thermal expansion, thermal conductivity, mechanical strength or dielectric constant [26]. Combining these quality related issues with economic aspects, material selection step requires variety of considerations to have a balance in between the function, cost, reliability and

processability. All these mentioned concerns make electronic packaging one of the most materials-intensive industrial applications today [27].

The materials used for encapsulation of electronic packages can be divided into three main types as metal-based, ceramic-based and polymer-based materials [26]. The terms hermetic (for metal and ceramic-based encapsulants) and non-hermetic (for polymer-based encapsulants) are also commonly used in literature [22]. The usage of polymer-based materials makes up to 90 % of the encapsulation applications in electronic systems with their advantages in cost, design flexibility, and lightweight [5]. Focusing only on microelectronics, the usage ratio of polymer-based encapsulation materials goes up to 99 % due to the process allowances of polymers [22]. Even in military industry, the metal and ceramic based high performance materials are being replaced with commercial off-the-shelf polymer materials in the recent years [22].

Polymer-based encapsulant materials consist of polymers and filler materials that combine the advantages of polymers in terms of processability and their lightweights, and the required properties of the filler materials such as thermal stability and electrical conductivity [26]. Epoxy molding compounds (EMCs) are the most commonly used compound polymers in electronic packaging applications. Apart from epoxies; silicones, polyurethanes and phenolics are also commonly employed as base materials [19,26]. Being the main material focused on in this thesis, the composition details and properties of EMCs are described in more detail in the following chapter.

2.1.2 Epoxy Molding Compounds

Epoxy molding compounds are the most common thermosetting materials for semiconductor encapsulation processes with their advantageous properties in terms of moldability, mechanical strength, thermo-mechanical matching with other elements in the package, small thermal shrinkage, and good moisture resistance. Their thermal stabilities in their cured states makes them favorable for electronic products that work under elevated temperatures. [5]

A variety of ingredients are included in an EMC composition to satisfy the required mechanical, thermal and chemical properties, as well as the desired processability for molding processes. Some examples to these ingredients are epoxy and hardener

resins, silica fillers and coupling agents [28]. A typical composition of an EMC material is illustrated in **Figure 2.4**.

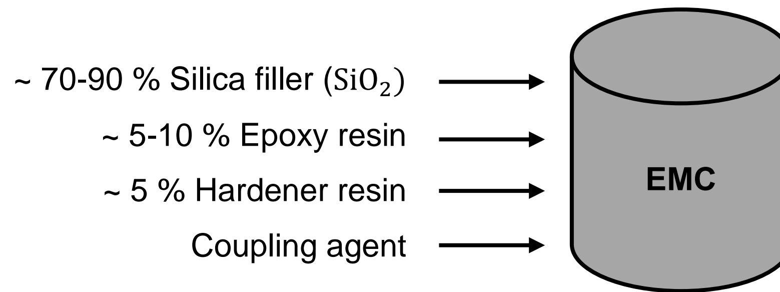

Figure 2.4: EMC composition with simplified list of ingredients – adapted from [28]

As they are named after, EMCs are epoxy resin based materials, which makes them well-balanced for molding applications [28]. Epoxy resins have high adhesion strength, low shrinkage rate and good chemical and moisture resistances [28]. In addition; their low melt viscosity, low curing temperature and short curing time make them really well-fitting for molding processes for mass production [28]. Epoxy resins require hardeners to start their cross-linking reactions for further curing. At elevated reaction temperatures, epoxy resin and hardener transforms into a densely cross-linked three dimensional network [29]. Phenol novolac resins are commonly employed as hardeners in EMC composition due to their performance on heat and moisture resistance, curing properties, and storage stabilities [2]. Chapter 2.1.6 describes the changes in material properties during curing in more detail.

A large amount of silica fillers are used to ensure a lower coefficient of thermal expansion (CTE), as well as a lower moisture absorption of the EMC [2]. They are also known for increasing the elastic modulus and decreasing the chemical shrinkage during curing [5,30]. Being able to decrease the thermal and chemical shrinkage, the amount of filler particles has a direct effect on part warpage as shown by Kiong et al [31]. Apart from the favorable aspects, the increasing amount of particles lead to an increase in the flow viscosity of the EMC, which decreases the moldability and promotes additional failures such as wire sweep, which is an encapsulation defect that is described in Chapter 2.1.4 [5,32]. Coupling agent is used to accelerate the interaction in between the epoxy resin and the filler silica particles [33].

When producing molded parts with EMCs, the cycle times for molding are kept as short as possible due to economic concerns. Thus, a close to full cross-linking density of the material is not reached during molding process [6]. It is well known from the industrial

experience that the EMCs have insufficient thermal stability in low chemical conversion levels, therefore the molded parts cannot be efficiently used in high temperature applications. To achieve increased mechanical properties and thermal stability, there is a need for the post-mold cure process which will be discussed in more detail in chapter 2.1.6. [7]

Due to time dependent mechanical and viscous properties of EMCs, it is a complicated task to analytically model and simulate their behavior. The parameter sets for EMC encapsulation methods are still mostly arranged in trial and error manners due to this complexity [13]. In addition, the reactive nature of epoxy resins may cause alterations in EMC properties during the storage of the raw material [34]. Another major problem that is faced in the industry is the inevitable deviations in material properties in between different batches of materials provided by the manufacturers [5]. Considering all these aspects, it is still a challenge to optimize EMC encapsulation processes despite being the most common material for such purposes [5].

2.1.3 Encapsulation Methods

Encapsulation process technologies for electronic packages can be divided into five main groups as molding, glob topping, potting, underfilling and printing as shown in **Figure 2.5**. The selection of the proper method is dependent on different parameters such as required cycle time, production volume, used material and package type [35]. For simplicity, the focus on this chapter is kept on the encapsulation methods for the EMCs.

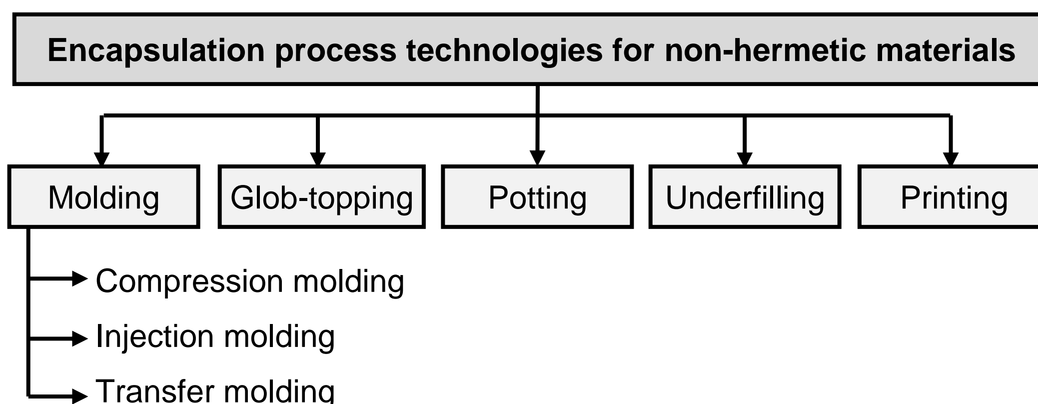

Figure 2.5: Encapsulation process technologies [35]

Since polymers are relatively cheap and easy to process in comparison to other encapsulant materials, they have been used for products with huge production volumes. Molding EMCs to encapsulate the semiconductors has been a dominant method and as a result, a relatively large know-how on this area is created according to specialized requirements of each application. [4]

In the earliest applications of EMC encapsulation, compression-molding was used in which the molding compound is heated and compressed inside a cavity [22]. **Figure 2.6** shows a process schematic for compression molding of a multichip module. In comparison to other molding methods, compression molding has the ease of die design, allowances of larger mold sizes and simple material flow paths leading to less issues which made it a favorable method [36]. On the other hand, high process and capital costs prevent compression-molding from being used in high volume mass encapsulation processes [36].

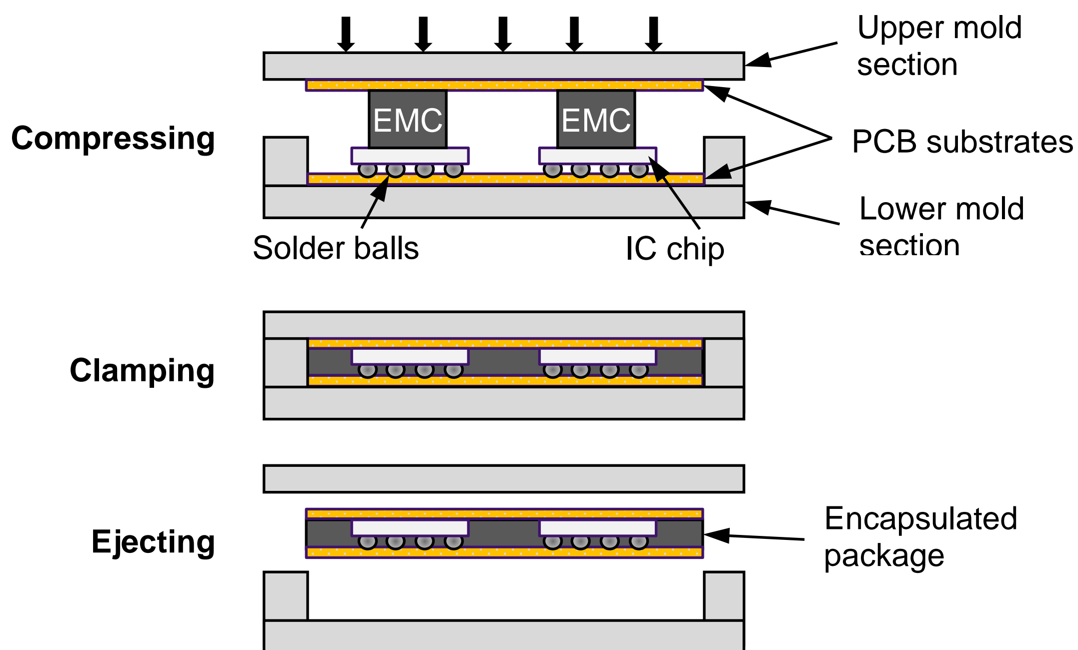

Figure 2.6: Compression molding of a multichip module package - adapted from [35]

Being initially developed for thermoplastic materials, injection molding can also be used for thermosetting materials with some modifications in the process. It is a faster method for mass production of encapsulants with the allowance of molding into multiple cavities at one shot. But the complexity of the tools makes the process an expensive alternative. Especially with thermosetting materials, the process gets more complex to adjust, optimize and monitor. [37]

With its advantages on allowing multiple cavities and processing thermosetting materials, transfer molding has gained worldwide acceptance and is the most common method today for large scale EMC encapsulation of electronic packages [38]. Compared to compression-molding, being able to mold into multiple cavities is a cutting-edge advantage for mass production purposes [35]. In comparison to injection molding, transfer molding process requires lower process pressures which makes it more suitable to encapsulate the delicate and intricate components [35]. Shorter sprues and runners make it possible to mold materials with shorter pot times [35].

During transfer molding, the preformed EMC pellet is first heated until it melts in a pot and then the transfer plunger compresses the molten EMC into the mold cavities in which the material is formed to the desired geometry as shown in **Figure 2.7** [35]. After filling the cavities with molten EMC, the packing pressure is increased. Due to high temperatures of the mold plates, the thermosetting material will start forming further cross-links and solidify as a result of curing [39].

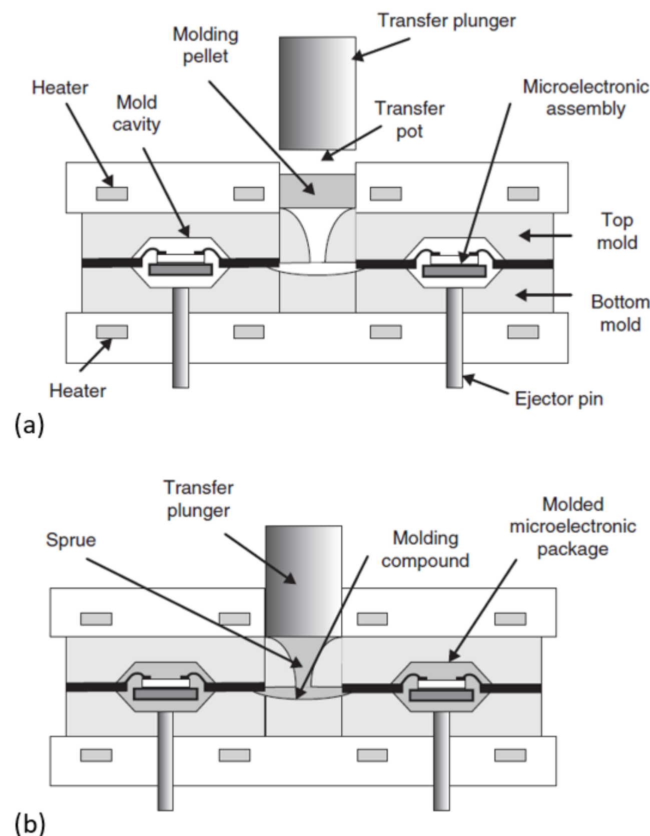

Figure 2.7: Schematic diagram of a transfer molding press (a) initial state and (b) during molding [35]

The curing of the material must be controlled carefully during molding process since a fast curing may lead to an excessive increase in viscosity and have a negative impact on the flow properties [2]. A fast hardening of the material will cause defects in the final geometry and possibly damage the semiconductor constituents in the package as a result of viscous flow [2]. A slow and insufficient curing on the other hand, will lead to an insufficient hardening of the material thus to problems during ejection of the molded encapsulant [40].

2.1.4 Quality Criteria of EMC Encapsulants

Being affected by various factors such as material properties, process parameters and environmental conditions, encapsulant defects might occur in any stage of manufacturing or assembly [5,41]. Examples to common process related defects include voids, polymer cracking, interface delamination, insufficient material cure and warpage [8]. In addition, the encapsulation process itself might lead to defects in non-encapsulant elements in the package such as the wires and the chips due to EMC flow and shrinkage [8]. **Figure 2.8** illustrates some of these defects. Since the package quality has a direct effect on the reliability of the electronic system, these manufacturing related defects can be highly critical for the whole product [42].

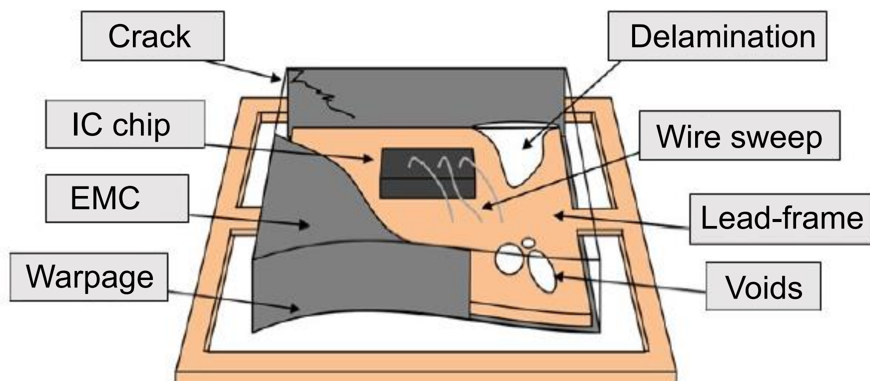

Figure 2.8: Schematic illustration for defect types in an encapsulant [5]

During molding processes of EMCs, encapsulation stresses occur as a result of cure-shrinkage [43]. In addition, thermal stresses are built due to thermal shrinkage while cooling back to room temperature after the molding process. These residual stresses might promote the mechanical failure of EMC encapsulants through polymer cracking [44]. Furthermore, the presence of materials with dissimilar CTEs lead to high interfacial stresses because of dissimilar thermal shrinkage, and the cure-shrinkage of

EMC also contributes to these interfacial stresses [45]. When these stresses exceed the adhesion strength, delamination occurs and the components detach from each other, leading to a failure of the electronic package due to structural and electrical integration issues [46,47]. With the accompaniment of excessive moisture, the so-called *popcorning* phenomenon might occur in which the encapsulant cracks as a result of high steam pressures acting on stress concentrated zones under high temperatures [45].

A major defect caused by the thermal mismatch is warpage, in which the part loses its planarity as a result of bending and deformation [8]. Warpage is a critical defect that leads to mechanical failures such as cracking and delamination, and also affects the package reliability in terms of thermal management, dimensional stability and integrity [10,48]. Being the focus point of this thesis, effecting mechanisms and reliability related influences of warpage is discussed in detail in Chapter 2.1.5.

Voids are formed via the trapped air or gas inside the encapsulant, which are mostly released by the volitant content in the EMC or caused by the porosity in the pellet [49]. **Figure 2.9** illustrates the void formation during transfer molding process. Corrosion formation on wires or other electronic components is caused by the voids in which the moisture condensation takes place [49]. Moreover, the gaps can induce stress concentrations and lead to polymer cracks [5].

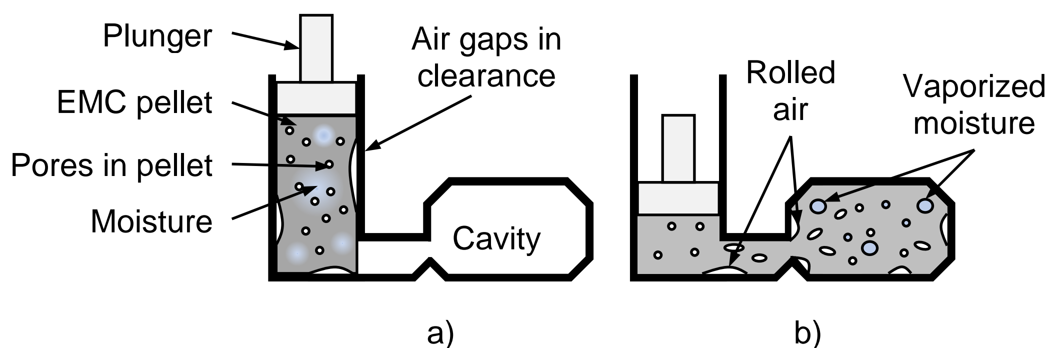

Figure 2.9: Void formation during transfer molding a) pores and surface defects on the pellet b) void propagation during transfer [49]

Wire sweep is the occurrence of wire displacements due to flow momentum of the molding compound [50]. Even though it is in theory possible to minimize wire sweep with adjustments in flow speed and flow paths, a low transfer rate with low viscosity is not possible in transfer molding of EMCs due to material curing [50]. The wire

displacement can cause short circuits in high density wire groups. Furthermore, the thin wires may even break under high tensile stresses [8].

Even though eliminating all these types of defects is not possible, they can be minimized with process parameter optimizations and additional steps applied after the molding process [8].

2.1.5 Warpage in EMC Encapsulants

Warpage is the defective phenomenon in which the part loses its planarity due to bending and deformation as introduced in Chapter 2.1.4. It is found to be the main failure mode (~75 %) in thin encapsulated electronic packages [51]. Failure analysis results show that over 90 % of the coplanarity in a package is formed just after the molding process before any usage or storage effects [51]. After the molding process, the package is exposed to residual stresses when cooling down to room temperatures, and with the presence of dissimilar materials with different CTE values in the package, warpage is formed due to different shrinkage values and the adhesive forces in between these components [9]. **Figure 2.10** illustrates the warpage formation in an EMC IC chip encapsulant with a PCB substrate on which the electronic components are mounted.

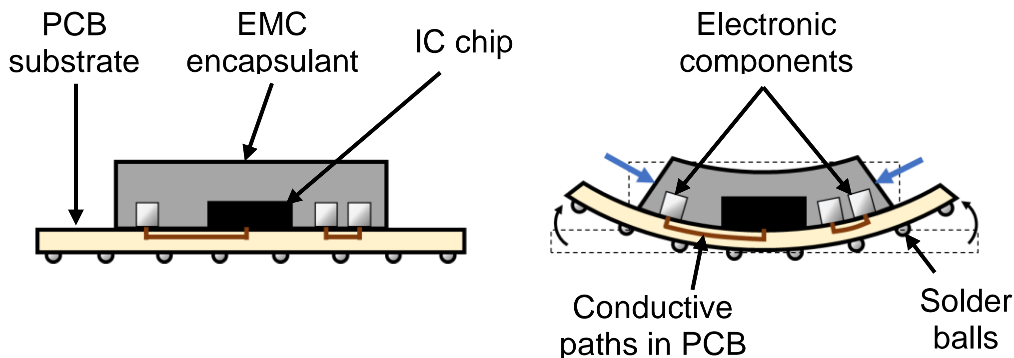

Figure 2.10: Warpage formation in an IC chip encapsulant

Even though the thermal stresses caused by the CTE mismatches were assumed to be the only source of warpage formation in the initial models, later studies show evidences to the effects of cure shrinkage on warpage [8,52]. During the material cure, the molten thermosetting materials transform into a viscoelastic solid and with further cross-link formation, the cure-shrinkage and cure related residual stress formations take place [53]. These considerations invalidate the assumption of stress-free

components mentioned above. While the part is being cured in mold cavity at high temperatures and its cure shrinkage is being restricted by geometrical constraints in the tool, residual stresses arise which later contribute to warpage formation when the part is taken out of the molding machine [43]. **Figure 2.11** shows the volumetric changes due to thermal shrinkage and cure shrinkage. As seen, the specific volume decreases during curing as a result of further cross-link formation in between the monomers while the part is being cured at T_{mold} [30]. In addition, the effects of changing glass transition temperature (increasing from T_{g0} to T_{g}) can also be seen, which leads to a lower thermal shrinkage than the thermal expansion due to earlier transition into glassy state with a lower CTE value of α_1 [9]. Chapter 2.1.6 gives detailed information about the cure effects on shrinkage, thus on warpage.

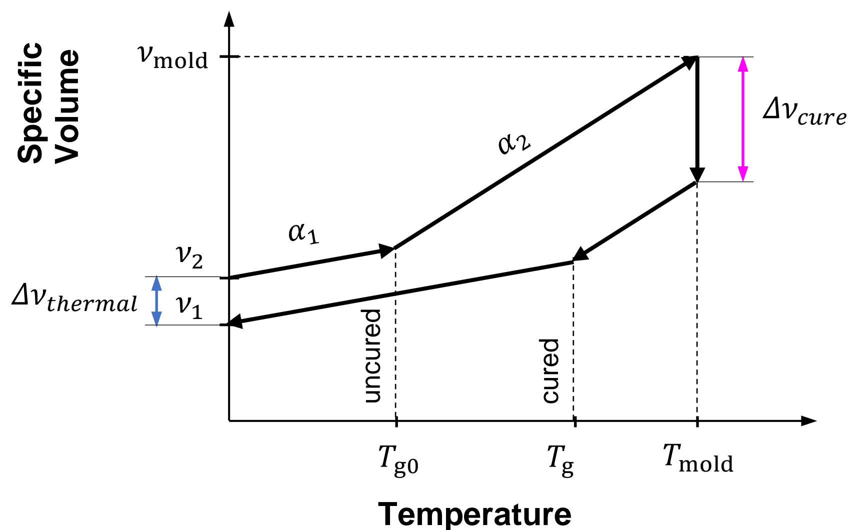

Figure 2.11: Volumetric shrinkage due to cooling and curing during molding [43]

Excessive warpage and residual stresses in the final encapsulant may cause device failure due to dimensional and geometrical instabilities [54]. A major problem arises considering the assembly of package to motherboards or to other packages [55]. **Figure 2.12** illustrates the connection infeasibilities and warpage driven defects in an IC package. As seen, the manufacturing related warpage prevents the full integrity by hindering the solder operations for pins on the high elevated sides of the PCB. Furthermore, additional warpage can be induced in the package due to heat flow during soldering. Therefore, tensile stresses occur in the solder joints [56]. The compressive stresses may damage the chip and other functional elements in the package, or the tensile stresses acting on encapsulant might lead to cracks. The conductive elements on the PCB substrate may break due to tensile stresses [52].

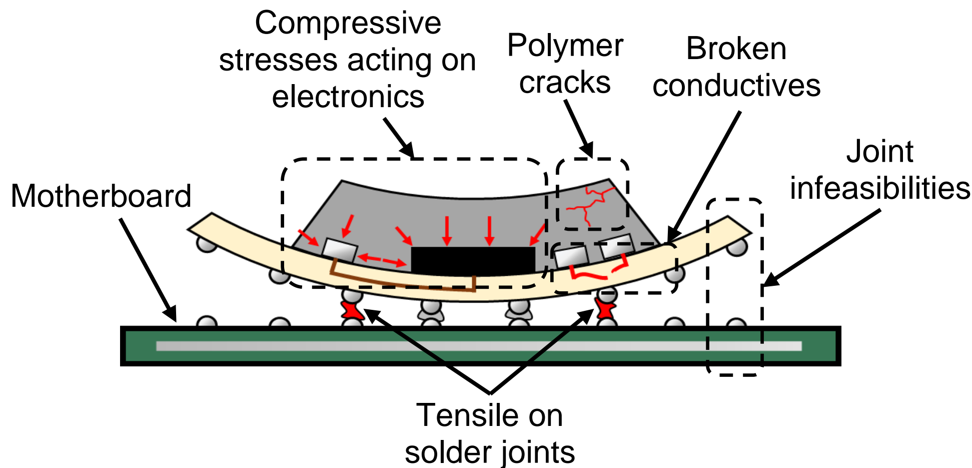

Figure 2.12: Schematic of warpage driven failures in an IC package

Warpage can be minimized with the optimization of process parameters. Selection of materials with similar thermal properties is another critical factor for reducing warpage [10]. In addition, the package geometry and assembly are design related optimization possibilities [8]. Reduced cooling rate is also shown to be an alleviating strategy for resultant warpage [43,51]. Last but not least, post-mold cure is another common process to reduce the manufacturing related warpage in electronic packages [11]. Being the main process focused on in this thesis, the post-mold cure step and its effects on warpage is discussed in the following chapter.

2.1.6 Post-mold Cure and Its Effects on Warpage

During the molding process of EMC materials in which the temperature exceeds the reaction temperature, the resin and hardener transforms into a densely cross-linked three-dimensional network from their initial oligomeric mixture which is shown in **Figure 2.13** (a) [29,57]. Molecular mobility in the network decreases with increasing cross-linking density as seen in **Figure 2.13** (b), leading to a decrease in reaction speed [29]. Once the increased cross-linking density hinders the molecular mobility to a critical level that the material is not able to flow anymore, the so-called *gelation* occurs which is illustrated in **Figure 2.13** (c). The material is not processable for molding operations beyond this point [57]. Furthermore, as the cross-linking level gets closer to fully dense structure, the reaction speed starts to decrease dramatically as a result of the very limited molecular mobility and low amount of free monomers left as shown in **Figure 2.13** (d) [58]. Due to this decrease in reaction speed, the cycle time for molding is usually kept shorter than the required time to reach desired high

cross-linking densities with economic concerns [59]. The parts are taken out of the mold once they reach a sufficient hardness for proper ejection, but their properties at this state are not optimal due to low degrees of cure [12,60]. Therefore, post-mold cure (PMC) is an essential step in electronic packaging to achieve the required cross-linking density, as well as the required thermal, dimensional, and chemical stability of the encapsulant material [12]. In PMC step, the part is exposed to elevated temperatures again after the molding process to ensure higher cross-linking densities of the molding compound.

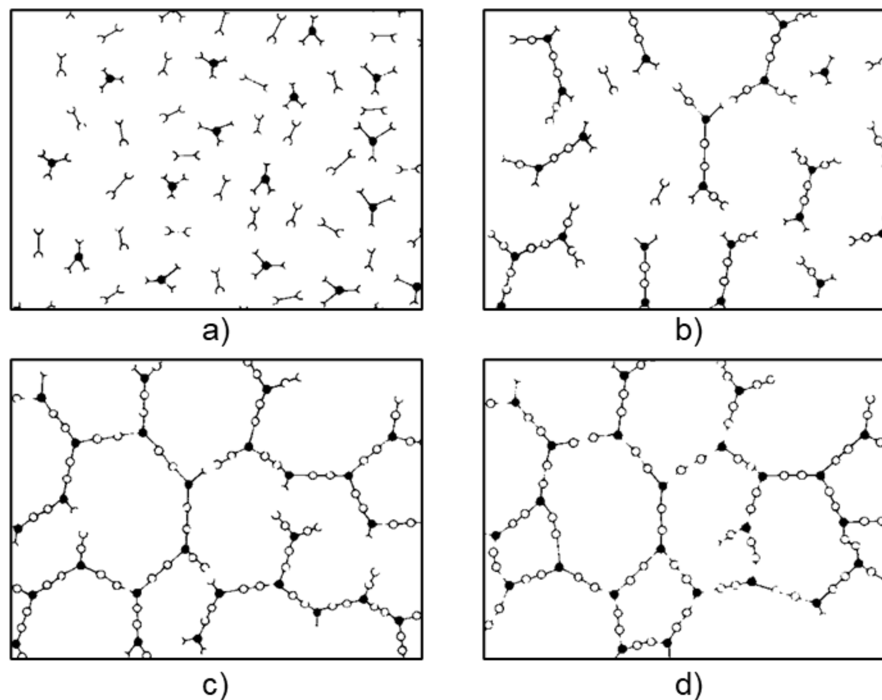

Figure 2.13: 2D representation of thermoset curing a) unreacted monomers b) formation of small-branched molecules c) gelled but incomplete cross-linked network d) cured network with almost no free monomers [57]

Ko and Kim [59] showed that the glass transition temperature, mechanical and adhesion strengths of EMC also increase after PMC with further increase in cross-linking density. With the mentioned changes during the process, the optimization of these material properties which were not completely fulfilled during molding is made possible through PMC [29,61]. In addition to optimization of mechanical properties, alleviating the resultant warpage is also amongst the goals of PMC in some cases [11].

The glass transition temperature (T_g) is a temperature range in which the molecular mobility increases due to the transition from the glassy state to gel/rubbery state of the material during heating [62]. **Figure 2.14** shows the time-temperature-transformation

(TTT) diagram that illustrates the different material states during curing. As the cross-linking density increases, the T_g of the curing thermoset increases from its initial value T_{g0}. If T_g exceeds the cure temperature T_c at some point during reaction, the reaction speed decreases dramatically since the diffusion rate drops suddenly due to the transformation from gel/rubbery state to glassy state. This dramatic decrease in molecular mobility and reaction speed is called vitrification [63]. Devitrification of the material is possible via further heating, and the cure can resume for partially cured thermosets after devitrification [57]. When the cure temperature and cure time are high enough, the cross-linking density will reach its highest limit with the final reachable glass transition temperature of $T_{g\infty}$ [62].

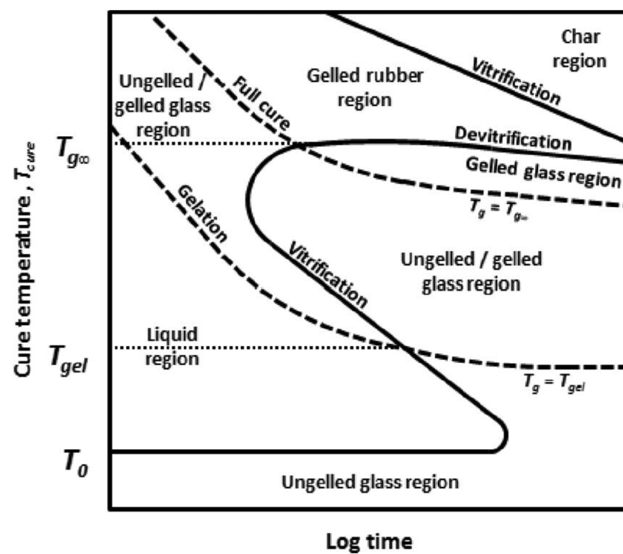

Figure 2.14: Time-temperature-transformation diagram for a thermosetting system [62]

The above-mentioned shift in T_g during PMC is a key point to decrease the manufacturing induced warpage of EMC encapsulants. **Figure 2.15** shows the effects of increased T_g on shrinkage. Assuming a process in which a part is transfer molded and then post-mold cured with the same temperature T_{cure} for both steps, the glass transition temperature of the EMC increases from T_{g0} to T_g during molding, and then from T_g to T'_g during post-mold cure. Since a big portion of the cross-links are formed during the molding operation, the cure shrinkage of the EMC is lower during PMC in comparison to its molding operation (see **Figure 2.11** for reference). $\Delta\nu_1$ being the thermal shrinkage after molding process while cooling back to room temperature before the PMC operation, it can be seen that the part is subjected to a smaller

shrinkage of $\Delta\nu_2$ after PMC due to earlier transition back into glassy state with a lower CTE value of α_1. As a result of lower shrinkage in EMC encapsulant, the thermal mismatch in between EMC and other package components with lower CTE values is compensated to some extent. Thus, the final warpage of the part is decreased. [9]

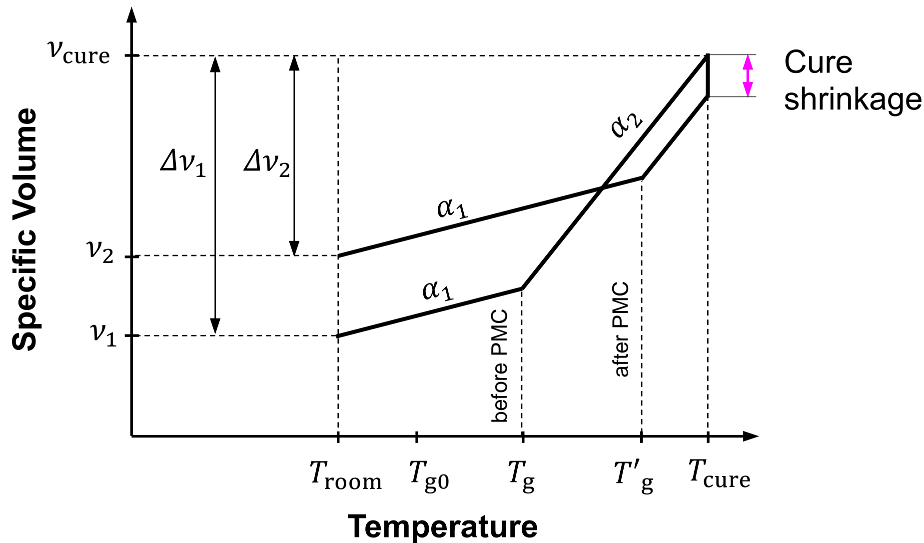

Figure 2.15: Specific volume change during PMC - adapted from [9]

Reducing the effects of thermal mismatch is also favorable for minimizing interfacial delamination [7]. As shortly mentioned in Chapter 2.1.5, additional thermal stresses are created during soldering operations of packages in the assembly step. Due to these additional stresses which occur while cooling back to ambient temperature after the soldering operation, interfacial delamination and encapsulation cracks may occur [8]. It was shown that the occurrence of delamination and cracks during this step is decreased with post-mold cure as a result of reduced thermal mismatch [59]. Increased adhesion and mechanical strengths are also benefits of PMC that increases the product reliability for further assembly operations and usage after manufacturing [64].

The optimal duration for post-mold cure is an open area for improvements in the industry. The common practice is to post-mold cure the samples in the temperatures around $T_{g\infty}$ for long times up to 16 hours to reach the ultimate $T_{g\infty}$ value for thermal stability [12]. But it was shown that post-mold cure operations with several hours do not have significant effects on warpage and material properties since the biggest portion of the curing reaction takes place in the very first one hour interval [59]. Longer PMC times might even have disadvantages due to polymer ageing or degradation [65].

2.2 Machine Learning

Machine learning is the study of statistical algorithms that enable the computers to simulate human learning activities [14]. As a more engineering-oriented definition, a computer is not programmed for a specific task, but it is programmed to learn for a task via extracting the underlying patterns from the data it was trained before [66].

Machine learning has been a major mainstay of information technologies over the last 30 years with increasing amount of data and newly developed algorithms being available [18,67]. It has been widely used in different areas such as finance, telecommunications and marketing with the purposes of classification and forecasting [14]. Spam e-mail recognitions, traffic jam predictions in online maps, personalized advertisements popping up in our devices are all examples of machine learning applications that are highly adapted into our daily lives [68]. With the advances in monitoring and other data collection techniques in medical applications, machine learning has become a well suited approach for quick and specialized diagnosis tasks also in the field of medicine [69]. Improved image processing techniques with machine learning are used for face recognition and autonomous driving [70]. Focusing on the industrial production, there is an increasing trend of machine learning usage for product and production optimization due to its capabilities on optimizing resource usage rates and production times [15].

With a simple approach, the workflow of a machine learning model can be divided into 5 main steps as data collection, data preprocessing, model building and training, model validation and execution as shown in **Figure 2.16** [14,71].

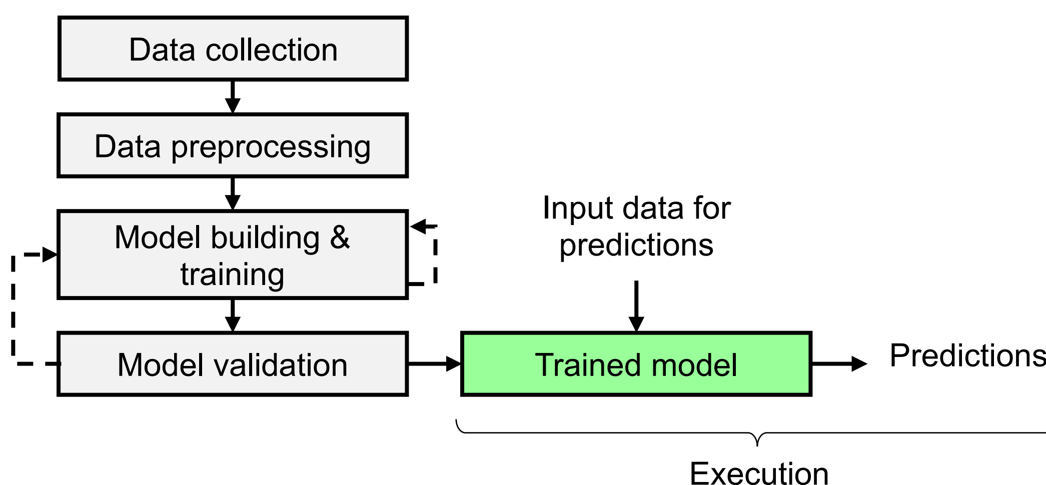

Figure 2.16: Simplified workflow of a machine learning model

The first step of creating a machine learning model is data collection. The data-driven decision making and prediction capability of a machine learning model functions by the mathematical algorithms which are built over the data that is used for training [72]. Thus, the key point to a sound machine learning model is having a reliable data set to fit the mathematical algorithm. The quantity and quality of the data set directly affect the accuracy of the final predictions [73].

A data set consists of multiple data points each representing an entity to analyze. Each point has the so-called features that describe the properties, such as the age, height, weight, and gender of a patient; or production parameters and materials used in an experimental sample. A complete set of features for a data point forms the feature vector for that point, and each one of these features comprises one dimension of the so-called feature space. The corresponding output values or classifications might also be included in the data set, such as the presence of weight abnormality in the patient or the resultant warpage values of the experimental samples. The classes or outputs for the given features are commonly called the labels. [67]

Since acquiring a flawless set of data is not possible in real life, the preprocessing step plays a critical role for the reliability of the model [74]. When the available data is irrelevant, redundant, or noisy, the model training step gets more difficult and inaccurate. Thus, preprocessing operations such as data cleaning, normalization, feature extraction and feature selection might be applied for better results [75].

Once the required data reliability is ensured, the following step is the training of the model. During training, the model extracts the relationship between the features and outputs (labels) of the train data if they are available. The weights and biases of the mathematical algorithm are updated iteratively to reach higher accuracy [74]. According to type and availability of data, different types of algorithms can be used in the model [66].

The next step before the final execution of the model is the validation phase in which the trained model is tested with unseen data. The trained model might make good predictions for the train data but the aim here is to validate how good the model predicts for unseen data points. Once the model validation is completed, the final step is to execute the trained model and predict for unseen data from real-world applications [74].

In this sub-chapter, the fundamentals and the state of the art for machine learning are explained. Starting with introducing the main algorithm types used in machine learning in Chapter 2.2.1, the model validation concept and common validations methods are explained in Chapter 2.2.2. Two common supervised learning algorithms, namely Partial Least Squares Regression (PLRS) and Artificial Neural Networks (ANN) are described in Chapters 2.2.3 and 2.2.4 respectively.

2.2.1 Types of Machine Learning Algorithms According to Data Availability

Machine learning algorithms can be divided into four main subcategories as supervised learning, unsupervised learning, semi-supervised learning and reinforced learning according to type and availability of the train data as shown in **Figure 2.17** [68].

Supervised learning methods consist of algorithms that use provided labels to train the data, while unsupervised learning algorithms aim to extract the similarities in features without the labels [66,67]. More details about these two methods are given in the following sub-chapters. Semi-supervised learning is a combination of both methods in which the data set has both labeled and unlabeled data [66]. Reinforcement learning algorithms on the other hand, are based on trial and error experiences in which the algorithm captures the reoccurrence patterns and create generalizations related to experienced occurrences [76,77].

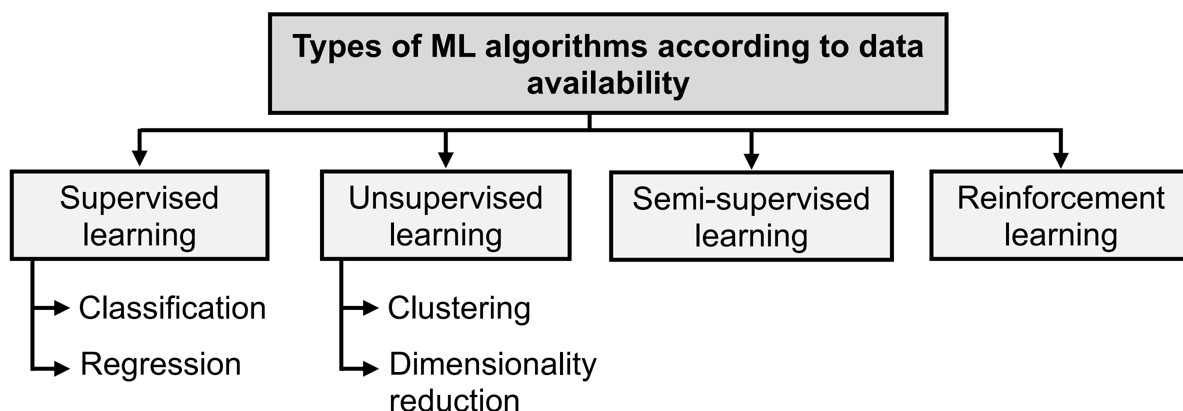

Figure 2.17: Overview of different types of machine learning algorithms [68]

Supervised learning

The term *Supervised learning* refers to the applications in which the model is trained with the labeled data, which means the true outputs (labels) are provided for each point in the train data set [66]. The goal of the algorithm in such cases is to create a rule by

extracting the patterns in the train data. After detecting the underlying pattern in train data, it is then made possible to predict for the data points with unknown labels [78].

Classification and regression are two main application areas of supervised learning which have categorical classes and numerical values as outputs respectively [67]. An example regarding classification with supervised learning is illustrated in **Figure 2.18**. Each point in train data has two features and a class as seen in **Figure 2.18** (a). The features for the test points are known (**Figure 2.18** (b)) and the model aims to predict the missing class information for these points as shown in **Figure 2.18** (c) [66].

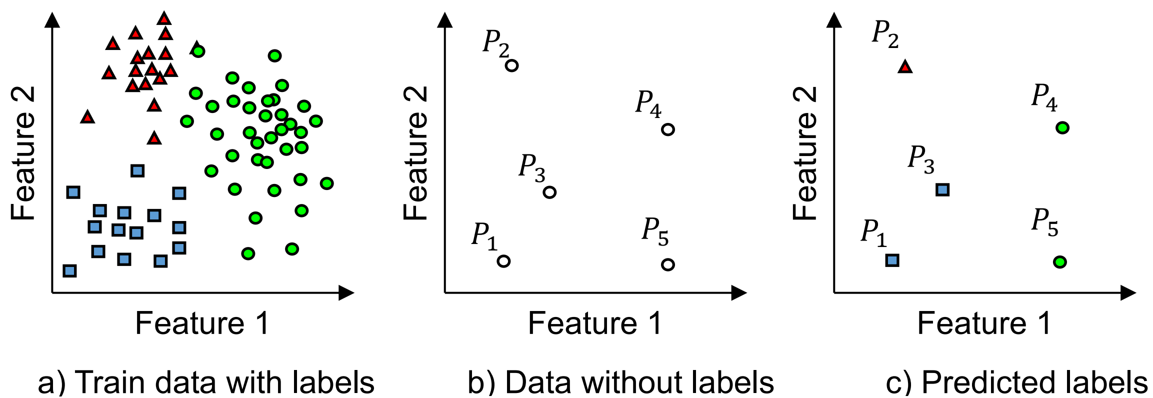

Figure 2.18: Classification algorithm a) train data with labels b) external test data without labels c) predicted labels - adapted from [66]

Curve fitting is another simple supervised learning example with the target of numerical outputs [66]. **Figure 2.19** shows an example in which the model fits a curve to train data with known labels and then predicts for an external data point.

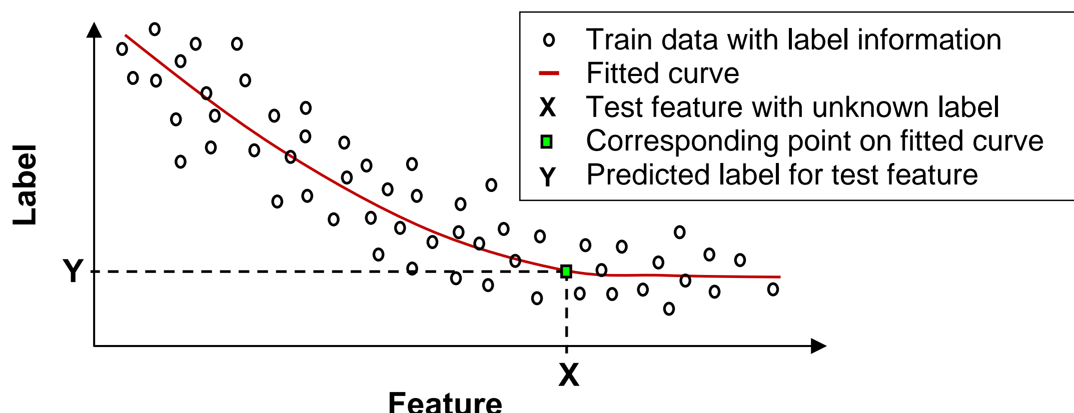

Figure 2.19: Regression through curve fitting - adapted from [66]

There are many different algorithms that can be used for supervised learning applications. Some common supervised learning algorithms for classification are called

Decision Trees, *Naive Bayes* and *Random Forest*. For regression on the other hand, *Support Vector Machines* (SVM), *Linear Regression*, *Partial Least Squares Regression* (PLSR) and *Artificial Neural Networks* (ANN) are commonly used [67,79]. Chapter 2.2.3 gives an overview about usage of PLSR and Chapter 2.2.4 describes the fundamentals and mathematical basis of ANN algorithms.

Unsupervised learning

In many real-life situations, labeling the data set is either difficult or not possible, therefore it is always a bottleneck to provide labeled data for machine learning applications [80]. In such cases, unsupervised methods are used to determine similarities and differences in between the data points without label information [81]. The sub-groups with similar features are called clusters and determining these feature affinities may consequently be helpful in finding unknown rules and correlations in a data set [82]. With reference to the clustering behavior, the normal behavior of a system can be represented, thus anomalies in the data set can be detected [66]. The clustering behavior can even be utilized for classification purposes without the label information [83].

As a simple example regarding clustering, an unlabeled data set with two attributes as Feature 1 and Feature 2 can be seen in **Figure 2.20**. In this example, the aim of the model would be grouping the similar types of data by detecting the clustering behaviors [66,83]. As the number of features increase, it gets harder for human perception to detect the clusters. That is why an algorithm-based approach is useful for such tasks with high-dimensional feature spaces [83].

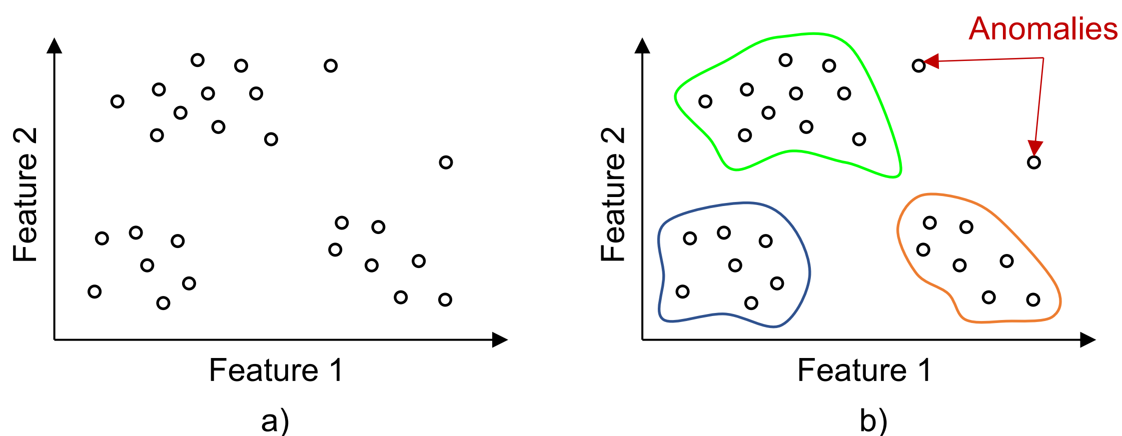

Figure 2.20: A two-dimensional domain with a) unlabeled clustered data points b) extracted clusters and anomalies - adapted from [66]

When working with high-dimensional data, unsupervised learning is widely used. Being able to detect the effects of each feature on clustering behavior, they can be put in a hierarchical order. The features that affect clustering the most can be extracted, while the features of low importance are being excluded from the feature space. Or some features with similar effects can be merged to one [67]. **Figure 2.21** shows a cluster that is mapped on a 2D space which consists of features that show similar proportional trends. To take advantage of this similarity and simplify the data set, Feature 1 and Feature 2 are projected on a regression line and merged as Feature´ [82]. This operation is called dimensionality reduction and aims to simplify the data set without losing too much information. With the simplified data set, the algorithm will run faster and the data will take less storage space [66]. In some cases, when the number of data points (or train samples) is less than the dimensions of the data, the implementation of regression models might fail due to the phenomenon called *Curse of Dimensionality* [84]. Therefore, it is a useful preprocessing technique that can increase both accuracy and productivity of the model even though the labels are known for the data set [66,85].

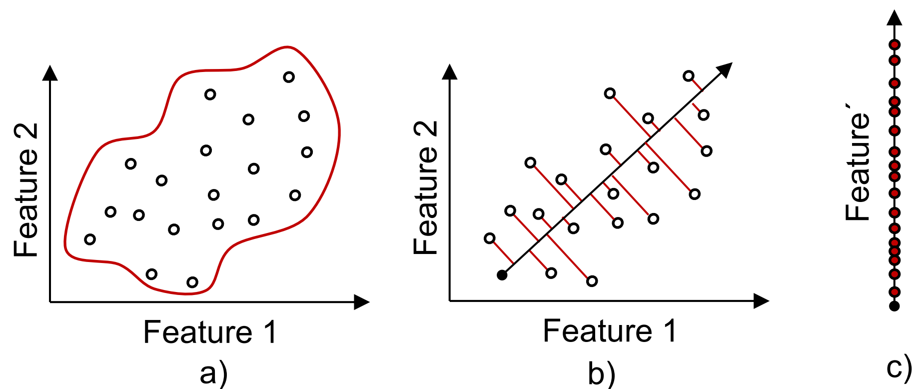

Figure 2.21: Dimensionality reduction via merging a) clustering data points in a 2D feature space b) point projections on regression line c) reduced feature space with projected features [82]

The common algorithms used for unsupervised learning include *K-means* and *Decision Trees* for clustering [86]. For dimensionality reduction on the other hand, *Principal Component Analysis* (PCA) and *t-distributed Stochastic Neighbor Embedding* are amongst the frequently used algorithms [67]. In this work however, supervised learning algorithms are used, and the focus will be on supervised methods from now on.

2.2.2 Model Validation

The trained model might have different conditions according to how well it fits to train data [67]. Underfitting is the case where the model is not trained good enough as illustrated in **Figure 2.22** (a) due to low number and quality of the train data or simple algorithms used [67]. Overfitting on the other hand, is the opposite problem in which a model performs really well on the samples from the training set but gives bad predictions for other external samples as illustrated in **Figure 2.22** (b) [87]. It is therefore important to evaluate the model with external data as well to ensure a good fitting model as in **Figure 2.22** (c) [88].

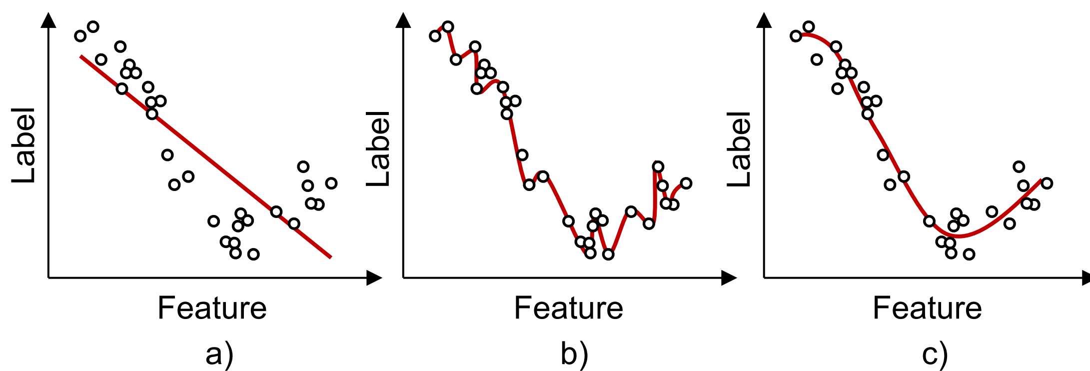

Figure 2.22: Schematic illustration of a) underfitting b) overfitting c) good fit [67]

A common practice to detect and avoid overfitting is splitting the available data, and holding a portion out of the training steps for further validation of the model with the unseen portion [89]. The portion that is used for building the model is called the train set. Train set contains samples with known labels. The validation set contains samples from the known provenance as well, but it is not used for training of the model. The model is then expected to perform similarly for train and validation data. Thus, the predictions for the validation set allows the accuracy assessment of the model [87]. The optimal ratio in between the train and validation data is still an area without a consensus despite the theoretical and numerical investigations. A commonly used ratio is 80 % and 20 % for train and validation portions respectively [89].

For many applications, the model assessments that are conducted only via validation data yields over-optimistic results as shown by Westerhuis et al [90]. Since higher order nonlinear functions are able to fit well to a data provenance, the risk is to create a model that fits well to data from the same provenance but performs poorly for the unseen data from separate data groups [88]. As the train and validation groups belong

to the same provenance, there is a chance of having a biased model assessment [88]. Therefore, having an additional unseen test data which has a good distribution over different input regions will ensure a better assessment of model performance also for different types of input data [88].

2.2.3 Partial Least Squares Regression

Partial Least Squares Regression (PLSR) is a method that includes regression, classification and dimensionality reduction techniques [91]. Following the same basis with Principle Component Analysis (PCA) which aims to reduce the dimensionality as shortly mentioned in Chapter 2.2.1, PLSR aims to predict a set of dependent outputs variables from the input set of independent variables via extracting the so-called *latent* variables, which stand for the input variables that has the biggest effect on the resultant outputs [92]. The extraction of latent variables reduces the multidimensionality of inputs to a lower number, thus allows predictions with a lower number of inputs [93]. For sake of simplicity, the governing equations for PLS are not introduced.

As a supremacy against linear and multiple linear regression, PLSR can analyze strongly collinear, noisy and numerous input variables [94]. Due to this allowances, PLSR has gained a big attention in chemometrics, where wide spectral chemical data is being used [91].

2.2.4 Artificial Neural Networks

Artificial neural networks (ANNs) are the mathematical representations of cerebral cortex part of animal brains. The structure and learning activities of the brain synapses are imitated via the large number of interconnecting elements and the mathematical functions assigned to these elements. [95]

The studies for ANNs began in early 1940’s with the studies of McCulloch and Pitts [96], but only in the last quarter of 20th century the algorithms became suitable for real-life applications [97]. Nowadays, the ANNs are commonly used in a lot of real-life problems. Pattern and speech recognition, financial forecasting, image processing applications are amongst the most common usage areas [98]. Also, focusing on industrial applications, ANNs have been successfully used for optimization and modeling of variety of industrial production processes [34,99,100].

Structure of an Artificial Neural Network

A neural network works as a series of nonlinear functions that transform a set of input values to a set of output values [98]. This transformation is governed via the parameters called synaptic weights and biases of the so-called neurons in the structure [101]. The determination of these weights and biases is called learning or training [101]. A neuron with multiple inputs and a single output is illustrated in **Figure 2.23**.

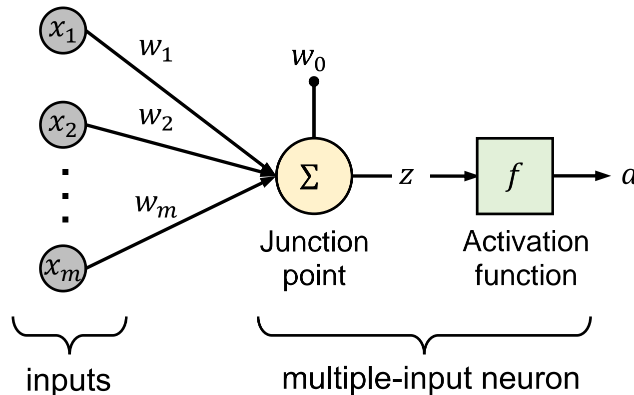

Figure 2.23: Schematic of a single artificial neuron with m inputs [88]

The inputs $x_1, x_1, \ldots, x_m$ are weighted by their corresponding elements $w_1, w_2, \ldots, w_m$ and the bias of the neuron w_0 is summed with the weighted inputs. The output from the junction point

$$z = \sum_{i=1}^{m} x_i \, w_i + w_0 \tag{2.1}$$

is then transformed by a nonlinear activation function f to form the output of a.

$$a = f(z) = f\left(\sum_{i=1}^{m} x_i \, w_i + w_0\right) \tag{2.2}$$

The main goal of activation functions is to introduce a nonlinearity so that the algorithm is able to tackle complex nonlinear problems as well. **Figure 2.24** illustrates some activation functions such as (a) *Sigmoid*, (b) *Rectified Linear Unit* (ReLU), and (c) *tanh* [102]. Normalization of the input data during preprocessing plays an important role at this step [103].

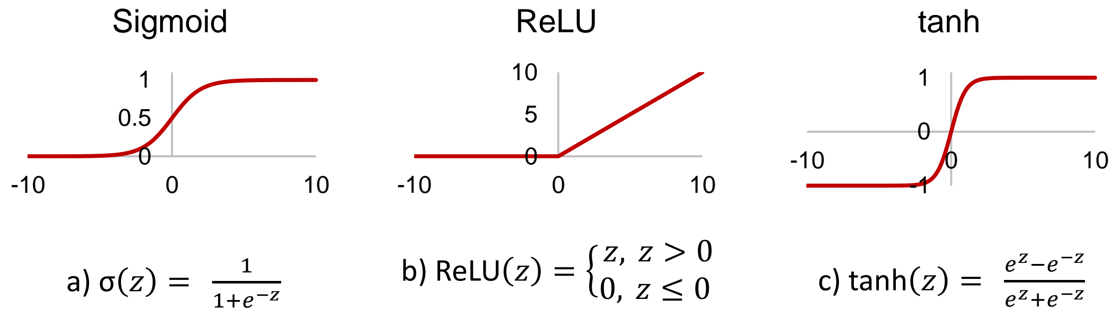

Figure 2.24: Commonly used activation functions a) sigmoid b) ReLU c) tanh [102]

An artificial neural network is built by organizing the neurons in layers. The layers in between input and output layers are called the hidden layers. The number of hidden layers and the number of neurons in each hidden layer assess the complexity of the network [101]. In feed-forward networks, the signal always flows from the input direction to the output direction. In other words, the output of a neuron is fed to the neurons of the following layer. [104]

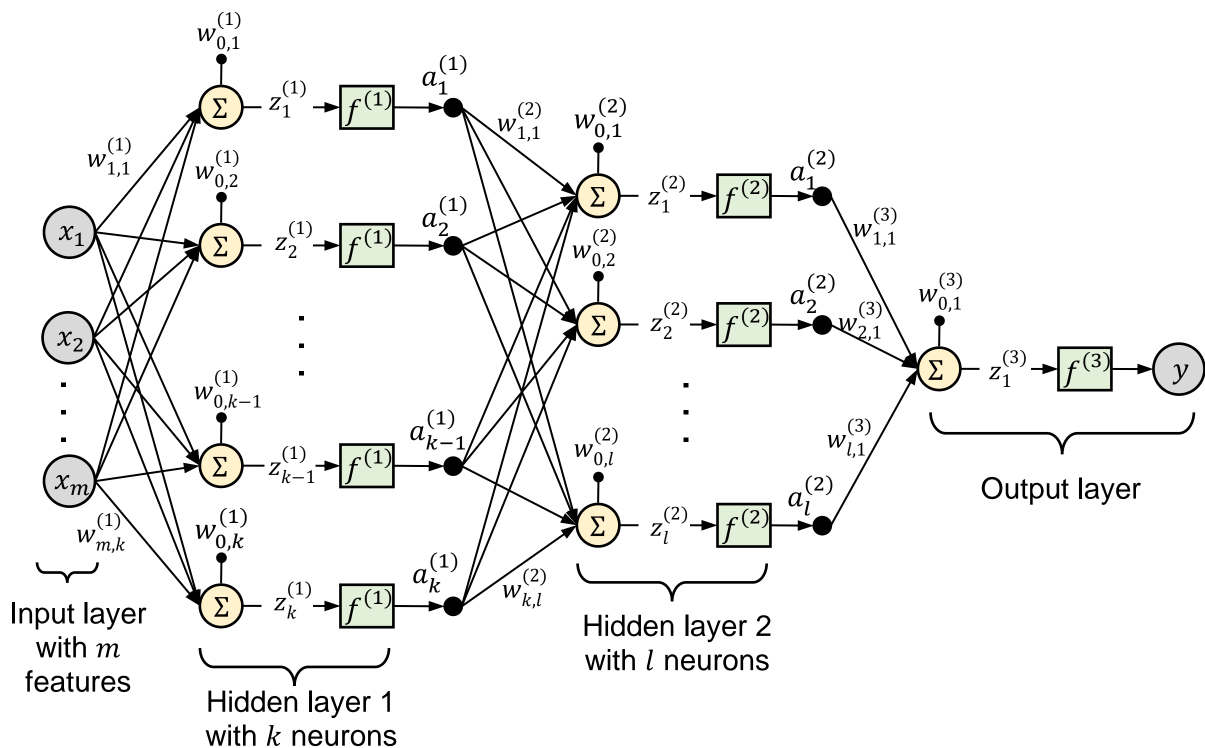

Figure 2.25: A multi-layered artificial neural network - adapted from [88]

Figure 2.25 illustrates a 3-layer artificial neural network with an input layer of m features, two hidden layers of k and l neurons, and an output layer of a single neuron. With reference to Equation (2.2), the output y can be expressed as

$$y = f^{(3)}\left(\sum_{i=1}^{l} a_i^{(2)} w_{i,1}^{(3)} + w_{0,1}^{(3)}\right) \tag{2.3}$$

where $f^{(3)}$ is the activation function used in the last layer, $a_i^{(2)}$ and $w_{i,1}^{(3)}$ are the outputs and weights of the ith neuron from previous the layer respectively. Networks with more layers can be built following the same principle, and the equation can be expanded using the open notations from previous layers.

Training Artificial Neural Networks

The terms training or learning for ANN refer to the problem of minimizing a defined loss function [105]. The mentioned loss function is characterized by the error in between the true and the predicted outputs for the train data. With reference to Equation (2.3), one can derive that the prediction error of the network is a function of the weights. The minimization of the loss function $E(w)$ is done iteratively and the weights of the network are updated after each iteration until reaching a defined number of iterations or an error rate [34,106].

During training, the first phase is called the *Forward pass* in which the outputs are predicted through Equation (2.3). The weights are assigned randomly for the first prediction [106]. The error value for the prediction is then calculated via the defined loss function. A common function used to assess the error in such cases is the *Sum of Squared Errors* (SSE) that can be expressed as

$$SSE = \frac{1}{2}\sum_{i=1}^{h}(y_i - y'_i)^2 \tag{2.4}$$

where h is the number of train data, y and y' are predicted and true output values respectively [106,107]. *Mean Absolute Error* (MAE) is another evaluation method which calculates the average absolute error in between true and predicted values as

$$MAE = \frac{1}{h}\sum_{i=1}^{h}|y_i - y'_i| \tag{2.5}$$

with the same notation as in Equation (2.4) [108].

Following the forward pass, the partial derivative of the loss function is calculated with respect to regression weights. This phase is called the *Back Propagation* [109]. To reach a local minimum of the loss function, the weights are updated in the opposite direction of the gradients, using the *Gradient Descent* method [106]. Every iteration cycle to update the weights is called an *epoch* [109]. **Figure 2.26** shows a simplified gradient descent example for a univariate loss function, the index t stands for the iteration number of the weight. When the derivate of the loss function is negative, the weight is updated in positive direction in the next iteration. Similarly, the update direction is positive when the derivative is negative [106].

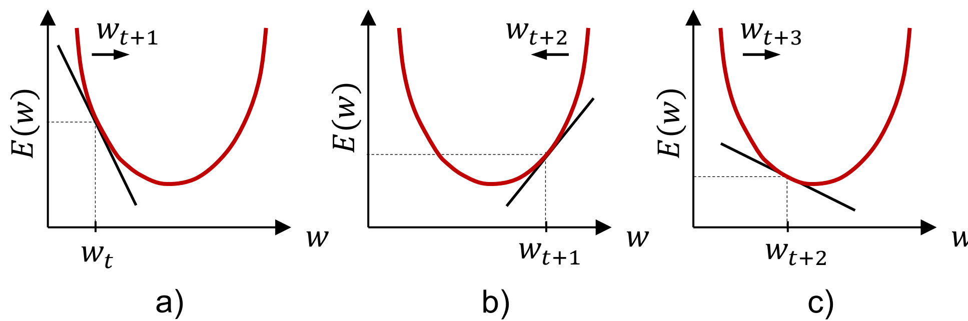

Figure 2.26: Basic idea of back propagation a) initial weight with negative derivative b) first iteration in positive direction c) second iteration in negative direction – adapted from [106]

The gradient of the differentiable loss function of a neural network is expressed as

$$\nabla E = \left(\frac{\partial E}{\partial w_1}, \frac{\partial E}{\partial w_2}, \dots, \frac{\partial E}{\partial w_n}\right) \tag{2.6}$$

where n is the total number of weights in the network. As seen, loss functions of the networks in real applications are composed of many weight variables. Therefore, chain rule is used for ease of calculations [110]. The difference in between the weight index n in Equation (2.6) and the iteration index t shown in **Figure 2.26** should be kept in mind at this point to avoid misunderstandings. The index n identifies the position and layer information of different weights that are connected to neurons in a network, whereas the index t identifies the iteration number for each one of these weights individually.

After each iteration, the weights are updated using the gradient descent method

$$\Delta w_{t+1} = -\eta \frac{\partial E}{\partial w_t} \tag{2.7}$$

$$w_{t+1} = w_t + \Delta w_{t+1} \tag{2.8}$$

where η is the so-called *Learning Rate* constant and t is the iteration index of the weight [110]. Learning rate is a constant that adjusts the iterative steps [107]. Higher learning rates increase the training speed, but the chances of finding the optimum point decrease [107]. Increased oscillation in iterative steps can also bring the risk of diverging from the optimal point [107]. Too low values on the other hand, lead to a slow and costly training and might end up with stagnating in a local minimum which is actually far from the global minima of the loss function [107].

To optimize the network training, additional terms can be included in gradient descent to tackle the above-mentioned issues with learning rate. A common method is adding a so-called *Momentum* term β, which is weighted with the weight change from the previous iteration step as

$$\Delta w_{t+1} = \beta \Delta w_t - \eta \frac{\partial E}{\partial w_t} \tag{2.9}$$

$$w_{t+1} = w_t + \beta \Delta w_t - \eta \frac{\partial E}{\partial w_t} \tag{2.10}$$

to smooth the oscillations in between iterations.

2.3 Summary & Motivation

Electronic packages are on high demand due to increased electronics usage in a broad scale of products. Encapsulation is a major process in electronic packaging since it directly affects the product's final quality and reliability. Molding is the dominant method used today for encapsulation and EMCs are the most common polymer materials used for producing molded encapsulants to electronics.

PMC is a common process in electronic packaging industry that is applied following the molding processes to reach increased properties of the encapsulant material via further curing. PMC is mostly conducted with the same parameters in industry, aiming for the highest possible cross-linking density. It is known that warpage is affected by PMC, but there are not many studies investigating the effecting mechanisms. Thus, the effects of different PMC parameters on resultant part warpage are not well-known.

Due to the reactive nature of EMCs, it is always a challenge to adjust the process parameters accordingly. As a contributing factor to these challenges, the properties of EMCs inevitably deviate for each batch of delivery from the material manufacturers. Despite being used for many years, transfer molding and post-mold cure processes are still far from being optimal. Most of the time, the parameter adjustments are done in trial-and-error manners. There is a search for alternative process optimization and adjustment methods due to these challenges. Being a proven method for such purposes, machine learning methods are successfully applied to industrial optimization problems.

With reference to above mentioned reasons, this thesis aims to investigate the effects of different post-mold cure parameters on the warpage of EMC encapsulated electronic packages and create a process knowledge in the light of the conducted experiments. Due to the challenges of EMC process simulations, it is aimed to propose a machine learning based algorithm to optimize the PMC parameters with the main concern of reducing the final warpage.

3 Materials, Instrumentation and Methods

This chapter describes the materials, the methods, and the instruments used in this work. Aiming for investigating the effects of different post-mold cure parameters on warpage in electronic packaging processes, sets of experiments are conducted. **Figure 3.1** illustrates the overall steps and workflow followed during this study. Material characterization tests are carried out to have an understanding of thermal and cure behaviors of used materials. Experiments with different PMC parameters are conducted with the goal of creating an overall understanding of the warpage reduction mechanism during PMC and correlating the resultant warpage values with the parameters used. Moreover, with the data set created from the experiments, a machine learning (ML) algorithm is created that predicts for the resultant warpage and cross-linking density after post-mold cure, and then optimizes the parameters with reference to predicted values.

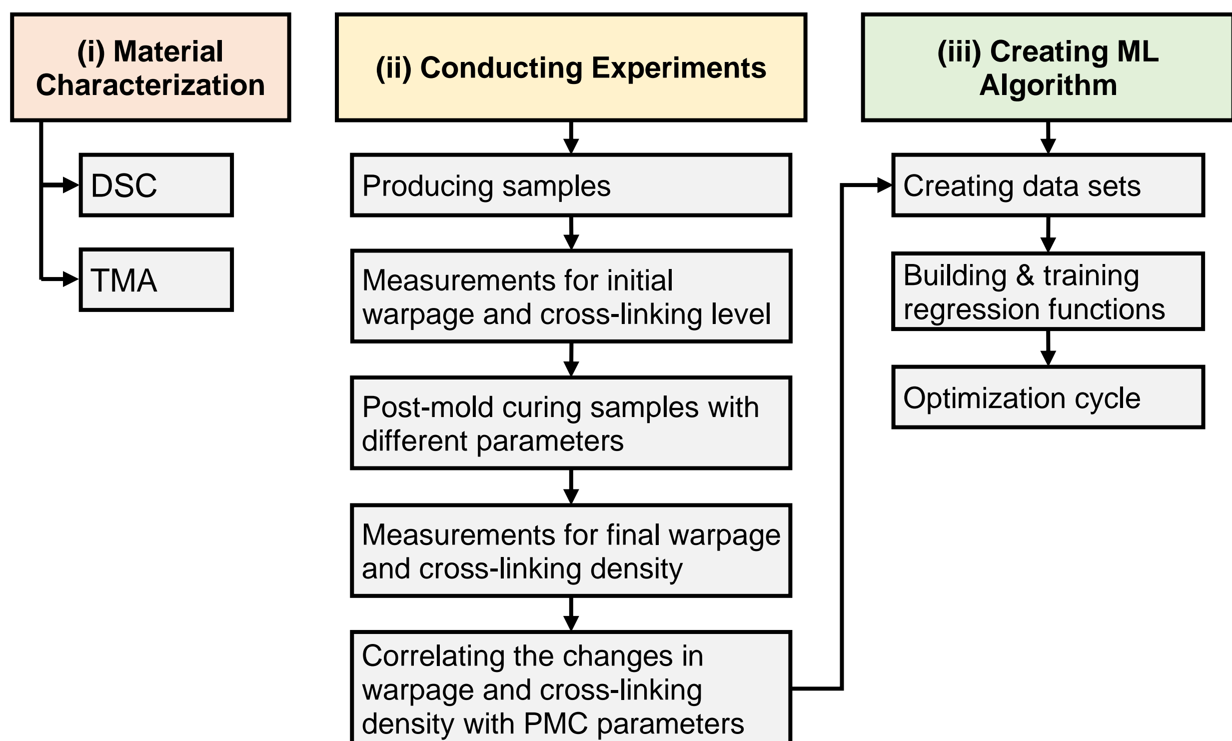

Figure 3.1: Overall steps and workflow of the thesis

Starting with Chapter 3.1, which introduces the two main types of materials used in the experiments, namely the epoxy-molding compound and the printed circuit board substrate; Chapter 3.2 describes the Differential Scanning Calorimetry (DSC) and Thermomechanical Analysis (TMA) tests which are performed to characterize their

thermal and cure behaviors. Following that, In Chapter 3.3, the overall workflow of the PMC experiments is described including the production of samples, conducted measurements for their warpage and cross-linking density values, the post-mold cure operations, and the experiment plan with different PMC parameters. Finally, the principles of the machine learning algorithm are explained in Chapter 3.4, including the subfunctions used for predicting the warpage and cross-linking density.

3.1 Materials

The samples that are used in the experiments consist of two main components. Namely, epoxy molding compound (EMC) and the printed circuit board (PCB) substrate as shown in **Figure 3.2**. The EMC is molded over the PCB substrate to form the encapsulant structure.

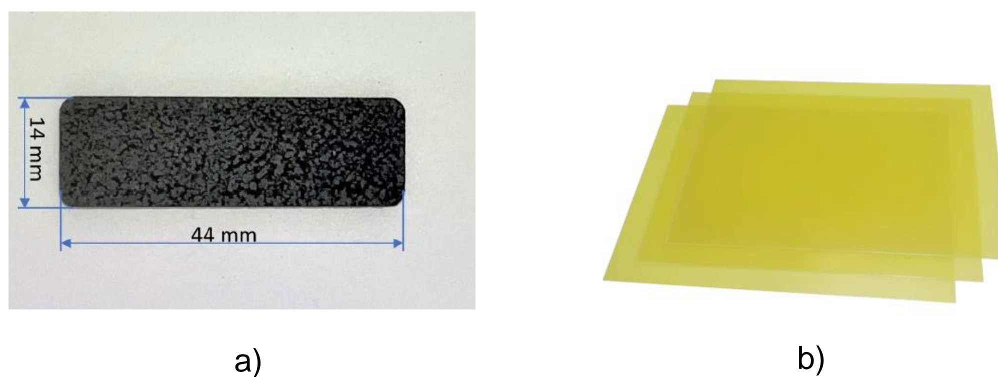

Figure 3.2: Main materials used in samples a) EMC granules pressed in sheet form b) PCB substrates without mounted electronics

The encapsulant material EMC consists of a highly filled epoxy resin with ~90 % filler content. The exact composition is not known since it is not published due to material manufacturers policies. It is available in granular form and pressed into sheets before molding as shown in **Figure 3.2** (a). The granules and the pressed sheets are stored in a freezer to avoid further curing reactions that might take place at room temperature.

The PCB substrate which is shown in **Figure 3.2** (b) is a single layer glass epoxy laminate with a thickness of 200 µm, which normally includes mounted sensor elements to be encapsulated. These elements are not included in the substrates used in the experiments for research purposes.

3.2 Material Characterization

This chapter explains the tests conducted for characterizing the thermal properties and cure behaviors of EMC and PCB substrate. Differential Scanning Calorimetry (DSC) measurements are conducted to assess the cure behaviors, degrees of cure and T_g values of EMC encapsulants with different cross-linking densities. In addition, to determine the T_g value and to inspect any possible curing reactions of glass epoxy, DSC measurements for PCB substrate are carried out as well. Chapter 3.2.1 gives a brief introduction to DSC method and explains the measurement details. As the second method due to some insufficiencies of DSC, Thermomechanical Analysis (TMA) is used to estimate the CTE and T_g values of fully cured EMC, which is described in detail in Chapter 3.2.2.

3.2.1 Differential Scanning Calorimetry

Differential Scanning Calorimetry is a thermal analysis technique that is commonly used to study the cure kinetics of thermosets [111]. For the measurements, the test and reference materials are placed in crucibles (mostly made of Aluminum) and then positioned on the heaters in a heating chamber [112]. An empty Aluminum crucible can be used as a reference as well [112]. According to ASTM Standard E473-22d [113], the difference in between the heat flows into test and reference samples is measured as a function of temperature while they are being exposed to a controlled temperature program. Important thermal properties such as the reaction start temperature, glass transition temperature T_g, melting temperature T_m and reaction enthalpy ΔH can be determined via DSC [114]. An example DSC curve is shown in **Figure 3.3**.

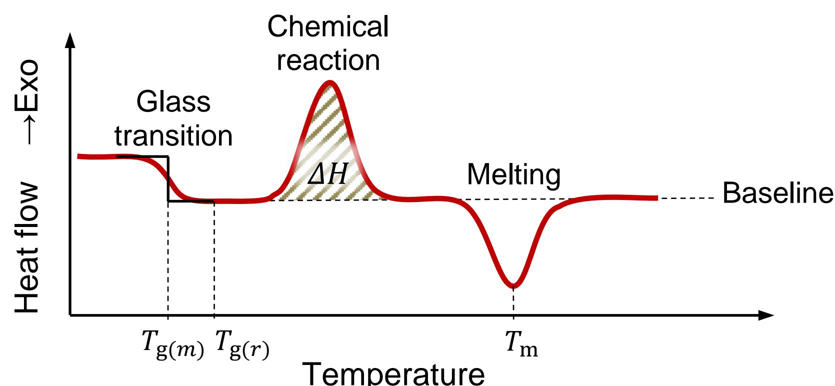

Figure 3.3: An example DSC curve - adapted from [112,115]

The degree of cure α of a thermoset can be calculated via the heat flow data acquired from the DSC curves as an indicator of cross-linking density [111]. The reaction rate $d\alpha/dt$ is assumed to be directly proportional to heat flow $dH(t)/dt$ [111]. Under this assumption, the degree of cure can be expressed as

$$\alpha = \frac{1}{\Delta H_{\text{total}}} \int_0^t \left(\frac{dH(t)}{dt}\right) dt \tag{3.1}$$

where ΔH_{total} is the total reaction enthalpy of the non-cured reactant during a full curing reaction from its non-cured state [5,57]. It should be noted that $dH(t)/dt$ corresponds to the ordinate of the DSC curve when plotted over time [57]. The reaction enthalpies of the samples are calculated via the area under the exothermic peaks of DSC curves, and then used as a measurable indicator of cross-linking density with reference to Equation (3.1). Moreover, the T_{g} values at the given states are estimated with reference to ASTM Standard E1356-08 [115]. Even though the midpoint temperature $T_{\text{g(m)}}$ is used more often in practice to determine glass transition temperature via DSC, return-to-baseline temperature $T_{\text{g(r)}}$ of the samples are also evaluated since it indicates the complete transition into gel state which is an important step for warpage reduction during PMC [115]. Chapter 4.2.2 gives a more detailed explanation about this point.

The measurements are conducted with the device DSC 204 F1 Phoenix from NETZSCH-Gerätebau GmbH, Selb, Germany. A non-isothermal temperature program with a heating rate of 10 °C/min is used within a temperature range of 0 °C to 300 °C. Liquid nitrogen is used as coolant.

To characterize the curing behavior of EMC, DSC specimens with 30±2 mg masses are extracted from the molded samples with different molding parameters and then placed in 25 µl Aluminum crucibles for measurements. For the non-cured state of the material, the granules are cold pressed into 30±2 mg sheets without the molding operation and then stored in a freezer to avoid further material curing.

Another set of DSC measurements are conducted to inspect the glass transition behavior of PCB substrate and to detect any possible curing in the glass epoxy during post-mold cure. Specimens with 30±2 mg masses are extracted from the PCB substrates of molded parts, and DSC measurements are conducted before and after their PMC operations.

3.2.2 Thermomechanical Analysis

Thermomechanical Analysis (TMA) is described as a thermal analysis technique that measures the deformation of a substance as a function of temperature while it is under non-oscillatory load and subjected to a controlled temperature program [113]. During TMA, the sample is placed in a heated container and a probe is positioned above in a way that it applies a load to the sample as shown in **Figure 3.4**. The displacement of the probe during the heat program is measured via a high precision detector (usually a linear variable displacement transducer), which is then used to determine the deformation of the loaded sample [116]. TMA method can be used to determine a variety of material properties such as CTE, melting temperature T_{m}, and glass transition temperature T_{g}.

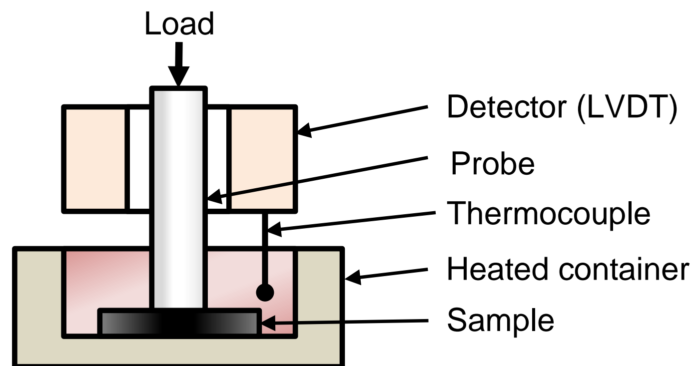

Figure 3.4: Schematic of a Thermomechanical Analyzer – adapted from [116]

In this work, the TMA method is used to obtain the T_{g} value of fully cured state of EMC since it is hard to obtain T_{g} information of highly cross-linked thermosets with DSC [116]. The measurements are conducted using the device TMA Q400 V22.5 from TA Instruments, New Castle, USA. A program with the temperature range -60 °C to 270 °C, a heating rate of 5 °C/min, and a compression load of 0.1 N is used. In accordance with ASTM Standard E831 – 19 [117], specimens are cut out from a 1 mm thick EMC wafer, which is molded with T_{mold} = 170 °C and t_{mold} = 120 s, and then post-mold cured with T_{PMC} = 175 °C for t_{PMC} = 4 h to reach an almost full cross-linking density of EMC. The cutting operation is conducted with air cooling to avoid further thermal effects despite the PMC operation. Since producing proper TMA samples out of granular EMC was not possible without molding, no TMA measurements are conducted for non-cured state of EMC.

3.3 Post-mold Cure Experiments

To inspect the effects of different PMC parameters on part warpage, sets of PMC experiments are conducted following the systematic illustrated in **Figure 3.1** (ii). Firstly, the samples are produced with the methods and details introduced in Chapter 3.3.1. Following the production, the IR spectra and initial warpage values of the samples are measured as described in Chapter 3.3.2. The samples are then post-mold cured with the considerations and parameter sets introduced in Chapter 3.3.3. Following the PMC operations, the measurements are repeated directly.

3.3.1 Sample Production

As the first step of experiments, the parts are produced to be post-mold cured later. This chapter introduces the special molding method that is used for molding the samples, followed by the explanation of preparation steps, and used molding parameters.

The experimental samples shown in **Figure 3.5** are produced using the method *Local Pressure Molding*, which is a combination of transfer molding and compression molding that enables covering a substrate with a mold material [118].

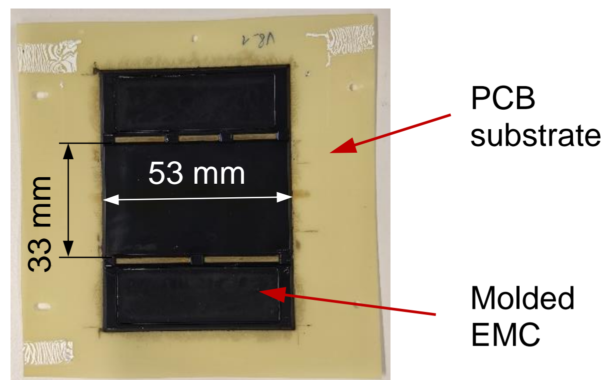

Figure 3.5: Molded sample

The hydraulic press machine that is used for molding operations (VSKO 25 from Lauffer GmbH & Co. KG, Horb am Neckar, Germany) consists of upper and lower heating plates as shown in **Figure 3.6**. The tool plates are assembled on heating plates. The lower plate houses the mold cavity and the compression stamp, which is geometrically divided into two main sections as rear and front.

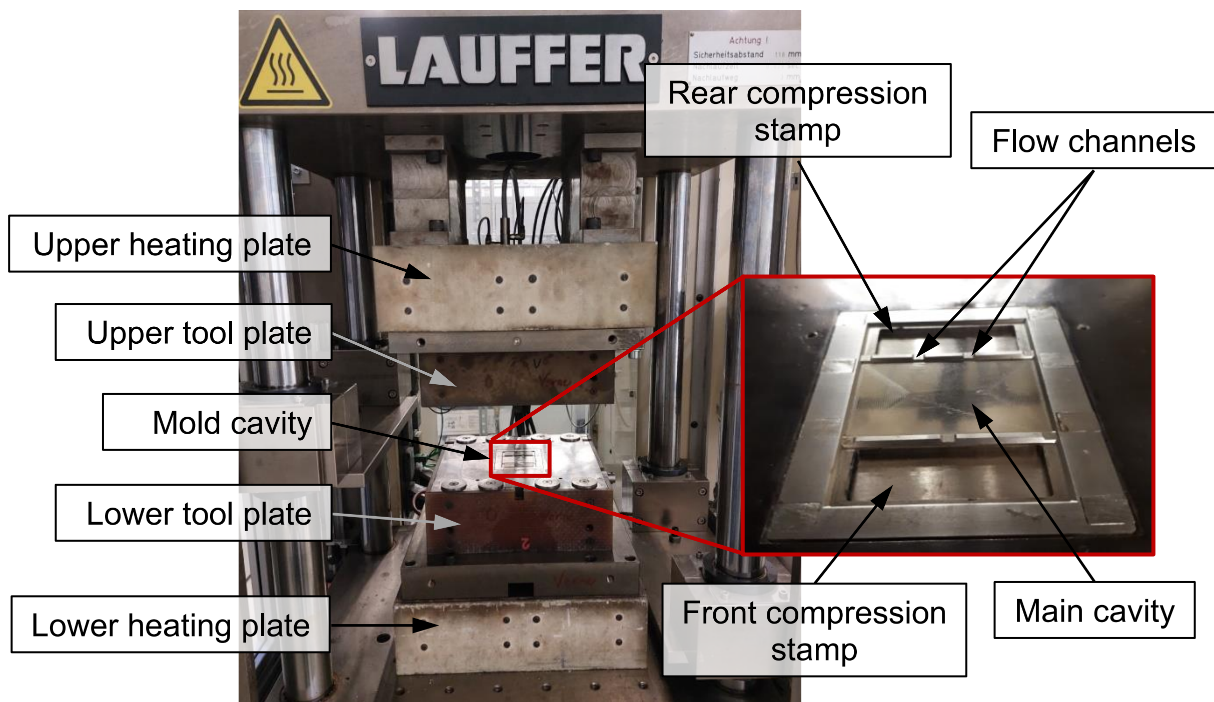

Figure 3.6: Press machine and the mold cavity used to produce samples

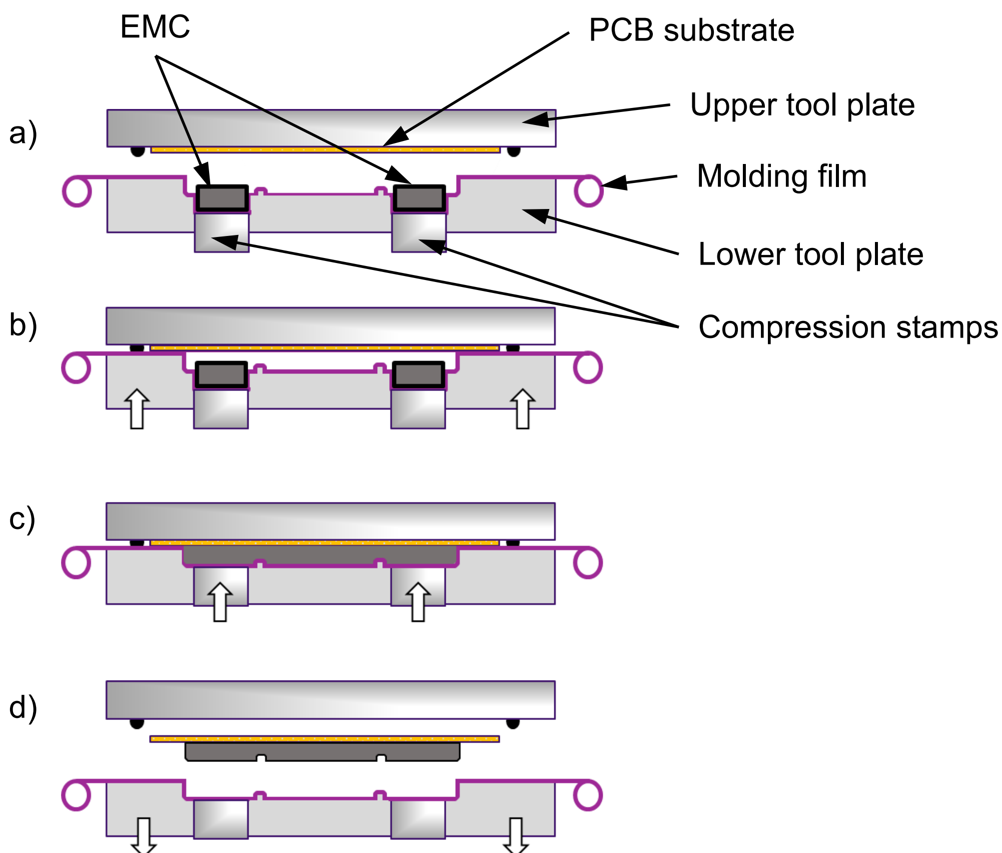

Figure 3.7: Schematic process steps in local pressure molding a) placing the EMC sheets and PCB substrate b) heating plates close c) compression stamps transfer and compress molten EMC d) solidified part is ejected

The molding procedure is illustrated in **Figure 3.7**. After the plates are heated up to mold temperature, a molding film is laid over the tool to prevent molten epoxy from sticking in the stamp clearances and on the surfaces. The PCB substrate is attached to upper tool plate using a heat resistant tape. The pressed EMC sheets are then placed on compression stamps as shown in **Figure 3.7** (a). After the plates are closed (**Figure 3.7** (b)), the compression stamps elevate and transfer the molten EMC into the main cavity through multiple flow channels while the material is still in pot time with low viscosity that enables proper material flow (**Figure 3.7** (c)). During packing, the EMC viscosity increases with further curing of EMC as mentioned in Chapter 2.1.3. Once the required hardness is reached the plates are opened, and the molded part is ejected and removed from the molding film (**Figure 3.7** (d)).

Even though the tool is designed to be used with two EMC sheets, one sheet is used to produce the experiment samples in this work in order to have a material flow from one side and avoid air gaps in the middle area where the flows from both sides meet. EMC granules are pressed into sheets with 6±0.05 g masses and kept in a freezer. Before molding, the sheets are brought to room temperature in a desiccator.

To have a larger process window to inspect the effects of PMC, it is aimed to have the largest possible amount of unreacted molecules remaining in EMC after molding process. However, a certain level of cross-linking density is required for proper ejection of the molded part as mentioned in Chapter 2.1.3. Therefore, the minimum mold cycle that allows a proper part ejection is investigated. Keeping other molding parameters constant as T_{mold} = 170 °C, material transfer velocity v = 0.2 mm/s, and packing pressure P = 50 bars, different mold cycle times are tested and t_{mold} = 65 s is determined as the minimum possible cycle time that yields a proper cavity filling and ejection. A total of 108 samples are molded with the parameters T_{mold} = 170 °C and t_{mold} = 65 s to inspect the effects of different PMC parameters on final warpages of the parts.

As an important point, it should be noted that the molding machine had an unexpected failure during production of the parts and had a repair operation after the 54$^{\text{th}}$ sample. Therefore, even though they are all produced with the same molding parameters T_{mold} = 170 °C and t_{mold} = 65 s, the samples are named as **Sample Group A** for the

first 54 samples before the machine failure, and **Sample Group B** for the last 54 samples which are produced after the repair.

Even though the main investigations for PMC effects on warpage are conducted using the sample groups A and B, some parts with different molding parameters are also produced to have more data points for the machine learning algorithm. **Sample Group C** consists of parts which are molded with T_{mold} = 170 °C and t_{mold} = 90 s to have a smaller amount of unreacted molecules left for PMC. It is aimed to include some parts with different cross-linking densities and different initial warpages to have a larger interval for the train data. The details about the PMC operations of these samples are explained in Chapter 3.3.3.

3.3.2 Measurements

To assess the effects of PMC on part warpage and EMC cross-linking density, 3D Optical Profilometer and Fourier Transform Infrared Spectroscopy measurements are conducted respectively, directly before and after the PMC operations. The measured values before and after PMC are then compared to correlate with the PMC parameters used.

Fourier Transform Infrared Spectroscopy for Determining Cross-linking Density

The basis of infrared spectroscopy is the interaction of infrared light with matter [119]. When a sample is subjected to a beam of infrared light, the chemical bonds vibrate at the wavelengths which match their vibration modes, and they absorb the radiation at these matching wavelengths while vibrating [120]. Thus, the chemical structure of the samples can be determined based on the infrared absorbance [119]. The peaks in the infrared spectrum indicates the presence of the corresponding chemical bonds or functional groups [121].

Fourier transformation infrared (FTIR) spectroscopy is a non-destructive method for identification of chemical compounds via IR spectra [122]. The principle of FTIR spectroscopy is based on the optical principle of Michelson interferometer, which uses a beam splitter and divides the IR beam coming from the source into two parts, one directed to a stationary mirror and the other one to a moving mirror [121]. After reflecting back from the mirrors, the recombined beams form an interference pattern,

and due to absorbed infrared energy, fluctuations in the intensity occur in time domain [123]. In FTIR spectroscopy, the time domain is transformed to the frequency domain using the Fourier transformation as shown in **Figure 3.8**. With this transformation, the evaluation of the functional groups that correspond to specific frequencies are made possible [121,124].

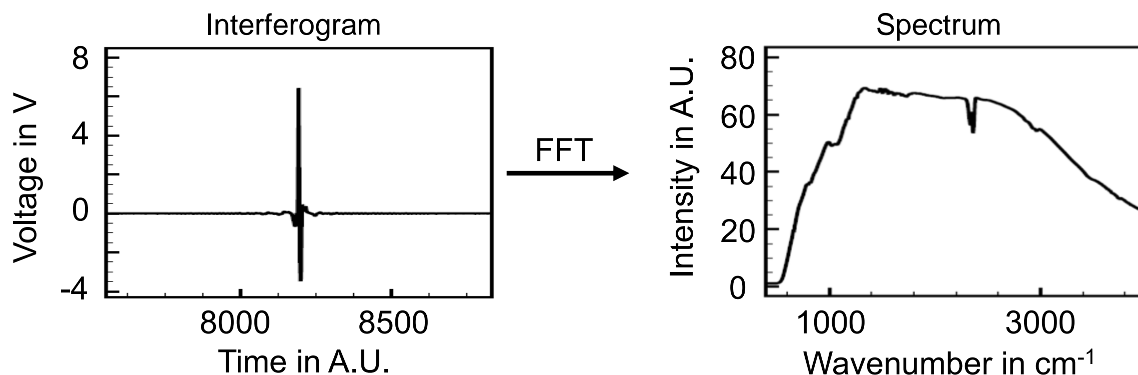

Figure 3.8: Fourier transformation of an interferogram to a spectrum [124]

There are three main sampling modes for FTIR Spectroscopy, namely transmission, reflectance and attenuated total reflectance [122]. Attenuated Total Reflectance (ATR) is a sampling method that allows fast and accurate IR spectroscopy analyses of liquid and solid specimens [125]. The IR beam is focused on an ATR crystal with a high refractive index, as a result the beam reflects internally in the crystal and these internal radiations penetrate the crystal surface a few microns towards the sample as shown in **Figure 3.9** [122]. Since the penetration depth is so small, a good contact in between the crystal and the samples must be ensured for proper absorption in sample [122,125].

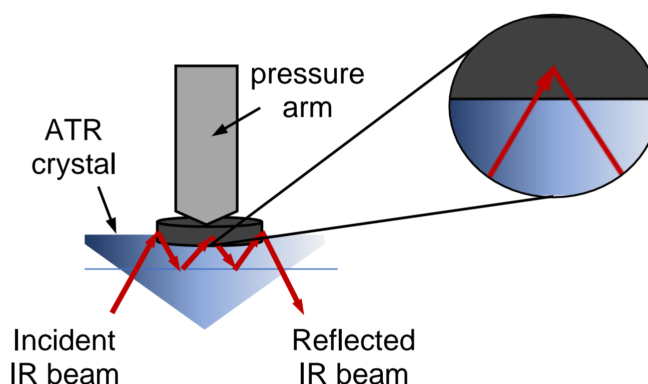

Figure 3.9: Attenuated Total Reflectance – adapted from [122]

In this work, the ATR-FTIR Spectroscopy method is used to monitor the changes in chemical composition of EMC material. The changes in IR spectra of EMC materials during molding are observed by Niegisch et al [126]. Ng and Hooi [61] detected the presence of PMC treatment in EMC encapsulated packages using FTIR spectroscopy. In the light of these studies, the main goal is to detect the changes that occur in IR spectrum during PMC, and to correlate these spectral changes with the changes in cross-linking density of the material.

The measurements are conducted using a Bruker ALPHA II type FTIR Spectrometer with a Platinum ATR tip. A measurement program with a resolution of 4 cm^{-1} over the wavenumber range 400 cm^{-1} – 4000 cm^{-1} is used. 12 scans were taken for each measurement and then these scans are averaged to achieve the resultant spectrum. The background spectrum is measured with the same settings before each measurement session. The first measurements for the samples are conducted directly after their molding processes, and then the parts are post-mold cured with different parameters. After PMC, the parts are kept in a desiccator until the second measurements. This waiting time is always kept under 12 hours.

Warpage Measurements and Evaluations

The warpage measurements are conducted using a 3D Optical Profilometer. Optical profilometry is a method that is used to study the height differences and surface topographies of objects. Instead of physical probes, the optical profilometers work on the principle of light reflection. It is a non-destructive and non-contact method that allows fast analyses. In this work, a device with Confocal Grid Structured Illumination (CGSI) is used. CGSI profilometry works with the principle of illuminating a grid pattern on the sample surface and detecting the distortions in the grid pattern with reference to a focus plane with undistorted grid [127,128]. **Figure 3.10** illustrates a schematic for a CGSI profilometry setup (a), the object image with illuminated grid pattern (b), and the resultant height map (c) which is computed via the deformed grid pattern.

The measurements for warpage evaluations are conducted with a VR5000 Optical 3D Profilometer (Keyence GmbH, Neu-Isenburg, Germany), which has a measurement tolerance of ±10 μm in vertical direction. Acquiring the raw surface topography of the warped samples, a Python script is used to compute the warpage values from the scanned height information. **Figure 3.11** illustrates the methodology used for warpage

evaluations. Considering the structure of the samples, only the warpage of middle area that is highlighted in **Figure 3.11** (a) is considered since the sensors and other electronics which are damaged by warpage are located in that area in real life applications. Starting with the scanned 3D height map (**Figure 3.11** (b)), a circle is fitted in the horizontal cross-section that passes through the symmetry center O as shown in **Figure 3.11** (c). The fitted radius and the length of the curvature is used to calculate the deflection δ which is shown in **Figure 3.11** (d). The cross-section is then rotated every 0.25° degrees around the center point O, and the cross-section with the biggest δ value is taken as the critical cross-section, and $\delta_{\max}$ at that cross-section is assigned as the critical warpage of the part.

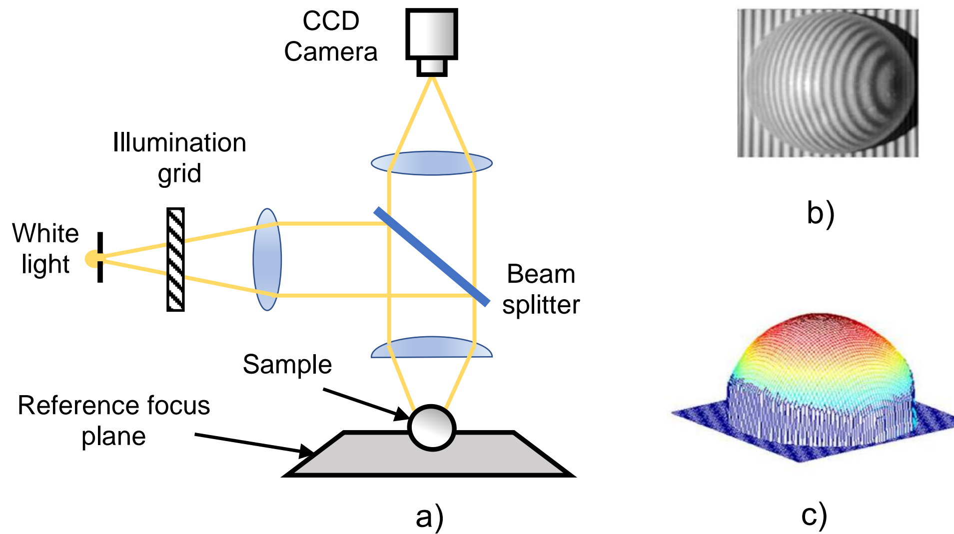

Figure 3.10: 3D Optical Profilometer a) schematic CGIS setup b) grid illuminated image c) 3D height map of the sample – adapted from [127,128]

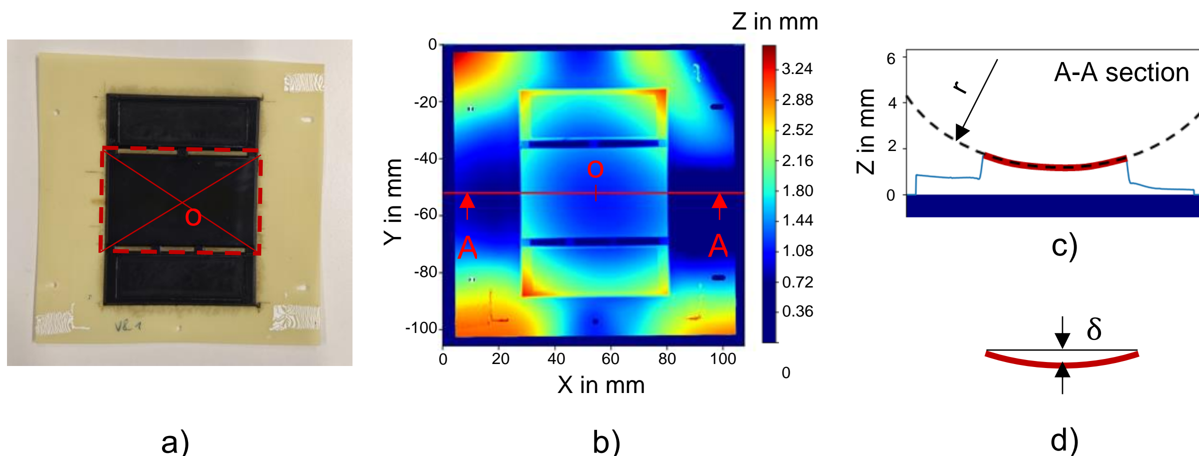

Figure 3.11: Principle of python script for warpage evaluation a) scanned image and area of interest b) 3D height data c) circle fitting in cross-section d) calculating the deflection via the curvature

3.3.3 Post-mold Cure Experiments

To investigate the effects of different PMC parameters on warpage, the molded samples are post-mold cured with varying temperatures and times. Before PMC, the samples are kept in ambient conditions and once warpage and IR spectra measurements are done, they are put in a fan oven for their PMC operations. **Table 3.1** summarizes the molding parameters used in different sample groups as described in Chapter 3.3.1, and the parameters used for their PMC operations are shown in **Table 3.2**. After PMC, the samples are brought back to room temperature in ambient conditions and the measurements for warpage and IR spectra are repeated directly.

Table 3.1: PMC experiments - mold parameters used in different sample groups

Sample Group	Mold temperature in °C	Cycle time in s	Quantity
A	170	65	54
B	170	65	54
C	170	90	27

Table 3.2: PMC experiments - PMC parameters used in the experiments for different sample groups

PMC Temperature	PMC time in min							
	5	10	20	30	60	120	180	240
185 °C	-	B, C	B	B, C	B, C	B	B	B
170 °C	A	A, C	A, B	A, B, C	A, B, C	A	A	A
160 °C	-	A, C	A	A, C	A, C	-	-	-
150 °C	-	A, B	A, B	A, B	A, B	B	B	B

3.4 Optimization Algorithm

As the second part of the thesis after experimentally investigating the effects of different PMC parameters on resultant warpage of an electronic package, a machine learning based optimization algorithm is created that predicts the optimal PMC temperature and PMC time which yield the desired part warpage and EMC cross-linking level after the PMC operation. Even though the main quality criterion focused on in this thesis is warpage, cross-linking density of the EMC encapsulant is the major consideration for real-life PMC operations. Therefore, the reaction enthalpy of EMC is also included in the optimization algorithm as a measurable indicator of cross-linking density with reference to the principle mentioned in Chapter 3.2.1.

Figure 3.12 illustrates the flowchart for PMC parameter optimization using the machine learning model. As seen, the part has the initial warpage δ_0 and initial reaction enthalpy h_0 before its PMC operation. The non-destructive 3D Profilometry and ATR-FTIR Spectroscopy measurements are conducted to correlate with these initial values and then the measurements are fed into the machine learning model also with the desired final warpage δ'_{desired} and desired final reaction enthalpy h'_{desired}. The model then iteratively predicts the optimal PMC parameters that yield the desired values.

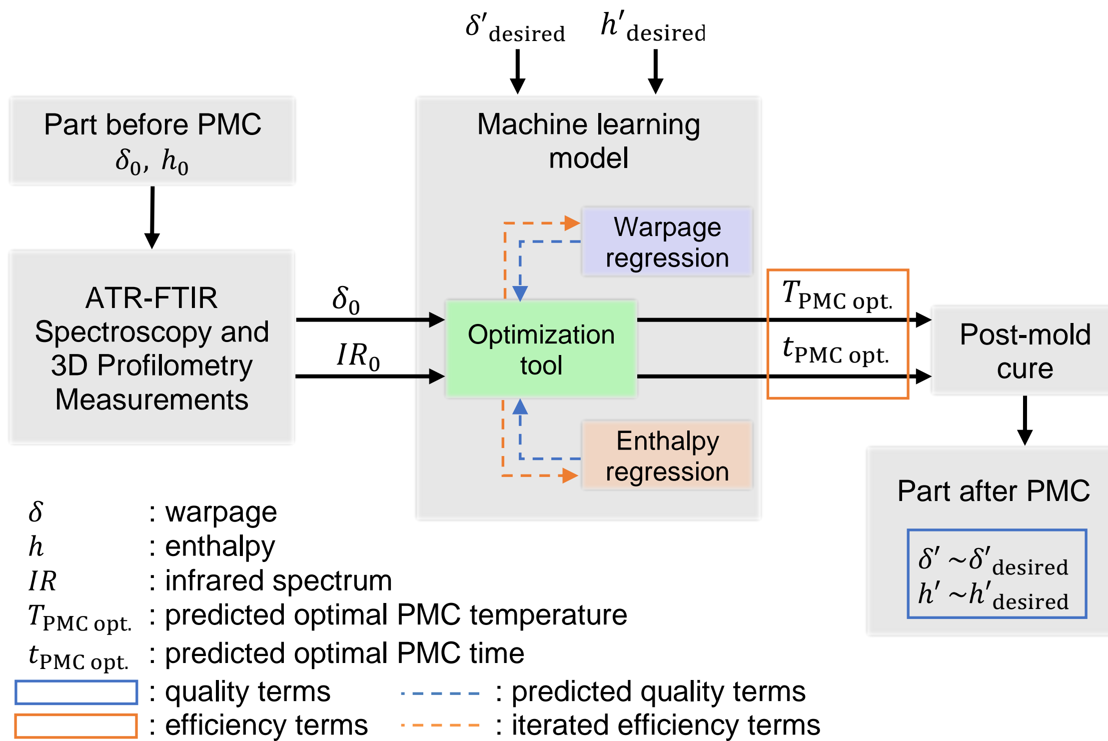

Figure 3.12: Flowchart for process optimization with the machine learning model

From the perspective of the machine learning model, PMC temperature T_{PMC} and PMC time t_{PMC} are the two efficiency terms. And similarly, the final warpage δ' and final reaction enthalpy h' are called the quality terms. Efficiency terms are independent variables that can be determined iteratively, and the quality terms are the dependent variables that are varied as a result of the selected efficiency terms. With reference to these properties, δ' and h' are the values that the machine learning algorithm has to predict, and T_{PMC} and t_{PMC} are the values that can be iteratively fed as the inputs for the predictions. Therefore, the machine learning model consists of two regression functions that predict for the quality terms δ' and h', and an optimization tool that feeds the regression functions with iterated PMC parameters as well as the measured initial quality terms δ_0 and IR_0.

This subchapter describes the principles behind the introduced machine learning model in detail. Starting with data preparation in Chapter 3.4.1, the algorithms used in warpage and enthalpy regression functions are described in Chapters 3.4.2 and 3.4.3 respectively. Following these, in Chapter 3.4.4, the principle of the optimization tool is introduced in which all other functions are used together with different tasks.

3.4.1 Data Preparation

To train the regression tools, two different data sets are used. The first data set is created via the PMC experiments conducted in context of the thesis. Having the measured warpage values δ_0 and δ' , the IR spectra IR_0 and IR', and knowing the used PMC parameters T_{PMC} and t_{PMC}, the data points for the molded samples are created. These data points include the warpage values and IR spectra before and after PMC, and different PMC parameters used for each one of these samples. This data set is used for training warpage and enthalpy regression functions. Some of the data points are listed as examples in **Table 3.3**. In accordance with the requirements, the features and the labels are selected accordingly for training.

Table 3.3: Data Set 1 – Data acquired from PMC experiments

Sample	Measurements before PMC					PMC		Measurements after PMC				
	δ_0 in µm	IR intensities in A.U.				T_{PMC} in °C	t_{PMC} in min	δ' in µm	IR intensities in A.U.			
		X_1	X_2	…	X_n				$\mathrm{X'}_1$	$\mathrm{X'}_2$	…	$\mathrm{X'}_\mathrm{n}$
PMC 1	388	0.96	0.96	…	0.12	170	5	275	0.96	0.96	…	0.21
…												
PMC 52	362	0.94	0.95	…	0.16	150	60	192	0.93	0.93	…	0.15
…												
PMC 106	644	0.98	0.98	…	0.35	185	240	506	0.96	0.96	…	0.22

Since DSC measurements require destructive sample preparation for molded parts, another previously prepared data set is used to train a subfunction for enthalpy regression. EMC granules are cold pressed into 30±2 mg sheets and then cured in a fan oven with varying temperatures and times. After curing, ATR-FTIR spectroscopy measurements are conducted to acquire IR spectra of the samples with different cross-linking densities. Following that, DSC measurements are performed to measure their reaction enthalpies. Having the measured IR spectra and reaction enthalpies for the samples with different cross-linking levels, a data set was created prior to this work. **Table 3.4** shows some of the data points included in this data set.

Table 3.4: Data Set 2 – External data acquired from DSC specimens

Sample	IR intensities in A.U.				Measured reaction enthalpy in J/g
	X_1	X_2	…	X_n	
DSC 1	0.97	0.96	…	0.89	28.85
…					
DSC 233	0.97	0.97	…	0.92	9.11
…					
DSC 386	0.96	0.96	…	0.79	1.76

Prior to being used for training the machine learning models, some preprocessing steps are applied to data sets. The IR spectra are normalized to overcome the measurement related deviations in between intensities, and the wavenumbers outside the range 600 cm^{-1} – 1800 cm^{-1} are cropped to decrease the spectral features that are irrelevant to EMC material. Similarly, the PMC parameters and measured warpage values are normalized to train the ANN algorithms.

3.4.2 Warpage Regression

Being the main quality criterion considered in this thesis, a machine learning based function is created to predict the final warpage of the parts after their PMC operations. As shown in **Figure 3.13**, the function takes the inputs of initial warpage, initial IR spectrum and the PMC parameters, and predicts for the resultant warpage after PMC.

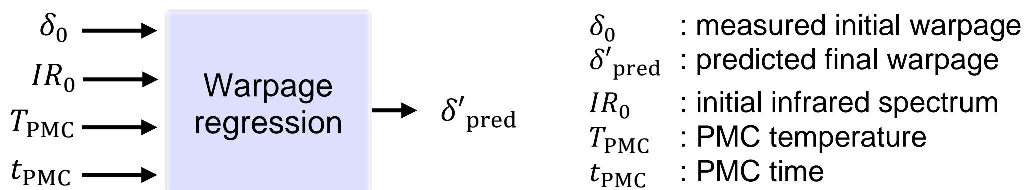

Figure 3.13: Flowchart for warpage regression function

To create the function, an ANN algorithm is used. The algorithm is trained with Data Set – 1, which was created via the PMC experiments. Different ANN architectures and learning parameters are tested to find a feasible architecture and a proper parameter set. Chapter 4.3.1 explains these parameters and the predicted values.

3.4.3 Enthalpy Regression

The enthalpy regression function consists of two functions for initial and final reaction enthalpy predictions as shown in **Figure 3.14**. Firstly, the function *Step I* predicts for

the initial reaction enthalpy h_0 with the input of initial infrared spectrum IR_0. Following that, *Step II* predicts for the final reaction enthalpy h' in accordance with the input parameters of PMC temperature T_{PMC} and PMC time t_{PMC}.

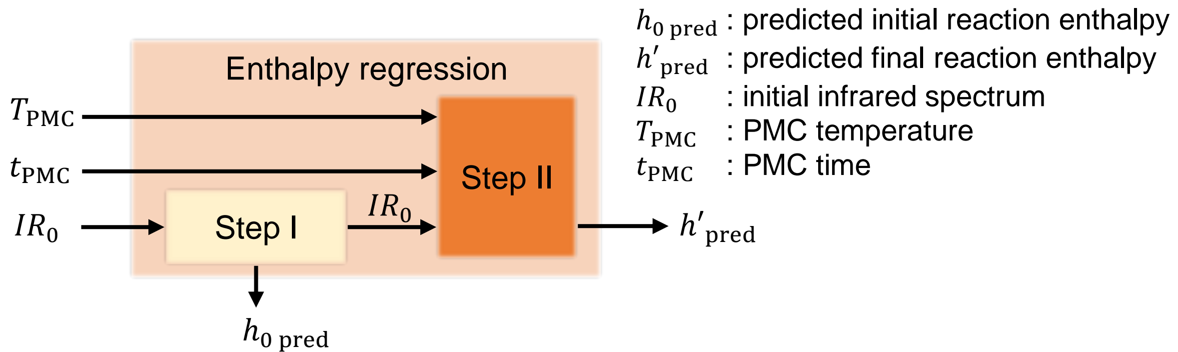

Figure 3.14: Flowchart for enthalpy regression function

Just like warpage regression function, an ANN algorithm is used to build the function. The architecture and learning parameters from warpage regression model are used for enthalpy regression as well since the input features are almost the same with only one numerical output. To train the enthalpy regression function, both Data Set – 1 and Data Set – 2 are used in combination for different steps.

3.4.4 Optimization Cycle

The parameter optimization is done iteratively via predicting the final warpage δ' and the final reaction enthalpy h' values for varying PMC parameters. The regression functions are used for predicting for the iterated steps as shown in **Figure 3.15**.

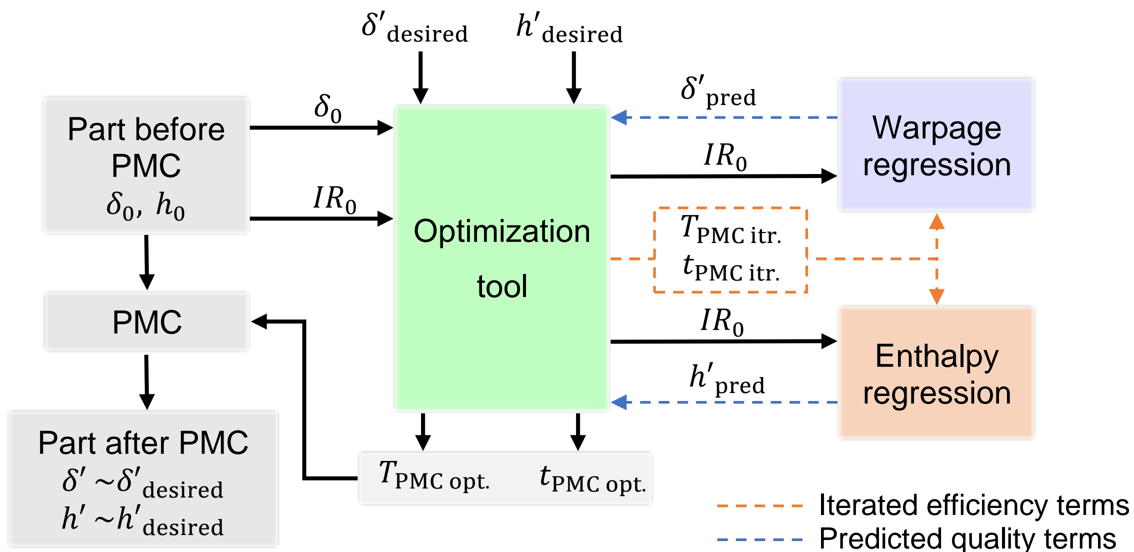

Figure 3.15: Flowchart for optimization tool

The initial warpage value δ_0 and initial IR spectrum IR_0 of the part are fed into the model as well as the desired values for final warpage δ'_{desired} and final reaction enthalpy h'_{desired}. The iterative loop is then started for PMC temperature $T_{\text{PMC itr.}}$ and time $t_{\text{PMC itr.}}$, and the final values for warpage and reaction enthalpy are predicted using the regression functions. The values for $T_{\text{PMC itr.}}$ and $t_{\text{PMC itr.}}$ are assigned randomly for the first iteration. The cost function

$$J = w_\delta\left(\delta'_{\text{desired}} - \delta'_{\text{pred}}\right)^2 + w_h\left(h'_{\text{desired}} - h'_{\text{pred}}\right)^2 + w_T(T_{\text{PMC itr.}})^2 + w_t(t_{\text{PMC itr.}})^2 \tag{3.2}$$

is defined where δ'_{pred} and h'_{pred} stand for predicted final warpage and predicted reaction enthalpy values, and w_δ, w_h, w_T and w_t are the corresponding weights to quality and efficiency terms to define their importance for the total cost. As seen, the cost function increases when the predicted quality terms diverge from the desired values. Similarly, the increasing time and temperature increase the cost since they increase the process costs for PMC operation. Using a modified gradient descent approach with reference to Giani's work [34], the PMC parameters are updated after each iteration in the direction of decreasing cost. When δ' and h' are lower than their desired values, then the weights w_δ and w_h are assigned as 0 for that iteration step since no costs will arise when the desired values are satisfied. Once the cost is stagnated at a point, or decreased to a defined level, the iteration cycle is stopped and the parameters that yield the minimum cost are presented as the optimal parameters $T_{\text{PMC opt.}}$ and $T_{\text{PMC opt.}}$. Appendix A1 gives a more detailed description of the optimization tool and the gradient descent method used for parameter iterations.

4 Results

Following the methods and experimental plan explained in Chapter 3, this chapter presents the acquired results. Starting with the findings from material characterization tests and their interpretation in Chapter 4.1, the results for post-mold cure experiments are presented and evaluated in Chapter 4.2. Finally, in Chapter 4.3, the prediction outcomes from the machine learning model are shown and discussed.

4.1 Results from Material Characterization Tests

This chapter presents the results from the material characterization tests, which are conducted to understand the thermal and cure behaviors of the materials. The results from DSC measurements and their interpretations are presented in Chapter 4.1.1. As following, the outcomes from TMA measurements are described in Chapter 4.1.2. The important points from both tests are then summarized and interpreted in Chapter 4.1.3.

4.1.1 Differential Scanning Calorimetry Measurements

Differential scanning calorimetry (DSC) measurements are conducted to measure the reaction enthalpies and the glass transition temperatures of EMC at different cross-linking densities. Also, measurements for PCB substrate are conducted to investigate the glass transition behavior and to detect any possible curing of PCB during PMC with the considerations and settings described in Chapter 3.2.1.

Results for EMC

Figure 4.1 shows the dynamic heating sections of the DSC curves acquired from the EMC specimens extracted from the samples that are molded at 170 °C with varying cycle times. The exothermic peaks can be seen in the temperature interval 120 °C – 210 °C. The curing behavior of EMC can be observed via the disappearing peaks as the mold cycle time increases. There are no clearly observable exothermic peaks left after 110 seconds of molding.

The reaction enthalpies of the samples are calculated via the areas under the exothermic peaks in DSC curves as described in Chapter 3.2.1. **Figure 4.2** shows the averaged measured reaction enthalpies. As seen, the reaction enthalpy of raw EMC is measured as 28.76 J/g, and then the values decreased rapidly until 65 seconds of

mold cycle time with a measured reaction enthalpy of 2.66 J/g. Then the slope of the curve flattened, and after 110 seconds the enthalpy values converged to 0 J/g.

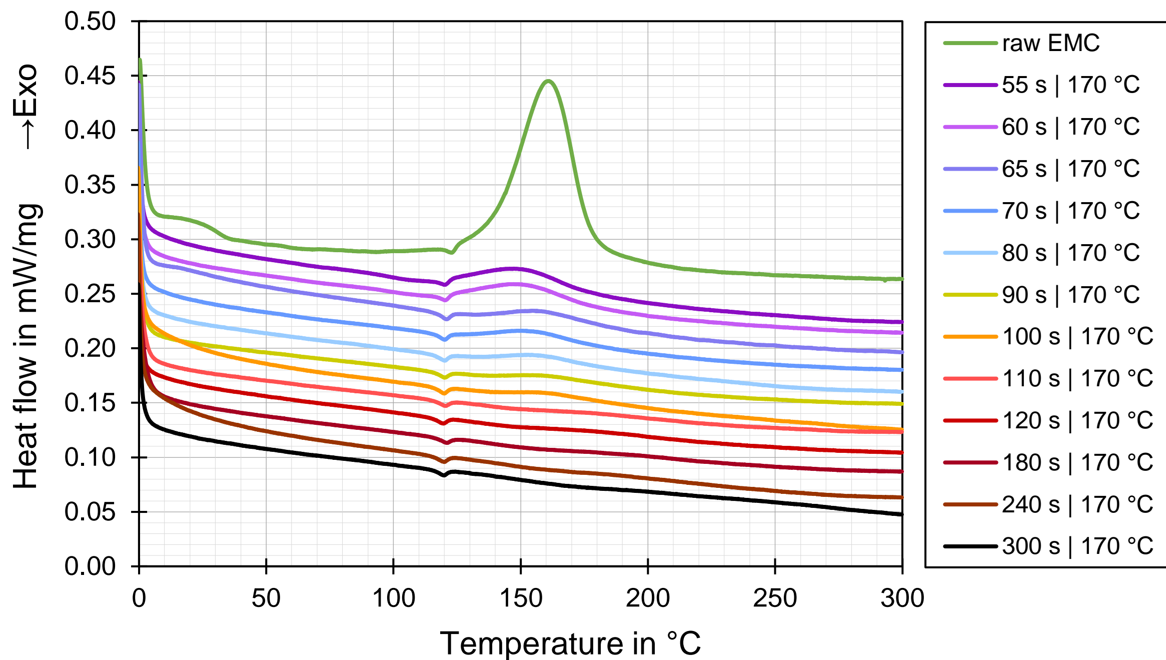

Figure 4.1: Dynamic heating section of DSC curves acquired from EMC samples molded at 170 °C with different cycle times

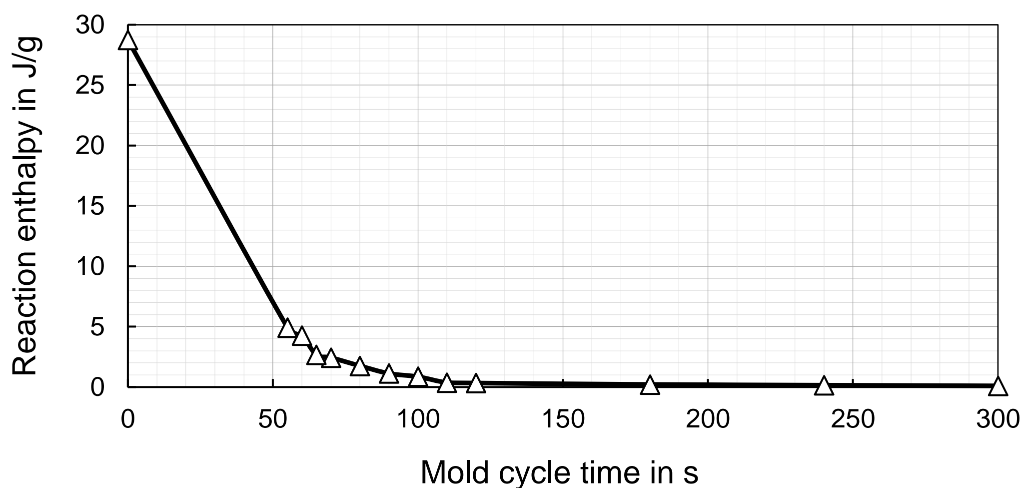

Figure 4.2: Measured reaction enthalpies of EMC samples molded at 170 °C with different cycle times

A glass transition range is clearly visible for raw EMC in **Figure 4.1**. As mentioned in Chapter 3.2.1, midpoint $T_{\mathrm{g(m)}}$ and return-to-baseline $T_{\mathrm{g(r)}}$ temperatures of the glass transition region are taken into consideration. Focusing on a smaller window in the graph for raw EMC, the midpoint temperature $T_{\mathrm{g(m)\,EMC\,raw}}$~28.5 °C and return-to-baseline temperature $T_{\mathrm{g(r)\,EMC\,raw}}$~42 °C can be determined as shown in

Figure 4.3 (a). With further curing of EMC during molding, these values increased to $T_{\mathrm{g(m)\ EMC\ 55s\ |\ 170°C}}$ ~97.5 °C and $T_{\mathrm{g(r)\ EMC\ 55s\ |\ 170°C}}$ ~110 °C after 55 seconds of molding at 170 °C as seen in **Figure 4.3** (b). Referencing Chapter 3.2.2, with increasing cross-linking density, the determination of glass transition interval via DSC gets difficult. No clear observations on glass transition were therefore possible via DSC measurements for the samples with cycle times longer than 55 s.

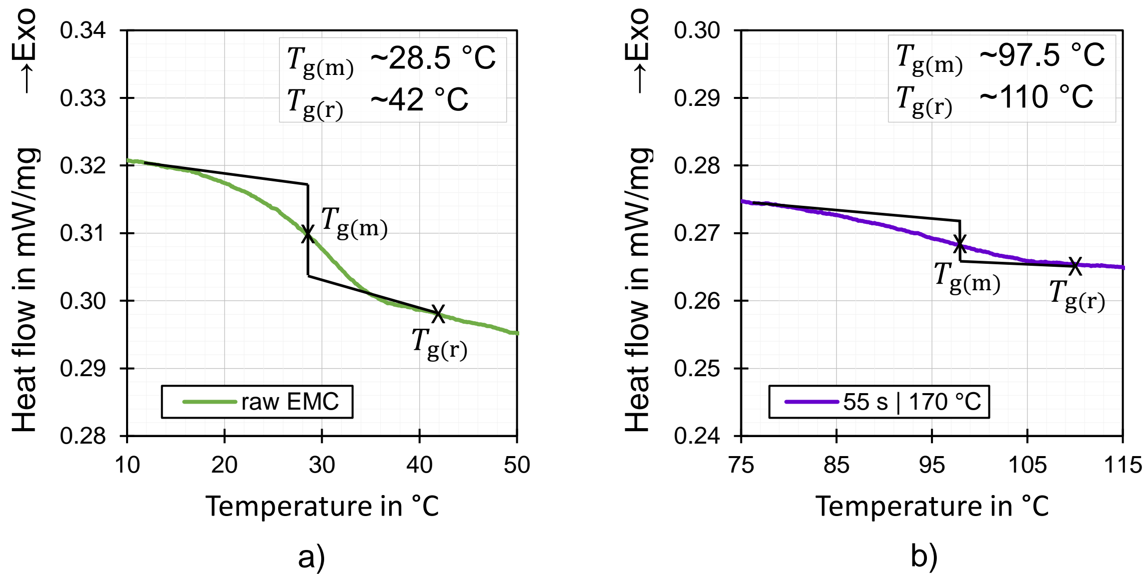

Figure 4.3: Glass transition temperatures acquired from DSC curves a) non-cured raw EMC b) EMC molded for 55 s at 170 °C

Results for PCB substrate

Even though the externally supplied PCB substrates were assumed to be at their full cross-linking densities, DSC measurements are conducted before and after their PMC operations to investigate the glass transition behavior and any possible curing of the glass epoxy. To do so, DSC specimens are extracted from the PCB substrates of the samples which are molded with T_{mold} = 170 °C and t_{mold} = 65 s as described in Chapter 3.3.1. Similarly, another group of specimens are extracted from the samples which are molded with the same parameters, but also post-mold cured with T_{PMC} = 185 °C and t_{PMC} = 120 min. Two measurements are conducted for each group.

Figure 4.4 shows the dynamic heating sections of the DSC curves acquired from the above-mentioned PCB specimens with and without PMC operations. When the PCB substrates are post-mold cured at 185 °C for 2 hours, the averaged midpoint temperature increased from the value of $T_{\mathrm{g(m)\ PCB}}$ ~ 145 °C to $T'_{\mathrm{g(m)\ PCB}}$ ~ 157 °C,

whereas the average return-to-baseline temperature has shifted from $T_{\text{g(r) PCB}}$ ~ 165 °C to $T'_{\text{g(r) PCB}}$ ~ 175 °C. The endothermic peaks at around 250 °C are present in all of the curves, which indicate material melting followed by the thermal degradation after ~ 270 °C. The endothermic peaks in the interval 80 °C – 110 °C on the other hand, are only seen in the curves of the samples without PMC, which indicate the presence of trapped humidity.

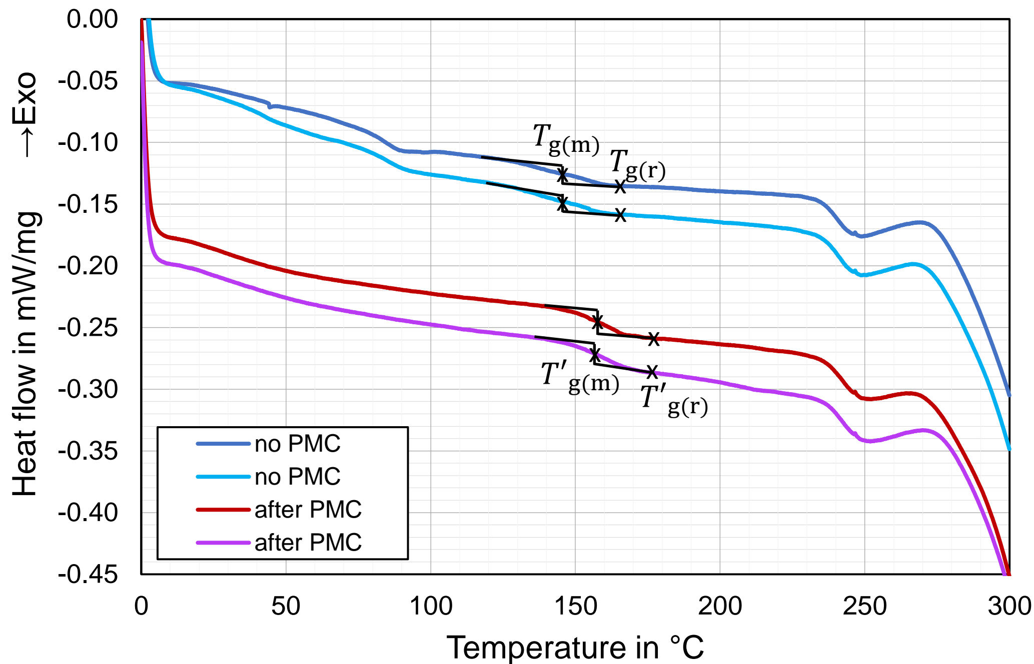

Figure 4.4: DSC curves of PCB substrate before and after 2 hours PMC at 185 °C

Since PCB substrate is not the main encapsulant material but only a base to the samples produced in context of this thesis, no detailed investigations are conducted regarding its cure behavior. No quality criterion was set considering the cross-linking density of the PCB substrates. Therefore, the DSC measurements were only conducted to roughly inspect the glass transition behavior and any possible curing of the substrates, which is related with warpage in the further steps.

4.1.2 Thermomechanical Analysis

Thermomechanical analysis (TMA) measurements are conducted using post-mold cured and fully cross-linked EMC specimens as described in Chapter 3.2.2. By this means, investigating the glass transition temperature of fully cross-linked EMC is aimed, which could not be observed via DSC measurements. **Figure 4.5** shows the measured dimension changes plotted over temperature. As seen, with the transition

from glassy state to gel state, the dimension change curve gets steeper due to increasing CTE value. Glass transition temperature can be determined via observing the change in the slope of the curve. From the conducted measurements, this point is determined as $T'_{\mathrm{g\,EMC}}$ ~ 177 °C for the EMC at its fully cured state after PMC. Also, as an additional information, the CTE values are measured as 8.73 μm/(m.°C) and 22.50 μm/(m.°C) at the temperatures 120 °C and 240 °C respectively.

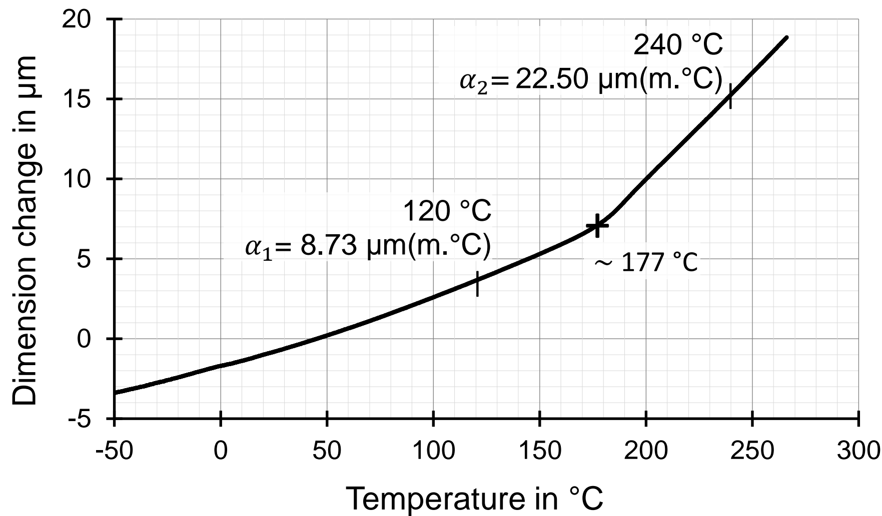

Figure 4.5: TMA measurements for highly cross-linked EMC

4.1.3 Overview and Interpretation of Material Characterization Tests

Starting with overviewing the results for EMC, DSC measurements showed that the EMC material reached a highly cross-linked structure after the molding operation with mold temperature T_{mold} = 170 °C and the cycle time t_{mold} = 110 s, after which the measured reaction enthalpy values converged to 0 J/g.

To determine the T_{g} values of different cross-linking levels of EMC, both DSC and TMA measurements are used which rely on different principles. Thus, a comparison of results from these methods could only be used for rough estimation purposes. Since return-to-baseline temperatures from DSC measurements yield a temperature value in which a complete transformation to gel state is achieved, the $T_{\mathrm{g(r)}}$ values from DSC are in better accordance with the T_{g} values acquired from TMA. Therefore, $T_{\mathrm{g(r)}}$ values are used for comparison purpose.

From the DSC curves, the return-to-baseline of the raw EMC is determined as $T_{\mathrm{g(r)\,EMC\,raw}}$ ~ 42 °C. With increasing cycle time, no observable glass transitions were

present in the DSC curves after $T_{\mathrm{g\,r\,EMC\,55s\,|\,170°C}}$ ~ 110 °C. For the almost fully cured EMC after PMC, the value $T'_{\mathrm{g\,EMC}}$ ~ 177 °C is acquired from TMA measurements. Since the experimental parts are molded with the cycle time t_{mold} = 65 s, the glass transition temperature for that point is roughly estimated as $T_{\mathrm{g(r)\,EMC\,65s\,|\,170°C}}$ ~ 120 °C via linear interpolation through the measured values with the rough asssumption of $T'_{\mathrm{g\,EMC}}$ is reached after 110 s. As an additional information, the CTE values of α_1 = 8.73 µm/(m.°C) and α_2 = 22.50 µm/(m.°C) are also acquired for glassy and gel states of fully cured EMC respectively.

Looking at the results from PCB measurements, no exothermic peaks were visible in the DSC curves, but an increase in average glass transition temperature was seen from the value $T_{\mathrm{g(r)\,PCB}}$~165 °C before PMC to $T'_{\mathrm{g(r)\,PCB}}$~175 °C after PMC. This indicated the presence of curing reactions in PCB substrate, which was considered to be at its fully cured state when delivered from the manufacturer. Even though only the cross-linking density of EMC is taken into consideration as a quality criterion for package reliability, the curing characteristics of PCB and its effects on warpage might be regarded as a future field of study.

4.2 Results from Post-mold Cure Experiments

This chapter presents and interprets the results acquired from the post-mold cure experiments which are described in Chapter 3.3.3. Starting with presenting the results about PMC effects on warpage in Chapter 4.2.1, these results are interpreted in Chapter 4.2.2. Following these, the results acquired from FTIR spectroscopy measurements are explained in Chapter 4.2.3.

4.2.1 Results on Warpage

As mentioned in Chapter 3.3.2, the warpage measurements for the parts are conducted before and after PMC operations. For sake of simplicity, only the results from the sample groups A and B, which are produced with the same parameters, are discussed to explain the effects of different PMC parameters on their final warpages.

Figure 4.6 shows the measured warpage values of the samples before and after PMC operations with different parameters. As seen, even though the parts were produced with the same molding parameters, the fluctuation in their initial warpage values (δ_0) is

relatively high due to instabilities in the machine. The manual operation of molding process considering raw material placements and part ejections is also another contributing factor to these deviations in δ_0. The effects of the failure in the machine, which is shortly mentioned in Chapter 3.3.1, can also be seen when the δ_0 values from both sample groups are compared. There is a large increase in δ_0 values of the samples from group B, even though the main molding parameters are kept the same. Due to this large difference, a direct comparison of the samples from the groups A and B were not possible. The small increase in average δ_0 for the samples 76-87 is caused since a new batch of the same PCB substrates is used for that interval.

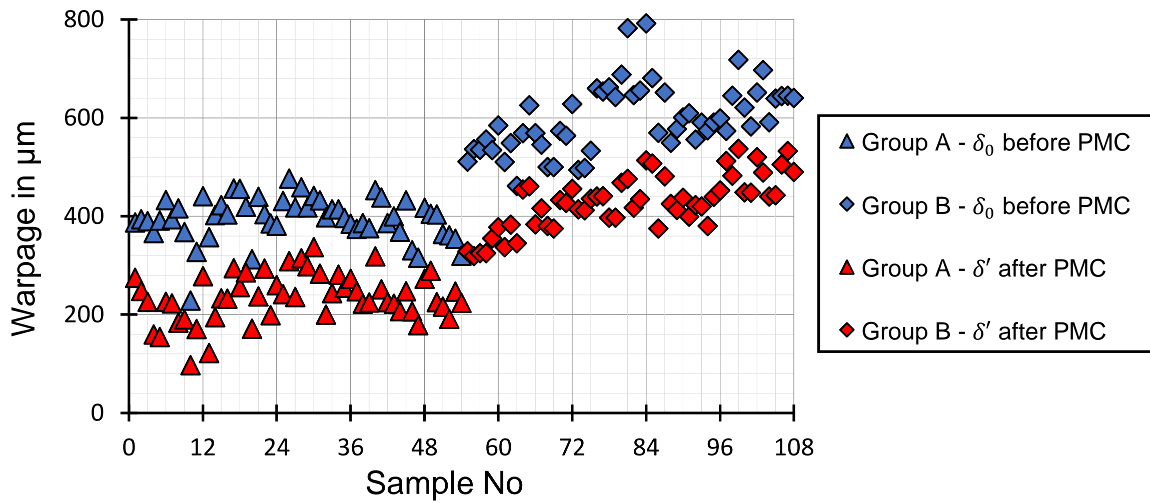

Figure 4.6: Measured warpage values of sample groups A and B

Looking at the values after PMC, it can be seen that PMC decreased warpage in all the samples with different PMC parameters. But, as the warpage value after PMC is also dependent on δ_0 value of the part, the fluctuated pattern is present in the final warpage δ' values as well, which makes it hard to compare these numerical values and to correlate a pattern in between them and the PMC parameters used. Due to these challenges in evaluating the numerical δ' values, the warpage reduction rate in percentage is used to compare the effects of different PMC parameters on warpage. The mentioned warpage reduction rates of the samples are calculated via

$$ {}^{T_{\mathrm{PMC}}}_{t_{\mathrm{PMC}}}\Delta\delta_X = \frac{\left(\delta_0 - {}^{T_{\mathrm{PMC}}}_{t_{\mathrm{PMC}}}\delta'_{\mathrm{X}}\right)}{\delta_0} \times 100 \qquad (4.1) $$

where δ_0 and ${}^{T_{\mathrm{PMC}}}_{t_{\mathrm{PMC}}}\delta'_{\mathrm{X}}$ are the warpage values in microns before and after PMC respectively. The upper and lower indices T_{PMC} and t_{PMC} on the left side indicate the

PMC temperature and PMC time used for the sample, and the lower index X on right specifies the group that the sample belongs to. As there are at least 3 parts in a sample group that are post-mold cured with the same parameters, the mean value of their reduction rates is taken into consideration during evaluations.

Figure 4.7 shows the average $\Delta\delta$ values acquired after post-mold curing the samples from Sample Group A with the parameter set given in **Table 3.2**. The error bars indicate the standard deviations in $\Delta\delta$ values. As seen, with varying T_{PMC} and t_{PMC}, different $\Delta\delta$ values are observed. Looking at the values for ${}^{170\,°\mathrm{C}}\Delta\delta_{\mathrm{A}}$, a peak can be seen at ${}^{170\,°\mathrm{C}}_{10\,\mathrm{min}}\Delta\delta_{\mathrm{A}}$ ~ 58 % which then shows a decreasing trend as t_{PMC} increases further. On the other hand, considering the measurement and production related tolerances, the fluctuations in ${}^{160\,°\mathrm{C}}\Delta\delta_{\mathrm{A}}$ and ${}^{150\,°\mathrm{C}}\Delta\delta_{\mathrm{A}}$ in the time interval of 10 – 60 minutes can be assumed as more or less a constant trend which does not show a clear increase or decrease unlike ${}^{170\,°\mathrm{C}}\Delta\delta_{\mathrm{A}}$. After 60 minutes, the decreasing ${}^{170\,°\mathrm{C}}\Delta\delta_{\mathrm{A}}$ reaches to the same level with other samples and follows a linear pattern until 180 minutes. Afterwards, another drop is seen at ${}^{170\,°\mathrm{C}}_{240\,\mathrm{min}}\Delta\delta_{\mathrm{A}}$ to ~ 28 %.

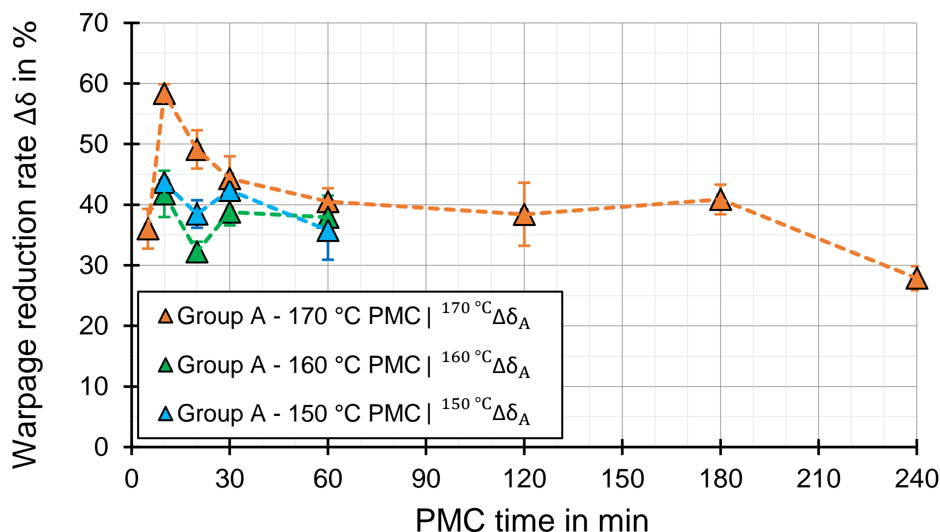

Figure 4.7: Average warpage reduction rates of samples from Group A

Figure 4.8 shows the average $\Delta\delta$ values of the samples from Sample Group B, which are post-mold cured with the parameter set shown in **Table 3.2**, with the aim of investigating a larger interval for T_{PMC} = 150 °C and the effects of a higher a T_{PMC} value of 185 °C. Due to different δ_0 values of Sample Groups A and B, a small control group with T_{PMC} = 170 °C is included in Group B to compare with ${}^{170\,°\mathrm{C}}\Delta\delta_{\mathrm{A}}$. Starting with comparing the control groups ${}^{170\,°\mathrm{C}}\Delta\delta_{\mathrm{A}}$ and ${}^{170\,°\mathrm{C}}\Delta\delta_{\mathrm{B}}$ in the time interval of

20 – 60 minutes in both figures, a difference in $\Delta\delta$ values can be seen. Despite these differences, a similar pattern is present in both ${}^{170\,°\mathrm{C}}\Delta\delta_{\mathrm{A}}$ and ${}^{170\,°\mathrm{C}}\Delta\delta_{\mathrm{B}}$ which decreases with increasing t_{PMC}. Moreover, ${}^{185\,°\mathrm{C}}\Delta\delta_{\mathrm{B}}$ shows the same pattern as ${}^{170\,°\mathrm{C}}\Delta\delta_{\mathrm{A}}$ with a peak at t_{PMC}= 10 min, which is followed by a decreasing slope and then a stay in almost in the same values until t_{PMC} = 180 min. When compared with ${}^{185\,°\mathrm{C}}\Delta\delta_{\mathrm{B}}$, ${}^{170\,°\mathrm{C}}\Delta\delta_{\mathrm{B}}$ yields better reduction rates in the time interval they are compared. Just like ${}^{150\,°\mathrm{C}}\Delta\delta_{\mathrm{A}}$, ${}^{150\,°\mathrm{C}}\Delta\delta_{\mathrm{B}}$ has an almost constant $\Delta\delta$ value over the time span until 180 minutes. The decrease at t_{PMC} = 240 min is then present for both ${}^{150\,°\mathrm{C}}\Delta\delta_{\mathrm{B}}$ and ${}^{185\,°\mathrm{C}}\Delta\delta_{\mathrm{B}}$, but the decrease in ${}^{150\,°\mathrm{C}}\Delta\delta_{\mathrm{B}}$ is smaller when compared to ${}^{185\,°\mathrm{C}}\Delta\delta_{\mathrm{B}}$.

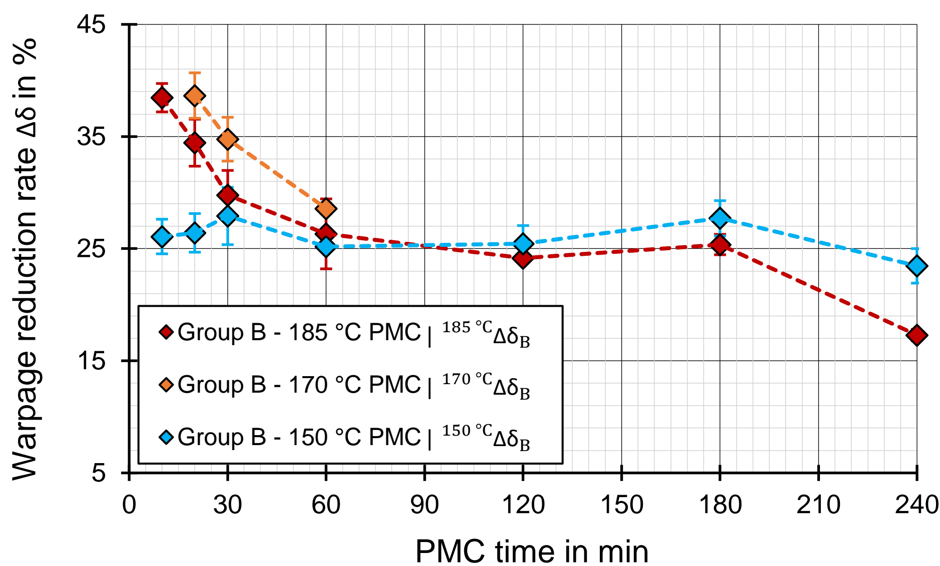

Figure 4.8: Average warpage reduction rates of samples from Group B

4.2.2 Interpretation of Results on Warpage

In the light of the results presented in previous chapter, this chapter explains the effecting mechanisms behind warpage reduction during PMC and reasons the different warpage reduction rates after different PMC parameters.

To begin with the warpage formation after molding process, let l be the horizontal length that passes through the symmetry center of the unwarped part and spans over the molded area as shown in **Figure 4.9** (a). **Figure 4.9** (b) illustrates the thermal shrinkage of EMC and PCB in this length during cooling down from the mold temperature T_{mold} to room temperature T_{room} after molding. Step 0 is the moment that the part is just taken out of the mold, and Step 1 represents the moment in which it reaches room temperature. When the lengths at Steps 0 and 1 are compared for both

components, it can be seen that the shrinkage Δl_{EMC} is larger than Δl_{PCB}, which causes a concave warpage formation when cooling from Step 0 to Step 1 as shown in **Figure 4.9** (c).

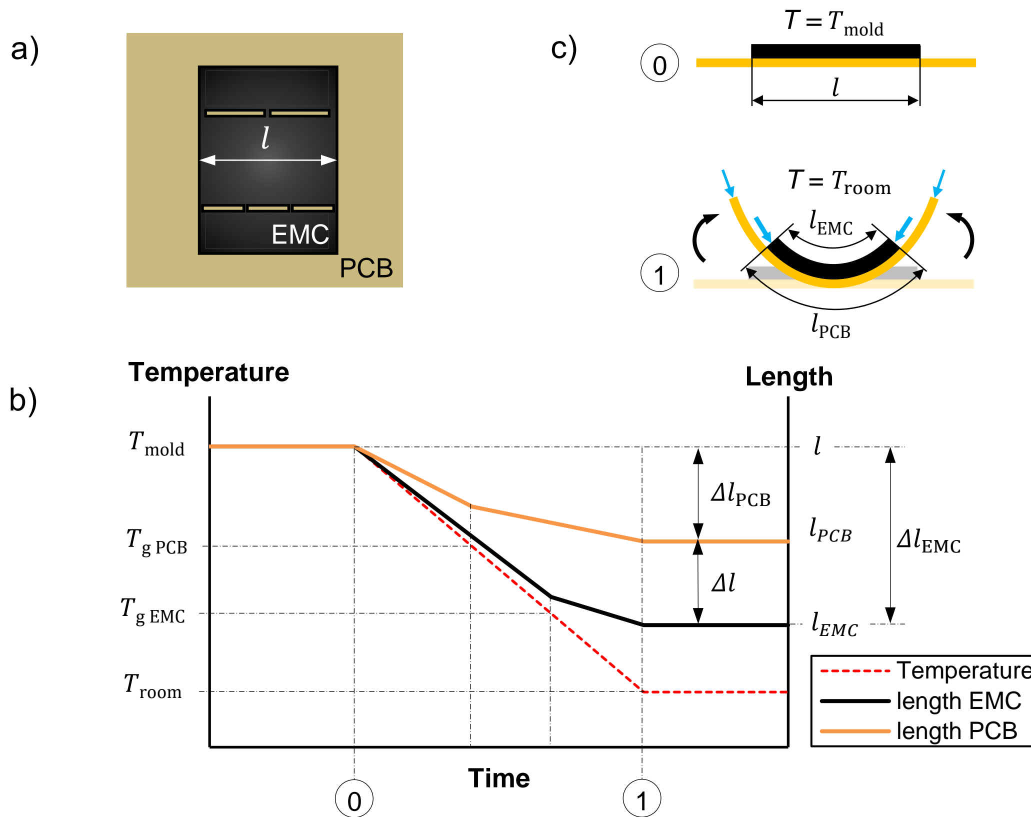

Figure 4.9: Warpage formation after molding a) part geometry and the reference length l b) shrinkage graph for EMC and PCB c) warpage modes just after ejection (0) and when cooled down to room temperature (1)

Figure 4.10 illustrates the warpage states during a PMC operation, which is conducted at $T_{\text{PMC}} \sim T_{\text{mold}}$ for a PMC time that suffices for an almost fully cured state of EMC but not for the full cure of the PCB substrate. Starting from Step 1 which was introduced above, the part temperature is elevated again. As a result of this increase in temperature, the concave warpage starts to decrease as the EMC expands more than the PCB substrate and leads to a decrease in Δl as illustrated in Step 2. Once $T_{\text{g EMC}}$ is reached during heating, the previously formed residual stresses in EMC are relaxed as a result of the transition from the glassy state to gel state as shown in Step 3. The similar stress relaxation effect happens again when $T_{\text{g PCB}}$ is reached as in Step 4. When the PMC temperature T_{PMC} is reached at Step 5, the temperature is kept constant for a while for further curing. During curing, the T_{g} values of both EMC and

PCB increase, and both components shrink chemically as a result of further cross-link formations. Due to fast curing behavior, EMC reaches a highly cross-linked structure at Step 6 and its glass transition temperature converges to T_{PMC}, leading to a stagnation in its curing reaction due to vitrification. The curing and chemical shrinkage of PCB on the other hand, continues until Step 7, which stands for the end of the PMC time after which the part is taken out of the oven. As EMC has a larger chemical shrinkage in the end of PMC, the concave warpage is increased in Step 7 in comparison to Step 5. As a result of the increased T_g values of $T'_{\mathrm{g\,EMC}}$ and $T'_{\mathrm{g\,PCB}}$, both materials shrink less while cooling back to room temperature for the second time. Due to this mechanism, the final warpage at Step 8 is relatively smaller when compared to Step 1 since $\Delta l'$ is smaller than Δl.

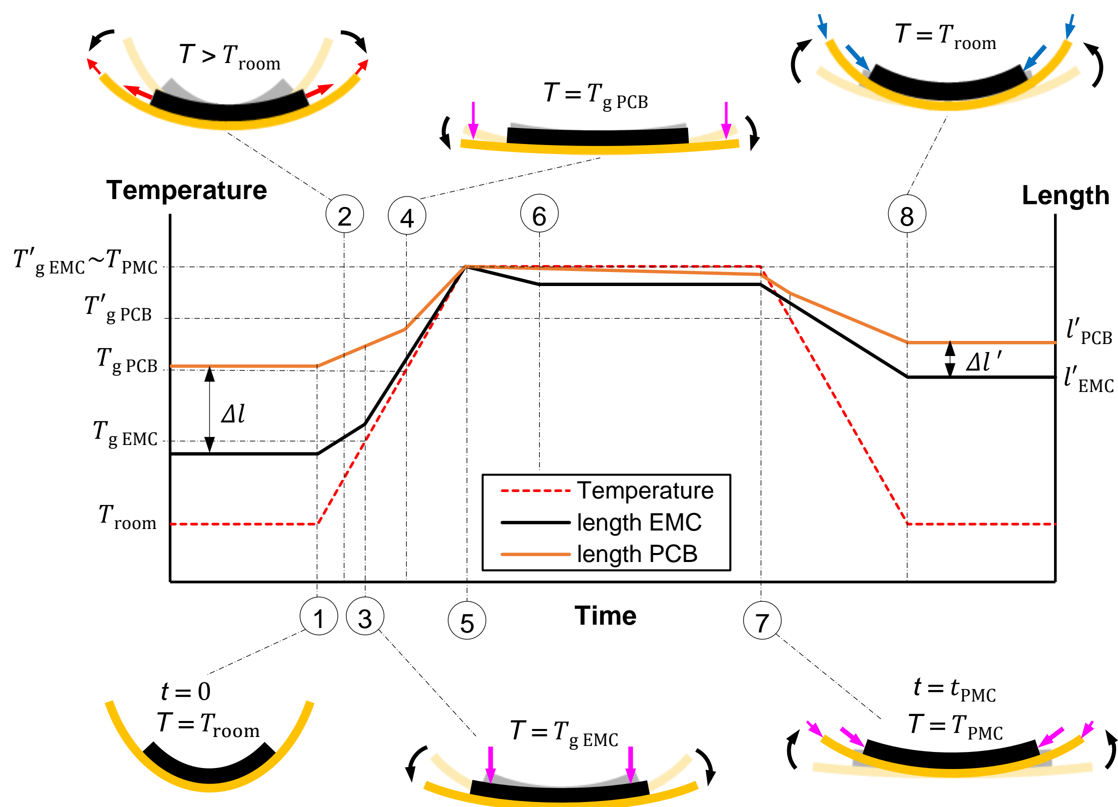

Figure 4.10: Warpage states during post-mold curing

With reference to above mentioned mechanisms acting during PMC, the results can be interpreted in a more meaningful way. **Figure 4.11** summarizes warpage reduction rates acquired from sample groups A and B, and divides them into three intervals according to patterns seen in $\Delta\delta$ curves.

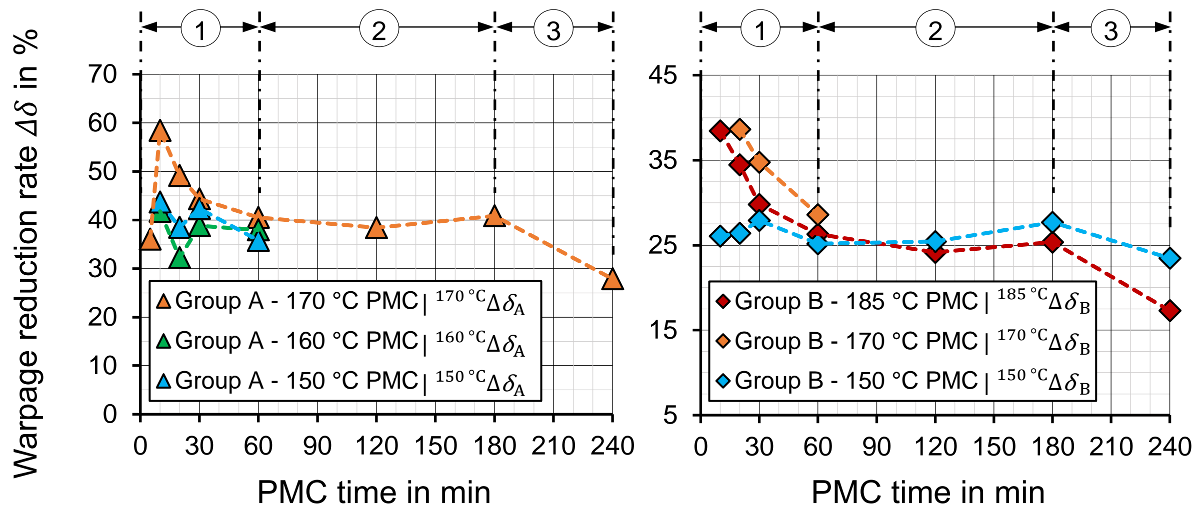

Figure 4.11: Intervals with different warpage reduction patterns

The first intervals of both groups consist of the range in which $\Delta\delta$ curves show both decreasing and constant trends depending on their T_{PMC} values. To begin with, the different values but same patterns for the control groups $^{170\,°C}\Delta\delta_{\mathrm{A}}$ and $^{170\,°C}\Delta\delta_{\mathrm{B}}$ can be explained with the different δ_0 values of both groups, which was shown in **Figure 4.6**. This difference with the same pattern shows that the $\Delta\delta$ value is also dependent on δ_0. The peaks in the curves with T_{PMC} values 170 °C and 185 °C in this interval can be explained via the stress relaxation of PCB substrates when they are post-mold cured with temperatures above $T_{\mathrm{g\,PCB}}$ = 165 °C (see **Figure 4.10** (4) for reference). For T_{PMC} values smaller than $T_{\mathrm{g\,PCB}}$, this effect cannot be utilized. Therefore, the samples with T_{PMC} values smaller than $T_{\mathrm{g\,PCB}}$ have smaller warpage reduction values within the first interval. Looking at the earliest point $^{170\,°C}_{5\,\mathrm{min}}\Delta\delta_{\mathrm{A}}$ in the same interval, it can be inferenced that due to the time needed for the part to heat up after being placed in the oven, the part could not reach $T_{\mathrm{g\,PCB}}$ yet. That means the PMC operation stopped in a point in between (1)-(4) in **Figure 4.10**, therefore ended up with a lower $\Delta\delta$. Narrowing down the focus only on the curves with a T_{PMC} larger than 165 °C in first interval, a decreasing $\Delta\delta$ pattern can be seen until t_{PMC} = 60 min. This decrease happens due to further curing of the PCB substrate at the temperatures above $T_{\mathrm{g\,PCB}}$. An increase in $T'_{\mathrm{g\,PCB}}$ decreases the thermal shrinkage of PCB substrate due to earlier transition into glassy state with a lower CTE value. Therefore, assuming a state in which l'_{EMC} has reached its maximum value due to maximized $T'_{\mathrm{g\,EMC}}$, a larger l'_{PMC} leads to a larger $\Delta l'$ value in concave warpage modes and ends up with a larger warpage value (see **Figure 4.10** Step 8 for reference). After 60 minutes of PMC time, the advantages of stress

relaxation in PCB substrate are neutralized due to this mechanism. The values for ${}^{T_{\mathrm{PMC}}>T_{\mathrm{g\,PCB}}}\Delta\delta$ are decreased almost to the same levels with the curves ${}^{T_{\mathrm{PMC}}<T_{\mathrm{g\,PCB}}}\Delta\delta$.

As next, the focus is switched on the second interval in which the $\Delta\delta$ curves show an almost constant behavior throughout the whole span. At this interval, both EMC and PCB cure reactions are stagnated for the curves ${}^{170\,°C}\Delta\delta_A$ and ${}^{185\,°C}\Delta\delta_B$, therefore no major effects related to chemical shrinkage or T_{g} shift are present. The small fluctuations might be due to the measurement and operation related tolerances and instabilities. For ${}^{150\,°C}\Delta\delta_{\mathrm{B}}$ on the other hand, the PCB cure is not reacted due to low temperature, and the cure reaction of EMC is again stagnated. The mentioned stagnation in cure reactions might be due to two main reasons: reaching the full cross-linking density in materials, or the occurrence of the vitrification phenomenon in which the material transforms into glassy state during curing due to T'_{g} value reaching T_{PMC} and causes a sudden decrease in reaction speed as introduced in Chapter 2.1.6.

Following the stabilized rates observed in the second interval, a decreasing $\Delta\delta$ trend was seen for all the sample groups within the third interval. When the PCB substrates are examined, a change in their color was seen as shown in **Figure 4.12** which indicate a degradation in PCB substrate that leads to a higher part warpage. In the substrates that are post-mold cured at 185 °C for 240 minutes, a relatively sharp change in color was seen, whereas the color change was not that obvious for the samples with T_{PMC} = 150 °C, which shows an accordance with the values ${}^{185\,°C}_{240\,\mathrm{min}}\Delta\delta_{\mathrm{B}}$ and ${}^{150\,°C}_{240\,\mathrm{min}}\Delta\delta_{\mathrm{B}}$.

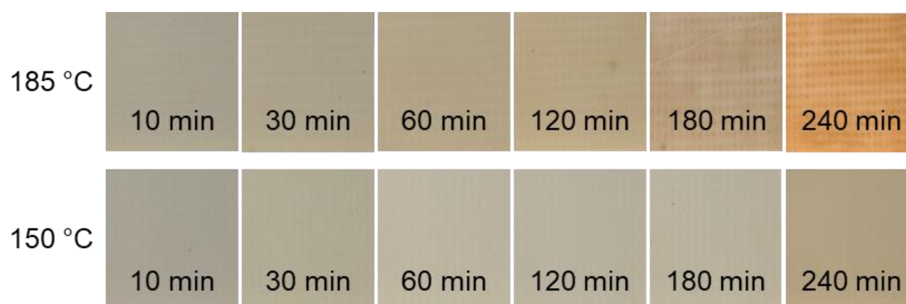

Figure 4.12: Color change in PCB substrates after PMC operations with different temperatures and times

Summarizing the results, it was seen that the maximum warpage reduction is reached after 10 minutes of PMC for the tested EMC encapsulants at the temperatures slightly higher than $T_{\mathrm{g\,PCB}}$. Further increase in T_{PMC} ended up with decreased warpage reduction performance. Similarly, increased t_{PMC} yielded lower warpage reduction

rates and even led to thermal degradation of the substrates. These results showed that the PMC operations conducted in industry for long PMC times (> 4h) can be optimized for lower final warpage values of the packages and a much faster PMC operation. Having lower PMC times and lower PMC temperatures, the cost per product is decreased as well while having a better quality in terms of package warpage. For the sample geometry and materials used in the experiments, the results showed that the optimal PMC parameters should be selected in a way that suffices for a full cure of EMC but prevents further curing of PCB substrate.

4.2.3 Results on Infrared Spectra

Even though investigating the cure behavior of EMC is not the focus point of this thesis, the cross-linking density of the encapsulant material is a major concern of PMC operations in general. Therefore, it should also be considered at some extend for PMC parameter optimization concerns as mentioned in Chapter 3.4. This chapter presents the IR spectra acquired from the ATR-FTIR Spectroscopy measurements for EMC material which are conducted before and after PMC. Looking at the differences in IR spectra, it is aimed to correlate the changes in chemical composition of the material with the changes in cross-linking density during PMC.

Figure 4.13 shows the IR spectra from EMC encapsulants of the molded samples before and after PMC operations at 170 °C for 30 minutes (a) and 120 minutes (b). As seen, the small changes in spectra are difficult to observe visually. Therefore, the detection of spectral changes is left to the machine learning algorithm.

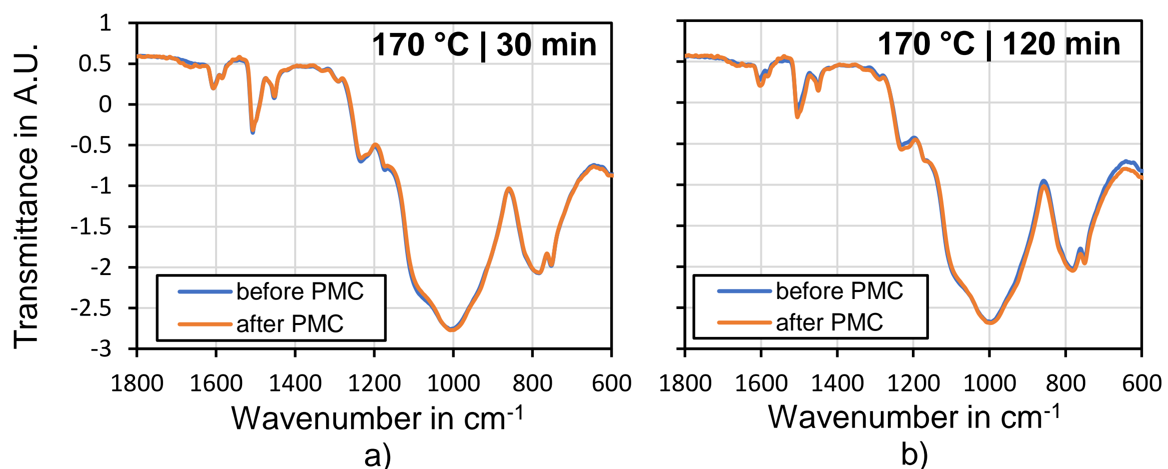

Figure 4.13: IR Spectra of EMC before and after PMC operations with a) 170 °C and 30 min b) 170 °C and 120 min

4.3 Results from Machine Learning Algorithm

After presenting and explaining the results acquired from the PMC experiments, the results from the machine learning model are shown and discussed in this Chapter. With reference to **Figure 3.15**, the reliability of the whole optimization cycle is dependent on the accuracies of predicted warpage and reaction enthalpy values. Therefore, to begin with, the outputs from warpage regression tool are presented and interpreted in Chapters 4.3.1 and 4.3.2 respectively. Similarly, the results for enthalpy regression are shown in Chapter 4.3.3, and then discussed in Chapter 4.3.4. Following them, Chapter 4.3.5 shows the predictions for optimal PMC parameters and Chapter 4.3.6 evaluates and reasons the results acquired from the whole optimization model.

4.3.1 Results from Warpage Regression

Being the focus point of the optimization algorithm, the predicted warpage values are presented, and prediction accuracy of the warpage regression tool is assessed in this chapter. As introduced in Chapter 3.4.2, an Artificial Neural Network is built to predict the final warpages δ' of the samples with the inputs of initial warpage δ_0 and initial IR spectrum IR_0 before PMC, and the PMC parameters T_{PMC} and t_{PMC}. **Figure 4.14** shows the data points that are used for training the warpage regression tool. Even in a simplified 3D domain that does not contain the features from IR spectra, no clear patterns on final warpages are present for visual inspections.

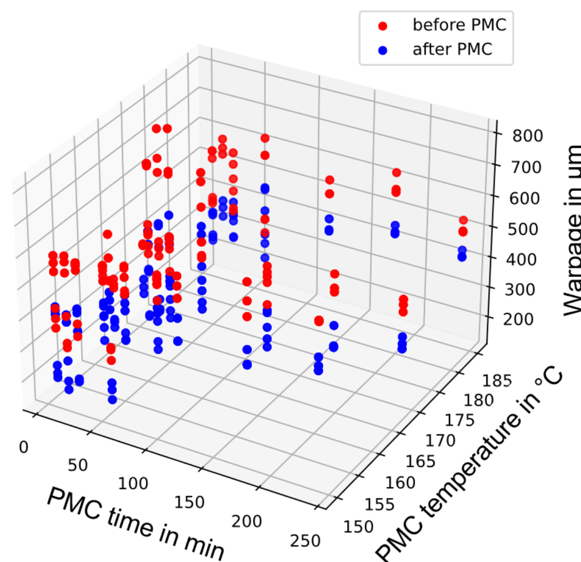

Figure 4.14: Data points used for training the warpage regression tool

To fit a pattern to above given scattered data points, different ANN architectures, learning rates and activation functions are tested for their accuracies as listed in **Table 4.1**. To assess their accuracies, their mean absolute error (MAE) values are calculated using Equation (2.5), with the values of measured and predicted final warpages. Looking at the MAE values, the parameter set ANN 2 is selected with the hidden layer structure of 150 : 90 : 40 neurons to be used in the model. The learning rate $1 \text{ x } 10^{-3}$ yielded the best result in comparison to other values tested. ReLU is used as the activation function due to its computational simplicity and slightly better results achieved with it.

Table 4.1: Tested ANN structures and their resultant mean absolute errors

Parameter set	Hidden layers	Learning rate	Epochs	Activation Function	MAE in µm
ANN 1	150 : 90 : 40	$1 \text{ x } 10^{-4}$	5000	ReLU	14.11
ANN 2	150 : 90 : 40	$1 \text{ x } 10^{-3}$	5000	ReLU	12.01
ANN 3	150 : 90 : 40	$1 \text{ x } 10^{-2}$	5000	ReLU	17.28
ANN 4	100 : 60 : 25	$1 \text{ x } 10^{-3}$	5000	ReLU	15.85
ANN 5	50 : 30 : 10	$1 \text{ x } 10^{-3}$	5000	ReLU	17.84
ANN 6	150 : 90 : 40	$1 \text{ x } 10^{-3}$	5000	sigmoid	12.33

The test samples are randomly selected amongst the samples from the PMC experiments and excluded from the train set. The final warpages of these samples are predicted using the initial measurements for warpage and IR spectrum, and the PMC parameters that are applied to them. The predicted values δ'_{pred} are then compared with the measured values δ' after their PMC operations in real-life. **Figure 4.15** shows the measured and predicted final warpage values of the test samples.

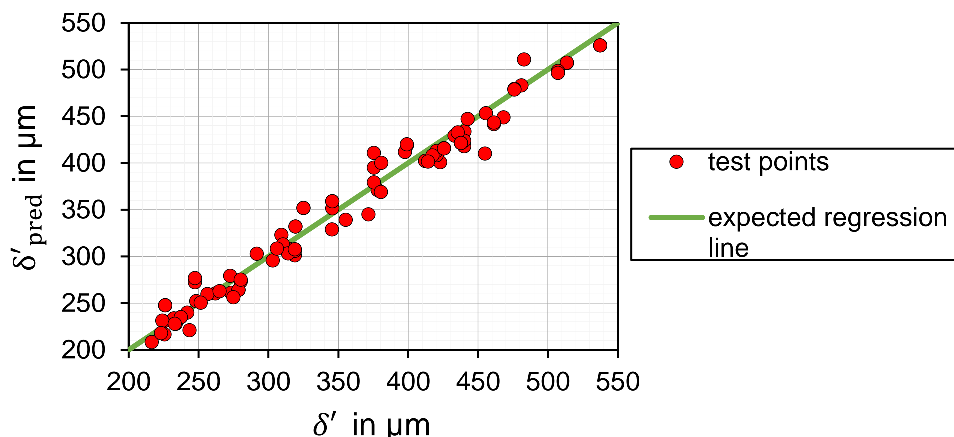

Figure 4.15: Comparison of measured and predicted final warpage values

4.3.2 Interpretation of Results from Warpage Regression

To begin with evaluating the ANN parameters, it can be seen in **Table 4.1** that the predictions got better as the complexity of hidden layer architecture increases. Having more weights and biases increased the model's ability to have better fits to nonlinear patterns. On the other hand, having more complex networks increases the computational costs. For the applications with large data sets, this is an important aspect to be considered. Since the data set used in this thesis is relatively small, the computational cost is not considered. The learning rate of $1 \text{ x } 10^{-3}$ yielded the best results in comparison to other values tested. This is mostly due to stagnating in a local minimum point of the loss function with the learning rate of $1 \text{ x } 10^{-4}$, and overshooting the optimal point with the learning rate of $1 \text{ x } 10^{-2}$.

The prediction accuracy of the selected ANN architecture with a MAE of 12.01 µm stands for only 3.4 % error with respect to average measured δ' value of 348 µm for the samples. Considering the 10 µm tolerance of the 3D Profilometer which was used for warpage measurements, the predictions showed a high accuracy despite the relatively low number of data points used for model training. These results showed that the usage of machine learning based methods for predicting the final warpage yields good results despite the manufacturing related tolerances and deviations of the parts, due to its allowances on including these uncontrollable and unmeasurable disturbances in the fitted pattern as well.

4.3.3 Results from Enthalpy Regression

Despite not being the main optimization concern in context of this thesis, the cross-linking density of the EMC had to be considered as a quality criterion as well since the main goal of PMC operations in general is reaching higher cross-linking densities. Therefore, using the residual reaction enthalpy values as a measurable indicator of cross-linking density, an enthalpy regression function is implemented in the optimization model, which predicts the residual reaction enthalpies as introduced in Chapter 3.4.3. This Chapter presents the results acquired from the enthalpy regression function which consists of two different subfunctions, to predict the initial and final residual reaction enthalpies of the samples as shown in **Figure 3.14**.

Before starting with presenting the predictions for EMC encapsulant, it should be remembered that the enthalpy prediction function was partly trained with an external data set that consists of DSC and ATR-FTIR measurements of pressed EMC samples with different cross-linking densities as mentioned in Chapter 3.4.1. Using the same ANN structure as the one used for warpage predictions, a model is created with a MAE of 1.27 J/g within the same external data set.

Since the model is trained with an external data set but not with the molded parts, its accuracy for molded parts should also be tested. To do so, the reaction enthalpy values of the molded DSC samples that are mentioned in Chapter 4.1.1 are predicted via their IR spectra using the first subfunction of the enthalpy regression tool (see **Figure 3.14** for reference), and then compared with their DSC measurements. **Figure 4.16** shows the measured and predicted reaction enthalpies plotted over molding time, with the error bars indicating the standard deviations in the DSC measurements. As seen, the predicted values show a similar pattern with the measured values, with a sudden decrease until the mold cycle time 110 s, and then a stagnation almost in the constant values throughout the whole span. An offset can be seen in between both curves, which becomes more stable and larger after 100 s. The overall MAE value is 0.67 J/g for the test interval of 55 s – 300 s. Since no parts could be produced with cycle times less than 55 s due to limitations on EMC solidification as mentioned in Chapter 3.3.1, the model could not be tested with molded samples with higher residual reaction enthalpies.

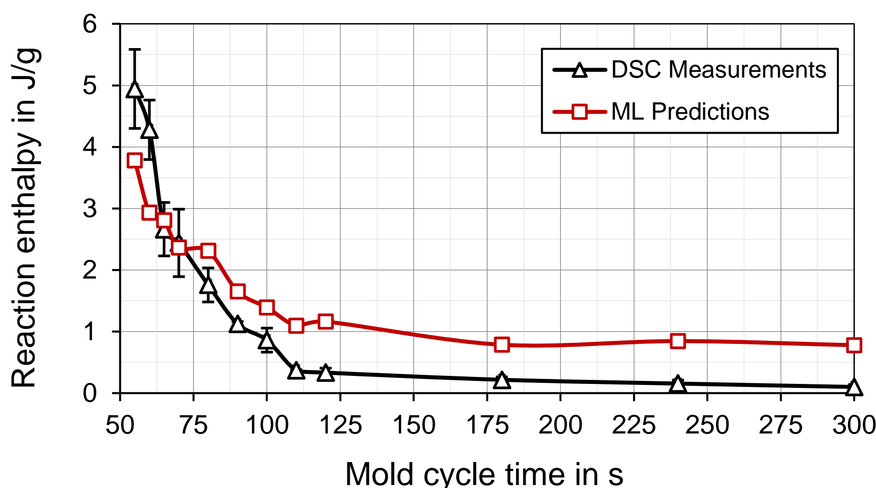

Figure 4.16: Predicted and measured reaction enthalpy values for EMC encapsulants of the molded parts

As the next step after assessing the accuracy of enthalpy prediction function, reaction enthalpies of the molded samples from PMC experiments are predicted for their states before and after PMC. The initial residual reaction enthalpies (h_0) of the parts are predicted through their initial IR spectra (IR_0) using the first subfunction of enthalpy regression tool (see **Figure 3.14** for reference). Referencing Chapter 4.2.3, no observable changes in IR spectra before and after PMC were present for visual investigations. To test if the machine learning algorithm is able to detect those minor changes, the final reaction enthalpies (h') are also predicted using the same subfunction but with the input of final IR spectra (IR'). **Figure 4.17** shows the predicted reaction enthalpy values before and after PMC, plotted over the corresponding sample numbers. As seen, the model was able to correlate the small changes in between IR_0 and IR' with the cross-linking densities and predicted lower h' values for IR'. But, since no IR' data will be available in real life applications before PMC, the second subfunction is used which predicts for h' values with the inputs of IR_0, T_{PMC}, t_{PMC} as explained in Chapter 3.4.3 (see **Figure 3.14** for reference). Comparing the h'_{pred} values acquired from both models, it can be seen that both were able to respond to increasing cross-linking densities in a meaningful way and predict lower values for h'.

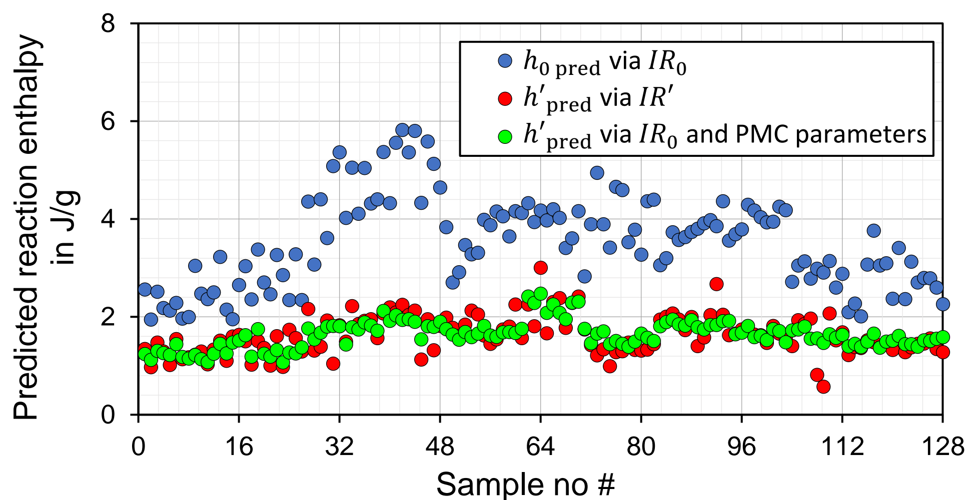

Figure 4.17: Average predicted reaction enthalpy values of molded parts before and after PMC

Despite being able to detect the changes in IR spectra, the predicted $h_{0\ pred}$ values are fluctuating a lot. Looking at the h'_{pred} values on the other hand, the fluctuations are diminished, and they are following a constant pattern in which no clear decrease or increase in predictions are seen in between different samples despite being post-mold cured with different parameters.

4.3.4 Interpretation of Results from Enthalpy Regression

After showing the prediction results for reaction enthalpies, this chapter interprets and explains the results. Starting with the first predictions shown in **Figure 4.16**, a similar pattern was seen in both measurements and predictions. The increased offset in between the curves after 100 s is mainly due to external train data which does not include many data points with really low reaction enthalpy values, which means that the algorithm was not trained with enough amount of spectra with almost ~0 J/g. Therefore, it was not able to predict accurately for the samples which converge to 0 J/g in their DSC measurements. The overall MAE value is calculated as 0.67 J/s, which stands for a good accuracy when compared with the tolerances of DSC method on measuring reaction enthalpies (up to 3 J/g).

Looking at **Figure 4.17**, meaningful predictions were made for reaction enthalpy values before and after PMC operation. Both methods to predict for h' yielded similar results with similar patterns, but the predicted $h_{0\ \mathrm{pred}}$ values fluctuated a lot. One main reason for these fluctuations in $h_{0\ \mathrm{pred}}$ is the manual molding operations. Even though a high attention was paid to conduct all the molding operations as equally as possible, unavoidable fluctuations in raw material placement and part ejection times might have ended up with different cross-linking densities due to fast curing behavior of the EMC material used. And since the molding machine is located in a workshop but not in a controlled laboratory, the small ambient changes might have minor effects on the cooling process as well when the parts are left for cooling in the ambient conditions after being ejected. The changes in $h_{0\ \mathrm{pred}}$ values can therefore be mostly reasoned by combination of these difficulties.

Looking at the predicted h'_{pred} values on the other hand, despite having different PMC parameters, all the predictions for different samples are more or less in the same levels since the EMC has already reached its almost fully cross-linked structure in the first 10 minutes of the PMC. Due to insufficient train data, the predictions were not able to converge to 0 J/g as mentioned above. This fast-curing behavior makes the algorithm hard to utilize for optimizing the cross-linking densities for the used EMC material. But, it can be seen that the fluctuations in predicted values are diminished once all the samples have the same cross-linking densities (their fully cured states in this case),

which shows the reproducibility and stability of the function once the above-mentioned production related tolerances are minimized.

To sum up, the reaction enthalpy predictions showed that the method is promising as it was able to detect small spectral changes that occur during PMC and make meaningful predictions for the decreasing final reaction enthalpies via PMC parameters and initial IR spectra. But, the fast curing behavior of the EMC material used in the experiments was a major limiting factor both for creating train data and producing molded test samples with different cross-linking densities. Therefore, no clear observations could be made for different PMC parameters and their effects on spectral changes to be correlated with the cross-linking densities. The use of a such prediction model for materials that have slower cure behaviors would yield more beneficial results and could be used in scenarios in which the effects of different PMC parameters on cross-linking density of the material could be investigated in a larger process window.

4.3.5 Results from Optimization Model

As described in Chapter 3.4.4, having the measured initial part warpage and IR spectrum of an EMC encapsulant, the machine learning model aims to predict the optimal PMC parameters that will yield the desired part warpage and EMC cross-linking density after the operation. As explained in Chapter 3.4.4, the optimization is done iteratively via predicting the final part warpage and final reaction enthalpy value of EMC encapsulant after PMC, with the cooperation in between regression functions and the optimization tool as illustrated previously in **Figure 3.15**. Following the chapters that present the prediction results and assess the accuracies of the mentioned regression functions, this chapter introduces the results acquired from the whole machine learning model for optimization.

Starting with setting the user defined parameters of the algorithm, the weights of quality and efficiency terms should be determined first to be used for the cost function that was introduced in Equation (3.2). There are no standard methods to determine the optimal weights as it highly depends on how the cost function was defined and how it is used in the algorithm. Therefore, the weights are determined in a trial-and-error manner, with the logic of having the cost function affected more by the important terms, and less by the terms of lower importance. In this case, warpage and PMC time are classified as the important terms since the main goal is to decrease them both. The

learning rate η and the momentum β terms are again determined through several tries to avoid overshooting or stagnating in irrelevant local minimum points. The limit Δ for the cost difference in between two iterations is defined, which stops the loop once the Δ gets smaller than the defined limit. **Table 4.2** lists the selected parameters for optimization model.

Table 4.2: Parameter set used for optimization model

w_δ	w_h	w_T	w_t	η	β	Δ
1.2	0.02	0.01	1.2	5 x 10^{-5}	0.1	1

Using the program defined with the above given parameter set, the prediction of optimal PMC parameters for a random sample from the Sample Group B is tested. **Table 4.3** shows the properties of the tested sample, such as its measured initial warpage δ_0, predicted initial reaction enthalpy $h_{0\ \mathrm{pred}}$, the PMC parameters that are applied to the sample during the experiments, and the resultant values δ' and h' acquired after its PMC operation in real life. For testing reasons, the desired values for quality criteria are set as $\delta'_{\mathrm{desired}}$ = 300 µm and h'_{desired} = 2 J/g, and the optimization cycle is then tested to predict for the optimal PMC parameters that would yield the desired values. To test the model's stability, different iteration start points $T_{0\ \mathrm{itr.}}$ for PMC temperature and $t_{0\ \mathrm{itr.}}$ for PMC time are used in different optimization tests.

Table 4.3: Predictions for optimal PMC parameters of a test sample

Properties of test sample					
δ_0	$h_{0\ \mathrm{pred}}$	T_{PMC}	t_{PMC}	δ'	h'_{pred}
573 µm	3.93 J/g	185 °C	180 min	433 µm	1.54 J/g
Desired quality criteria					
$\delta'_{\mathrm{desired}}$		300 µm	h'_{desired}		2 J/g
Optimization tests					
Parameters		**(1)**	**(2)**	**(3)**	**(4)**
$T_{0\ \mathrm{itr.}}$		150 °C	185 °C	150 °C	185 °C
$t_{0\ \mathrm{itr.}}$		0 min	240 min	0 min	240 min
$T_{\mathrm{PMC\ opt.}}$		204 °C	237 °C	174 °C	174 °C
$t_{\mathrm{PMC\ opt.}}$		0 min	0 min	15 min	15 min
δ'_{pred}		310 µm	279 µm	341 µm	341 µm
$\Delta\delta_{\mathrm{pred}}$		45.8 %	51.3 %	40.6 %	40.6 %
h'_{pred}		3.6 J/g	3.2 J/g	1.8 J/g	1.8 J/g
iterations #		2885	3499	723	1092

Looking at the results for optimization tests (1) and (2) at **Table 4.3**, it can be seen that the model predicted the optimal parameters as $T_{\mathrm{PMC\,opt.}}$ = 204 °C and $t_{\mathrm{PMC\,opt.}}$ = 0 min for test (1), and $T_{\mathrm{PMC\,opt.}}$ = 237 °C and $t_{\mathrm{PMC\,opt.}}$ = 0 min for test (2). After seeing that the model was not able to detect the boundary condition for t_{PMC} = 0 min as *no PMC* state, artificial data points are added to train set in order to define important boundary conditions for *no PMC* condition t_{PMC} = 0 min, or safety limit for electronic components T_{PMC} = 200 °C. It was aimed to increase the cost towards these regions with dummy data so that the algorithm does not try to exceed them while iterating. The predictions are then repeated after training the warpage regression tools with the extended data set. As seen in tests (3) and (4), despite starting their iterations from different points, both (3) and (4) ended up with the same PMC parameters $T_{\mathrm{PMC\,opt.}}$ = 174 °C and $t_{\mathrm{PMC\,opt.}}$ = 15 min, and predicted the same values h'_{pred} = 1.8 J/g and δ'_{pred} = 341 µm, which corresponds to a warpage reduction rate of ${}^{174\,°\mathrm{C}}_{15\,\mathrm{min}}\Delta\delta_B$ = 40.6 %. **Figure 4.18** shows the predicted optimal point for the tested sample and the results acquired from the PMC experiments of Sample Group B, in which the tested sample belongs to.

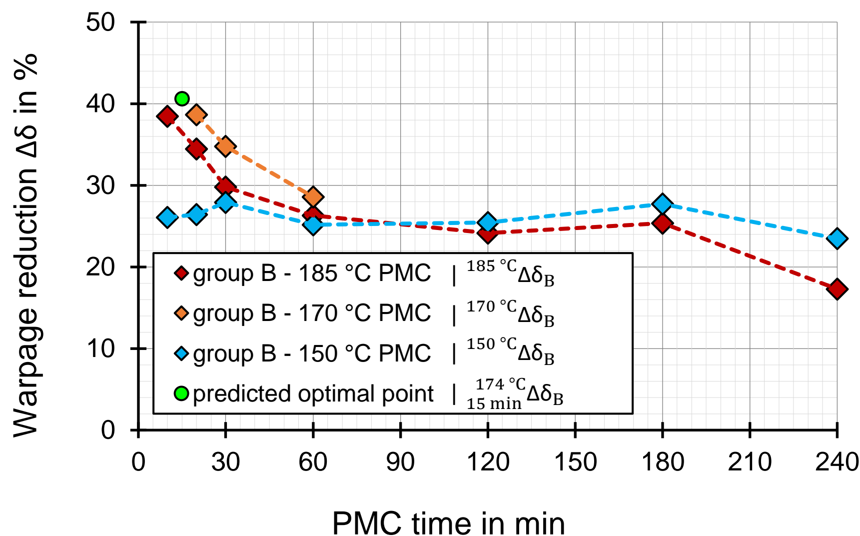

Figure 4.18: Comparison of predicted optimal point with experimental results

4.3.6 Interpretation of Results from Optimization Model

To visualize the optimization process for ease of interpretation, the iteration steps for optimization test (2) and (4) are plotted in **Figure 4.19** (a) and (b) respectively. The most upper plots show the iterated PMC parameters T_{PMC} and t_{PMC} over the iteration steps. In the graphs below them, the predicted δ'_{pred} and h'_{pred} values for these

iterated PMC parameters are shown, as well as the calculated cost for the corresponding steps in the most bottom subplot. The iterations were started from the values $T_{0\,\mathrm{itr.}}$ = 185 °C and $t_{0\,\mathrm{itr.}}$ = 240 min as defined in **Table 4.3**, and after each iteration these parameters are updated in the direction of reducing cost.

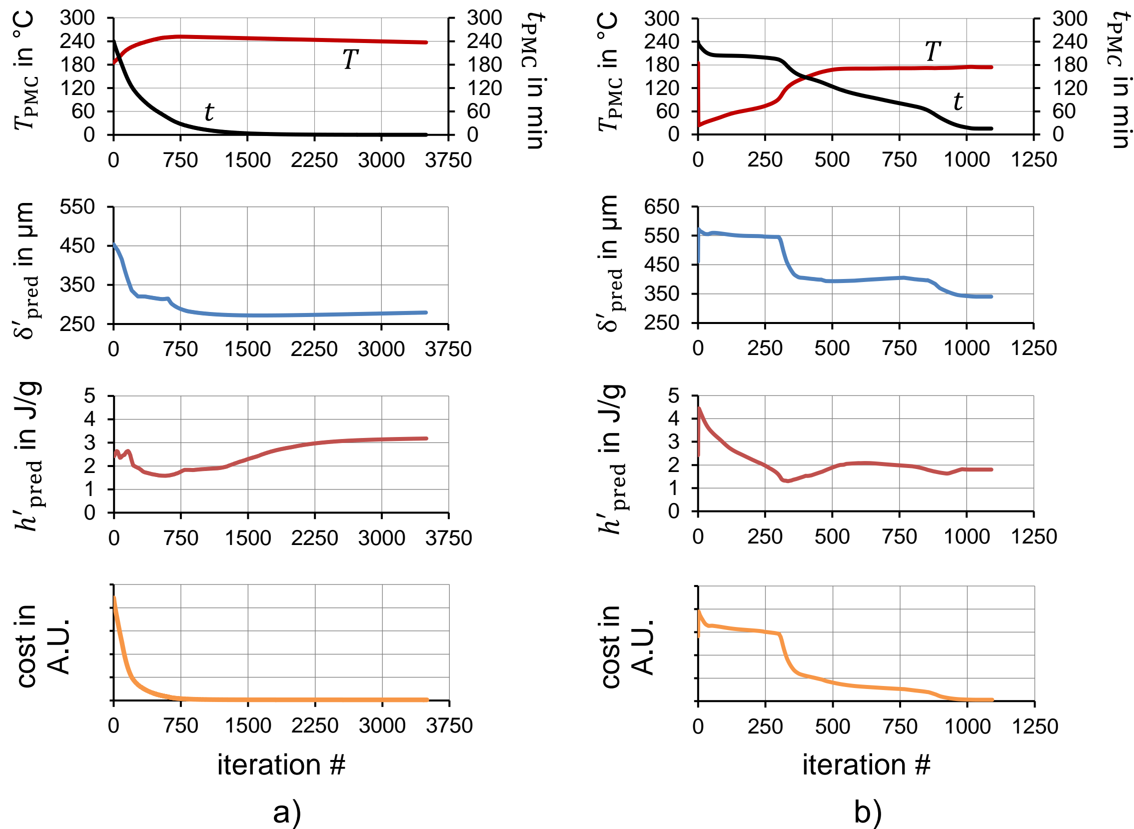

Figure 4.19: Iteration steps from a) optimization test (2) without dummy data and b) optimization test (4) with dummy data

As seen in **Figure 4.19** (a), since the upper temperature limit (~200 °C) for safety of electronic components was not defined before the dummy data points, the model tried to increase T_{PMC} further than that value. Also, when T_{g} values of both EMC and PCB components are reached, further increases in T_{PMC} values yielded larger final warpages in experiments as described in Chapter 4.2.2. But, since no samples with T_{PMC} > 185 °C were included in the train set, the model was not able to correlate that relationship from the iteration start point of $T_{0\,\mathrm{itr.}}$ = 185 °C and predicted higher warpage reduction rates for T_{PMC} values higher than 240 °C. As t_{PMC} = 0 min was not defined as *no PMC* condition in train data, it was recognized as a proper PMC operation which is not the case in reality. Therefore, the model also predicted a warpage reduction for t_{PMC} = 0 min and decreased the iterated t_{PMC} to 0 min to

minimize the cost which was defined as in Equation (3.2). Similarly, due to these insufficiencies, the predicted reaction enthalpy values showed no clear accordance with the expected pattern.

To overcome this training related issues, artificial points such as ${}^{170\,°\mathrm{C}}_{0\,\mathrm{min}}\Delta\delta = 0$ % or ${}^{200\,°\mathrm{C}}_{10\,\mathrm{min}}\Delta\delta = 0$ % added to train data to define the important boundary conditions. When the dummy points are also included in train set, the predictions got much more accurate as seen in **Figure 4.19** (b). The iterated T_{PMC} value is first decreased rapidly from 185 °C to ~30 °C and then slowly increased to 174 °C. The iterated t_{PMC} followed a constantly decreasing trend until t_{PMC} = 15 min and then stagnated. The predicted δ'_{pred} values are decreased with decreasing t_{PMC} to the value 341 µm in the end. The corresponding warpage reduction rate at the predicted optimal point ${}^{174\,°\mathrm{C}}_{15\,\mathrm{min}}\Delta\delta_B$ = 40.6 % showed an accordance with the patterns seen in experimental values as previously shown in **Figure 4.18**. Even though the desired warpage value of 300 µm could not be reached due to limitations of PMC on warpage reduction, an optimal parameter set is suggested that yields the closest possible value to that.

Looking at h'_{pred} values, the predictions were higher when the iterated T_{PMC} was decreased to ~ 30 °C, which is below $T_{\mathrm{g\,EMC}}$. Since no curing is expected at that temperature, having the predicted h'_{pred} value close to initial value $h_{0\,\mathrm{pred}}$ showed that the algorithm was able to fit a reasonable pattern for reaction enthalpy predictions. With increasing iterated T_{PMC}, the predicted h'_{pred} value decreased. When the iterated PMC parameters that suffice for an almost full cure of EMC are reached, the predictions are stagnated more or less in the same levels. That means, once the required T_{PMC} is reached during iterations, all the iterated steps corresponded to almost fully cross-linked states since all the iterated t_{PMC} values were above 10 min, which suffices for a full cure of EMC. And even though the expected reaction enthalpy value was 0 J/g for these fully cured states, the model could not converge to that point due to insufficiencies in train data as explained in detail in Chapter 4.3.4, and stagnated in 1.5 – 2 J/g interval which stands for the minimum prediction offset as previously shown in **Figure 4.17**.

These results pointed out one weakness of ANN algorithms: despite being able to fit accurate patterns to nonlinear data distributions, it is still a challenge for an ANN algorithm to make good predictions for the data regions which are outside their train

ranges. As an example, the warpage regression tool showed a good accuracy with a MAE of 12.01 µm for the PMC parameters which were used in the train data that was created via PMC experiments. But when the model is asked to predict for PMC parameters which are outside its train range during the iterations for optimization, the prediction accuracy got worse. Then, when the dummy data are included in train set, the predictions got much more accurate. These results showed that once the required train data is available, the ANN is a proper method that can yield reliable predictions for such tasks. To achieve this, task specific boundary conditions had to be defined which pointed out the need for having a large amount of train data from different provenances, feature regions and boundaries, which makes the use of ANN costly for many applications.

To sum up, despite the low amount of train data, the optimization model was able to find the optimal PMC parameter region once the important boundary conditions were defined with artificial points. The predicted warpage values showed an accordance with the experiment results. And the predictions for residual reaction enthalpies showed a decreasing pattern as expected, but predictions with high accuracies could not be reached due to insufficient training. The fast-curing nature of the used EMC material was a major limitation on this due to challenges of manually producing samples with different post-mold cure durations in such limited time windows. Due to these issues, the cross-linking density of EMC could not be properly implemented in the optimization algorithm as a limiting criterion. With the use of different encapsulant materials with slower cure behaviors, the suggested methodology can be tested for more complex optimization problems.

5 Conclusion & Outlook

In this thesis, the effects of post-mold cure (PMC) process on warpage of electronic packages are investigated. The real-life industrial PMC operations have the main goal of reaching full cross-linking densities of thermoset encapsulants in general, and therefore mostly conducted at the same elevated temperatures (~175 °C) for long times (> 4h). But the effects of different PMC parameters on warpage are not well known. Focusing on this lack of process knowledge, the aim of this study was to investigate the effects of PMC operation on warpage and to optimize PMC parameters accordingly. To do so, PMC experiments are conducted with different PMC temperatures and times, and a machine learning based optimization algorithm is created using the data acquired from these experiments. Warpage is considered as the main quality criterion to optimize the PMC parameters for, and the cross-linking density is also regarded as the secondary optimization concern since it is the main goal of PMC operations in general.

Experiments are conducted with molded samples that consist of two main components, namely the EMC encapsulant and PCB substrate. To characterize their thermal and cure behaviors, DSC and TMA measurements are conducted. The DSC measurements showed that the main encapsulant material EMC has a fast cure reaction so that it reaches an almost fully cross-linked structure after molding for 110 s at 170 °C. The glass transition temperature of raw EMC is evaluated as ~42 °C, and it increased to ~177 °C after reaching its full cross-linking density. The reaction enthalpy values acquired from DSC measurements are used as a measurable indicator of EMC cross-linking density. The externally supplied PCB substrates were assumed to be finished products at their fully cured states in context of this work prior to their DSC measurements. But DSC measurements revealed the cure reactions in glass epoxy of PCB substrate, with an increase in glass transition temperature from ~165 °C to ~175 °C during the PMC operations.

During PMC experiments, different PMC temperatures (T_{PMC}) and PMC times (t_{PMC}) are tested to investigate their effects on part warpage. The results showed that the PMC operations decreased the warpage in all the samples due to decreased thermal mismatch in between EMC and PCB substrate. It was also observed that selecting the T_{PMC} values above $T_{\mathrm{g\,PCB}}$ ~ 165 °C had an improving effect for warpage reduction by

benefiting the stress relaxation in PCB substrate as a result of transitioning into its gel state. However, longer t_{PMC} ended up with worse results for part warpage when T_{PMC} was above $T_{\mathrm{g\,PCB}}$. The reason behind this pattern was found to be the PCB curing. This pointed out that the optimal PMC parameters are also dependent on the possible curing of the substrates in the package, other than the main encapsulant material. With all the mentioned considerations, the results showed that the PMC parameters for this material set should be selected in a way that it suffices for the full cross-linking density of EMC but hinders further curing of PCB substrate to reach a maximum reduction in warpage. Since the samples are heated without conductive heat transfer in a fan oven during PMC, the additional time needed to heat the sample from T_{room} to T_{PMC} caused a relatively large difference in t_{PMC} and t_{mold} values required to reach full cross-linking densities. The parameter set T_{PMC} = 170 °C and t_{PMC} = 10 min yielded the best results in terms of warpage reduction for the materials and part geometry used in the experiments, which are molded with T_{mold} = 170 °C and t_{PMC} = 65 s.

As the next step, it was aimed to detect the changes in IR spectra before and after PMC, and then correlate these changes with further cross-linking of the material during PMC as a non-destructive method to evaluate cross-linking density. But, no clear changes were detected in IR spectra before and after PMC via visual inspections and further investigations are tested via the machine learning algorithms.

To predict the final part warpage, and also the final reaction enthalpy as an indicator of EMC cross-linking density, machine learning based regression functions are created using artificial neural network (ANN) algorithms as described in Chapter 3.4. The predicted warpage values had the mean absolute error (MAE) of only 12.01 µm for the test points that are randomly selected amongst the samples from PMC experiments. The enthalpy predictions showed an accuracy of 0.67 J/g in MAE and the model was able to detect the differences in between the IR spectra before and after PMC, which was not detectable with visual inspections. The decreases in reaction enthalpies after PMC were successfully predicted. As a major limiting factor regarding this point, due to fast-curing behavior of the EMC material, it was not possible to properly observe the changes in cross-linking density of EMC within a very limited time window for PMC. As a result, the studies on cross-linking density as a function of PMC parameters could not be properly implemented in the model since all the samples that are post-mold cured have reached their almost full cross-linking densities even after t_{PMC} of

10 minutes. Therefore, the predicted final reaction enthalpy values were always in the same levels for all the PMC parameters tested in the experiments.

The optimization of PMC parameters is done iteratively using the above-mentioned regression functions. Despite the assessed prediction accuracies in their train intervals, when the regression functions were asked to predict for iterated PMC parameters which are outside their train ranges, the prediction results for final warpage and enthalpy got worse. Thus, the predicted optimal PMC parameters did not show any accordance with the observations from the experiments. This pointed out one major weakness of ANN algorithms, that they require large amount of train data from different feature regions and sample provenances. To overcome this issue, artificial data points are included in the train data to define some important boundary conditions as described in Chapter 4.3.5. With the extended train data, the prediction quality for optimal PMC parameters improved and converged to parameters of T_{PMC} = 174 °C and t_{PMC} = 15 min which showed an accordance with the results acquired from the experiments. With larger train data from different parameter regions, one can expect predictions which show better accordance with the real-life applications.

As an outlook, a detailed inspection of PCB substrate curing and its effects on resultant part warpage is left as an open field of research for future studies. Package designs with more than two thermosetting components would lead to different aspects for further investigations on effecting mechanisms and design related optimization concerns. The fast-curing nature of the EMC material used in this study led to a relatively simple optimization problem, in which the optimal points for both warpage reduction and process costs were in the same direction of decreasing PMC time. Further studies on the materials with much slower cure behaviors might therefore find answers to more complex optimization problems, in which the elongated t_{PMC} needs for desired EMC cross-linking density also contributes to cost. Furthermore, studies on such materials would also allow further investigations regarding the relationship in between the IR spectra and cross-linking densities before and after PMC. To improve the prediction accuracies, different machine learning algorithms such as reinforcement learning can be used to overcome the limitations of supervised methods for predicting for the values that are outside their train ranges.

References

[1] Barrett J., 1998,“Electronic systems packaging: future reliability challenges,” Microelectronics Reliability, **38**(6-8), pp. 1277–1286.

[2] Komori S., and Sakamoto Y., 2009, Development Trend of Epoxy Molding Compound for Encapsulating Semiconductor Chips, *Materials for Advanced Packaging*, Lu D., and Wong C. P., eds., Springer US, Boston, MA, pp. 339–363.

[3] Sun Z., Huang W., and Xie Z., 2020,“Failure Analysis on Packaging Materials Fault for Plastic package IC,” *2020 21st International Conference on Electronic Packaging Technology (ICEPT)*, IEEE, pp. 1–3.

[4] Buchwalter S. L., 1999, Microelectronic Encapsulation and Related Technologies, *Conductive Polymers and Plastics*, Elsevier, pp. 121–125.

[5] Kaya B., 2018,“Concept Development and Implementation of Online Monitoring Methods in the Transfer Molding Process for Electronic Packages,” Dissertation, Technical University of Berlin, Berlin.

[6] Stark W., Goering H., Michel U., and Bayerl H., 2009,“Online monitoring of thermoset post-curing by dynamic mechanical thermal analysis DMTA,” Polymer Testing, **28**(6), pp. 561–566.

[7] Mengel M., Mahler J., and Schober W., 2004,“Effect of Post-mold Curing on Package Reliability,” Journal of Reinforced Plastics and Composites, **23**(16), pp. 1755–1765.

[8] Ardebili H., Zhang J., and Pecht M. G., 2019, Encapsulation defects and failures, *Encapsulation Technologies for Electronic Applications*, Elsevier, pp. 259–315.

[9] Kong J., Kim J.-K., and Yuen M., 2003,“Warpage in plastic packages: effects of process conditions, geometry and materials,” IEEE Trans. Electron. Packag. Manufact., **26**(3), pp. 245–252.

[10] Sun P., Leung V. C.-K., Xie B., Ma V. W., and Shi D. X.-Q., 2008,“Warpage reduction of package-on-package (PoP) module by material selection & process

optimization," *2008 International Conference on Electronic Packaging Technology & High Density Packaging*, IEEE, pp. 1–6.

[11] Shue C.-H., Hwang S.-J., Lee H.-H., Huang D.-Y., and Lee Y.-J., 2008,"Post-mold cure process simulation of IC packaging," *2008 International Conference on Electronic Materials and Packaging*, IEEE, pp. 106–110.

[12] Chew S., and Tan A.-C., 2003,"Evaluating the influence of post-mould cure and additives on the viscoelastic properties of a low stress epoxy mould compound," IEEE Trans. Electron. Packag. Manufact., **26**(3), pp. 211–215.

[13] Tong K., Kwong C., and Ip K., 2003,"Optimization of process conditions for the transfer molding of electronic packages," Journal of Materials Processing Technology, **138**(1-3), pp. 361–365.

[14] Wang H., Ma C., and Zhou L., 2009,"A Brief Review of Machine Learning and Its Application," *2009 International Conference on Information Engineering and Computer Science*, IEEE, pp. 1–4.

[15] Weichert D., Link P., Stoll A., Rüping S., Ihlenfeldt S., and Wrobel S., 2019,"A review of machine learning for the optimization of production processes," Int J Adv Manuf Technol, **104**(5-8), pp. 1889–1902.

[16] Tercan H., Guajardo A., and Meisen T., 2019,"Industrial Transfer Learning: Boosting Machine Learning in Production," *2019 IEEE 17th International Conference on Industrial Informatics (INDIN)*, IEEE, pp. 274–279.

[17] Prisacaru A., Guerrero E. O., Gromala P. J., Han B., and Zhang G. Q., 2019,"Degradation Prediction of Electronic Packages using Machine Learning," *2019 20th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE)*, IEEE, pp. 1–9.

[18] Smola A., and Vishwanathan S., 2020, Introduction to Machine Learning, *Machine Learning Refined*, Watt J., Borhani R., and Katsaggelos A., eds., Cambridge University Press, pp. 1–18.

[19] Pecht M., Ardebili H., and Zhang J., 2018. *Encapsulation Technologies for Electronic Applications*, William Andrew Publishing.

[20] Pecht M. G., Nguyen L. T., Hakim E. B., and Rafanelli A. J., 1997,“Plastic Encapsulated Microelectronics; Materials, Processes, Quality, Reliability, and Application,” Journal of Electronic Packaging, **119**(2), pp. 144–145.

[21] Hammond P. A., 2004,“Encapsulation of a liquid-sensing microchip using SU-8 photoresist,” Microelectronic Engineering, **73-74**, pp. 893–897.

[22] Ardebili H., Zhang J., and Pecht M. G., 2019, Introduction, *Encapsulation Technologies for Electronic Applications*, Elsevier, pp. 1–45.

[23] Chung D. D. L., 1995. *Materials for electronic packaging*, Butterworth-Heinemann, Boston.

[24] Kelly G., Alderman J., Lyden C., and Barrett J., 1997,“Microsystem packaging: lessons from conventional low cost IC packaging,” J. Micromech. Microeng., **7**(3), pp. 99–103.

[25] Liu Y., and Kinzer D., 2011,“Challenges of power electronic packaging and modeling,” *2011 12th Intl. Conf. on Thermal, Mechanical & Multi-Physics Simulation and Experiments in Microelectronics and Microsystems*, IEEE, 1/9-9/9.

[26] Wan Y.-J., Li G., Yao Y.-M., Zeng X.-L., Zhu P.-L., and Sun R., 2020,“Recent advances in polymer-based electronic packaging materials,” Composites Communications, **19**, pp. 154–167.

[27] Tong X. C., Ed., 2011. *Advanced Materials for Thermal Management of Electronic Packaging*, Springer New York, New York, NY.

[28] Sasajima H., Watanabe I., Takamoto M., Dakede K., Itoh S., Nishitani Y., Tabei J., and Mori T., 2017, New Development Trend of Epoxy Molding Compound for Encapsulating Semiconductor Chips, *Materials for Advanced Packaging*, Lu D., and Wong C. P., eds., Springer International Publishing, Cham, pp. 373–419.

[29] Belton D., 1987,“The Effect of Post-Mold Curling Upon the Microstructure of Epoxy Molding Compounds,” IEEE Trans. Comp., Hybrids, Manufact. Technol., **10**(3), pp. 358–363.

[30] Vreugd J., 2011,“The effect of aging on molding compound properties,” PhD Thesis, TU Delft, Netherlands.

[31] Kiong T. S., Ruzaini I., Kesvakumar, and Seng F. C., 2010,“Package warpage challenges for LQFP 144 lead CMOS 90 device and it's impact to lead coplanarity,” *2010 34th IEEE/CPMT International Electronic Manufacturing Technology Symposium (IEMT)*, IEEE, pp. 1–6.

[32] Kim W. G., and Lee J. Y., 2002,“Curing characteristics of epoxy resin systems that include a biphenyl moiety,” J. Appl. Polym. Sci., **86**(8), pp. 1942–1952.

[33] Liu H., Tan W., Li L., Qiu C., Cheng X., and Han J., 2014,“Influence of Coupling Agent on the Performance of Epoxy Molding Compound,” ECS Trans., **60**(1), pp. 781–786.

[34] Giani M., 2021,“Optimierung von augewählte Prozessparametern für das Transfer Molding Verfahren durch Machine Learning basierten Modellen.,” Master's Thesis, Institut für Kunststoffverarbeitung, RWTH Aachen, Aachen.

[35] Fang K., 2019, Encapsulation process technology, *Encapsulation Technologies for Electronic Applications*, Elsevier, pp. 123–181.

[36] Ma S., Li M., Xiao Z., Huang X., and Yu D., 2017,“3D wafer level compression molding process development for image sensor package,” *2017 18th International Conference on Electronic Packaging Technology (ICEPT)*, IEEE, pp. 505–509.

[37] Ardebili H., Zhang J., and Pecht M. G., 2019, Injection molding, *Encapsulation Technologies for Electronic Applications*, Elsevier, pp. 183–194.

[38] Becker K.-F., Joklitschke D., Braun T., Koch M., Thomas T., Schreier-Alt T., Bader V., Bauer J., Nowak T., Bochow-Ness O., Aschenbrenner R., Schneider-Ramelow M., and Lang K.-D., 2011,“Transfer molding technology for smart power electronics modules - materials and processes,” International Symposium on Microelectronics, **2011**(1), pp. 83–89.

[39] Spoelstra A. B., 1992,“Transfer moulding of reactive materials: application to encapsulation of integrated circuits,” EngD Thesis, Stan Ackermans Instituut, Technische Universiteit Eindhoven, Netherlands.

[40] Garete A. J. H., Fadullo M. F., and Roscain R. J. S., 2019,“Epoxy Mold Compound Curing Behavior and Mold Process Cure Time Interaction on Molded Package Performance,” *2019 IEEE 21st Electronics Packaging Technology Conference (EPTC)*, IEEE, pp. 311–315.

[41] Memis I., 2005,“Testing to eliminate reliability defects from electronic packages,” *2005 IEEE International Reliability Physics Symposium, 2005. Proceedings. 43rd Annual*, IEEE, pp. 462–465.

[42] Frear D. R., Ramanathan L. N., Jang J.-W., and Owens N. L., 2008,“Emerging reliability challenges in electronic packaging,” *2008 IEEE International Reliability Physics Symposium*, IEEE, pp. 450–454.

[43] Ernst L. J., Jansen K., Saraswat M., van 't Hof C., Zhang G. Q., Yang D. G., and Bressers H., 2006,“Fully Cure-Dependent Modeling and Characterization of EMC's with Application to Package Warpage Simulation,” *2006 7th International Conference on Electronic Packaging Technology*, IEEE, pp. 1–7.

[44] Sadeghinia M., Jansen K., and Ernst L. J., 2012,“Characterization and modeling the thermo-mechanical cure-dependent properties of epoxy molding compound,” International Journal of Adhesion and Adhesives, **32**, pp. 82–88.

[45] Lee D., and Kim J.-K., 1998,“Measurement of residual stresses in encapsulant of a PQFP,” *Proceedings of 2nd Electronics Packaging Technology Conference (Cat. No.98EX235)*, IEEE, pp. 286–290.

[46] Fan H. B., and Yuen M. M., 2010,“A multi-scale approach for investigation of interfacial delamination in electronic packages,” Microelectronics Reliability, **50**(7), pp. 893–899.

[47] Su C.-T., Chiang T.-L., and Chiao K., 2005,“Optimizing the IC delamination quality via six-sigma approach,” IEEE Trans. Electron. Packag. Manufact., **28**(3), pp. 241–248.

[48] van Driel W. D., Zhang G. Q., Janssen J., Ernst L. J., Su F., Chian K. S., and Yi S., 2003,"Prediction and verification of process induced warpage of electronic packages," Microelectronics Reliability, **43**(5), pp. 765–774.

[49] Ichimura S., Kinashi K., and Urano T., 1990,"Relation between inner voids of plastic IC packages and non-Newtonian flow characteristics of resin encapsulant," *40th Conference Proceedings on Electronic Components and Technology*, IEEE, pp. 641–645.

[50] Nguyen L. T., and Lim F. J., 1990,"Wire sweep during molding of integrated circuits," *40th Conference Proceedings on Electronic Components and Technology*, IEEE, pp. 777–785.

[51] Wang L., Zhao Z., Wang J., Wen L., Wang Q., and Lee J., 2009,"Investigation on TSOP warpage mechanism and improvement method," *2009 16th IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits*, IEEE, pp. 747–750.

[52] Kelly G., Lyden C., Lawton W., Barrett J., Saboui A., Pape H., and Peters H., 1995,"The importance of molding compound chemical shrinkage in the stress and warpage analysis of PQFPs," *1995 Proceedings. 45th Electronic Components and Technology Conference*, IEEE, pp. 977–981.

[53] Vreugd J. de, Jansen K., Ernst L. J., Bohm C., Kessler A., and Preu H., 2008,"Effects of Molding Compound Cure on Warpage of Electronic Packages," *2008 10th Electronics Packaging Technology Conference*, IEEE, pp. 675–682.

[54] Yang D. G., Jansen K., Ernst L. J., Zhang G. Q., van Driel W. D., Bressers H., and Fan X. J., 2004,"Prediction of process-induced warpage of IC packages encapsulated with thermosetting polymers," *2004 Proceedings. 54th Electronic Components and Technology Conference (IEEE Cat. No.04CH37546)*, IEEE, 98-105 Vol.1.

[55] Nakazawa T., and Samara A., 2014,"Three-dimensional inline inspection for substrate warpage and ball grid array coplanarity using stereo vision," Applied optics, **53**(14), pp. 3101–3109.

[56] Tan W., Cheng C., Liu H., Duan Y., Liu L., Cheng X., and Li L., 2020,"Warpage Control Method in Epoxy Molding Compound," *2020 China Semiconductor Technology International Conference (CSTIC)*, IEEE, pp. 1–3.

[57] Prime R. B., 1981, Thermosets, *Thermal Characterization of Polymeric Materials*, Elsevier, pp. 435–569.

[58] Pascault J.-P., Sautereau H., Verdu J., and Williams R. J. J., 2002. *Thermosetting Polymers*, CRC Press.

[59] Ko M., and Kim M., 1998,"Effect of postmold curing on plastic IC package reliability," J. Appl. Polym. Sci., **69**(11), pp. 2187–2193.

[60] Chew S., 1996,"Thermal and viscoelastic characterization of transfer-molded epoxy encapsulant during simulated post-mold cure," *1996 Proceedings 46th Electronic Components and Technology Conference*, IEEE, pp. 1032–1038.

[61] Ng Y.-S., and Hooi W.-K., 2007,"FTIR Method to Distinguish Mold Compound Types and Identify the Presence of Post-Mold-Cure (PMC) Treatment on Electronic Packaging," *2007 9th Electronics Packaging Technology Conference*, IEEE, pp. 652–656.

[62] Carbas R. J. C., Marques E. A. S., Da Silva L. F. M., and Lopes A. M., 2014,"Effect of Cure Temperature on the Glass Transition Temperature and Mechanical Properties of Epoxy Adhesives," The Journal of Adhesion, **90**(1), pp. 104–119.

[63] Fraga I., Montserrat S., and Hutchinson J. M., 2008,"Vitrification during the isothermal cure of thermosets," J Therm Anal Calorim, **91**(3), pp. 687–695.

[64] Xu C., Zhong Z. W., and Choi W. K., 2017,"Epoxy molding compound effect on fan-out wafer level package strength during post-mold thermal process," *2017 16th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*, IEEE, pp. 1388–1392.

[65] Carbas R., Da Silva L., Marques E., and Lopes A. M., 2013,"Effect of post-cure on the glass transition temperature and mechanical properties of epoxy

adhesives," Journal of Adhesion Science and Technology, **27**(23), pp. 2542–2557.

[66] Kang M., and Jameson N. J., 2018, Machine Learning: Fundamentals, *Prognostics and Health Management of Electronics*, Pecht M. G., and Kang M., eds., John Wiley and Sons Ltd, Chichester, UK, pp. 85–109.

[67] Badillo S., Banfai B., Birzele F., Davydov I. I., Hutchinson L., Kam-Thong T., Siebourg-Polster J., Steiert B., and Zhang J. D., 2020,"An Introduction to Machine Learning," Clinical pharmacology and therapeutics, **107**(4), pp. 871–885.

[68] Ray S., 2019,"A Quick Review of Machine Learning Algorithms," *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, IEEE, pp. 35–39.

[69] Kononenko I., 2001,"Machine learning for medical diagnosis: history, state of the art and perspective," Artificial Intelligence in Medicine, **23**(1), pp. 89–109.

[70] Fujiyoshi H., Hirakawa T., and Yamashita T., 2019,"Deep learning-based image recognition for autonomous driving," IATSS Research, **43**(4), pp. 244–252.

[71] Peter H., Ghani N., and Kalegele K., 2017,"On the Development of Machine Learning – Based Application Framework for Enhancing Performance of Livestock Mobile Application Systems in Poor Internet Service Areas," IJCA, **179**(1), pp. 45–54.

[72] BISHOP C. M., 2006. *Pattern recognition and machine learning*, Springer, New York.

[73] Jain A., Patel H., Nagalapatti L., Gupta N., Mehta S., Guttula S., Mujumdar S., Afzal S., Sharma Mittal R., and Munigala V., 2020,"Overview and Importance of Data Quality for Machine Learning Tasks," *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, ACM, New York, NY, USA, pp. 3561–3562.

[74] Mohammed J., and Mahmud M. J., 2020,"Selection of a machine learning algorithm for OSHA fatalities," *2020 IEEE Technology & Engineering Management Conference (TEMSCON)*, IEEE, pp. 1–5.

[75] S. B. Kotsiantis, D. Kanellopoulos, and P. E. Pintelas, 2007,"Data Preprocessing For Supervised Leaning,"

[76] Barto A. G., 1997, Reinforcement Learning, *Neural Systems for Control*, Elsevier, pp. 7–30.

[77] Foster J. M., and Jones M., 2017,"Reinforcement Learning with Analogical Similarity to Guide Schema Induction and Attention,"

[78] Nilsson N. J., 1998. *Introduction to Machine Learning*, Stanford, CA.

[79] Gao K., Mei G., Piccialli F., Cuomo S., Tu J., and Huo Z., 2020,"Julia language in machine learning: Algorithms, applications, and open issues," Computer Science Review, **37**, p. 100254.

[80] Hady M. F. A., and Schwenker F., 2013, Semi-supervised Learning, *Handbook on Neural Information Processing*, Bianchini M., Maggini M., and Jain L. C., eds., Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 215–239.

[81] Ripley B. D., 2014, Unsupervised Methods, *Pattern Recognition and Neural Networks*, Ripley B. D., ed., Cambridge University Press, pp. 287–326.

[82] Andrea Beck, and Marius Kurz, 2020,"A Perspective on Machine Learning Methods in Turbulence Modelling,"

[83] Kubat M., 2021. *An Intriduction to Machine Learning: Unsupervised learning*, Springer Nature Switzerland.

[84] Zimek A., Schubert E., and Kriegel H.-P., 2012,"A survey on unsupervised outlier detection in high-dimensional numerical data," Statistical Analysis, **5**(5), pp. 363–387.

[85] Ezzat A., Wu M., Li X.-L., and Kwoh C.-K., 2017,"Drug-target interaction prediction using ensemble learning and dimensionality reduction," Methods (San Diego, Calif.), **129**, pp. 81–88.

[86] Tomašev N., and Radovanović M., 2016, Clustering Evaluation in High-Dimensional Data, *Unsupervised Learning Algorithms*, Celebi M. E., and Aydin K., eds., Springer International Publishing, Cham, pp. 71–107.

[87] Xu Y., and Goodacre R., 2018,“On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning,” Journal of analysis and testing, **2**(3), pp. 249–262.

[88] Hagan M. T., Demuth H. B., Beale M. H., and Jesús O. de, 2016. *Neural network design,* 2nd ed.

[89] Joseph V. R., 2022,“Optimal ratio for data splitting,” Statistical Analysis, **15**(4), pp. 531–538.

[90] Westerhuis J. A., Hoefsloot H. C. J., Smit S., Vis D. J., Smilde A. K., van Velzen E. J. J., van Duijnhoven J. P. M., and van Dorsten F. A., 2008,“Assessment of PLSDA cross validation,” Metabolomics, **4**(1), pp. 81–89.

[91] Rosipal R., and Krämer N., 2006, Overview and Recent Advances in Partial Least Squares, *Subspace, Latent Structure and Feature Selection*, Saunders C., Grobelnik M., Gunn S., and Shawe-Taylor J., eds., Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 34–51.

[92] Abdi H., 2003. *Partial Least Squares (PLS) Regression*.

[93] Carrascal L. M., Galván I., and Gordo O., 2009,“Partial least squares regression as an alternative to current regression methods used in ecology,” Oikos, **118**(5), pp. 681–690.

[94] Wold S., Sjöström M., and Eriksson L., 2001,“PLS-regression: a basic tool of chemometrics,” Chemometrics and Intelligent Laboratory Systems, **58**(2), pp. 109–130.

[95] Uhrig R. E., 1995,“Introduction to artificial neural networks,” *Proceedings of IECON '95 - 21st Annual Conference on IEEE Industrial Electronics*, IEEE, pp. 33–37.

[96] McCulloch W., and Pitts W., 1944,“A logical calculus of the ideas immanent in nervous activity. Bulletin of mathematical biophysics, vol. 5 (1943), pp. 115–133,” J. symb. log., **9**(2), pp. 49–50.

[97] Meireles M., Almeida P., and Simoes M. G., 2003,“A comprehensive review for industrial applicability of artificial neural networks,” IEEE Trans. Ind. Electron., **50**(3), pp. 585–601.

[98] Bishop C. M., 1994,“Neural networks and their applications,” Review of Scientific Instruments, **65**(6), pp. 1803–1832.

[99] Brinksmeier E., Tönshoff. H. K., Czenkusch C., and Heinzel C., 1998,“Modelling and optimization of grinding processes,” Journal of Intelligent Manufacturing, **9**(4), pp. 303–314.

[100] Bai H., Kwong C. K., and Tsim Y. C., 2006,“Process modelling and optimisation using artificial neural networks and gradient search method,” Int J Adv Manuf Technol, **31**(7-8), pp. 790–796.

[101] Rajendra P., Sudheer K., and Boadh R., 2017,“Design of a Recognition System Automatic Vehicle License Plate through a Convolution Neural Network,” IJCA, **177**(3), pp. 47–54.

[102] Feng J., He X., Teng Q., Ren C., Chen H., and Li Y., 2019,“Reconstruction of porous media from extremely limited information using conditional generative adversarial networks,” Physical review. E, **100**(3-1), p. 33308.

[103] Ioffe S., and Szegedy C., 2015,“Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,”

[104] Abraham A., 2005, Artificial Neural Networks, *Handbook of Measuring System Design*, Sydenham P. H., and Thorn R., eds., John Wiley & Sons, Ltd, Chichester, UK.

[105] Barzilai J., 2020, On neural-network training algorithms, *Human-Machine Shared Contexts*, Elsevier, pp. 307–313.

[106] Günther F., and Fritsch S., 2010,“neuralnet: Training of Neural Networks,” The R Journal, **2**(1), p. 30.

[107] Erb R. J., 1993,“Introduction to backpropagation neural network computation,” Pharmaceutical research, **10**(2), pp. 165–170.

[108] Chai T., and Draxler R. R., 2014. *Root mean square error (RMSE) or mean absolute error (MAE)*.

[109] Goh A., 1995,“Back-propagation neural networks for modeling complex systems,” Artificial Intelligence in Engineering, **9**(3), pp. 143–151.

[110] Rojas R., 1996, The Backpropagation Algorithm, *Neural Networks*, Rojas R., ed., Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 149–182.

[111] Roşu D., Caşcaval C., Mustąţă F., and Ciobanu C., 2002,“Cure kinetics of epoxy resins studied by non-isothermal DSC data,” Thermochimica Acta, **383**(1-2), pp. 119–127.

[112] Montenegro R., 2003,“Crystallization, Biomimetics and Semiconducting Polymers in Confined Systems,” Dissertation, Mathematisch-Naturwissenschaftlichen Fakultät der Universität Potsdam, Universität Potsdam, Potsdam, Germany.

[113] ASTM,“Standard Terminology Relating to Thermal Analysis and Rheology,” ASTM E437-22d.

[114] Menczel J. D., Judovits L., Prime R. B., Bair H. E., Reading M., and Swier S., 2009, Differential Scanning Calorimetry (DSC), *Thermal Analysis of Polymers*, Menczel J. D., and Prime R. B., eds., John Wiley & Sons, Inc, Hoboken, NJ, USA, pp. 7–239.

[115] ASTM,“Standard Test Method for Assignment of the Glass Transition Temperatures by Differential Scanning Calorimetry,” ASTM E1356-08.

[116] James J., 2017, Thermomechanical Analysis and Its Applications, *Thermal and Rheological Measurement Techniques for Nanomaterials Characterization*, Elsevier, pp. 159–171.

[117] ASTM,“Standard Test Method for Linear Thermal Expansion of Solid Materials by Thermomechanical Analysis,” E831-19.

[118] Schaaf U., and Kugler A.,“Vorrichtung und Verfahren zur Herstellung von mit einer Gießmasse zumindest bereichsweise überdeckten Bauelementen,” WO 2019/057388 A1.

[119] Glassford S. E., Byrne B., and Kazarian S. G., 2013,“Recent applications of ATR FTIR spectroscopy and imaging to proteins,” Biochimica et biophysica acta, **1834**(12), pp. 2849–2858.

[120] Dutta A., 2017, Fourier Transform Infrared Spectroscopy, *Spectroscopic Methods for Nanomaterials Characterization*, Elsevier, pp. 73–93.

[121] Ismail A. A., van de Voort F. R., and Sedman J., 1997, Chapter 4 Fourier transform infrared spectroscopy: Principles and applications, *Instrumental Methods in Food Analysis*, Elsevier, pp. 93–139.

[122] Abidi N., Ed., 2021. *FTIR Microspectroscopy*, Springer International Publishing, Cham.

[123] Mohamed M. A., Jaafar J., Ismail A. F., Othman M., and Rahman M. A., 2017, Fourier Transform Infrared (FTIR) Spectroscopy, *Membrane Characterization*, Elsevier, pp. 3–29.

[124] Baker D. J., 2011,“Surface Modification and Characterization of Zinc Oxide for Use in Organic Photovoltaic Devices,”

[125] Grdadolnik, and Joze, 2002,“ATR-FTIR Spectroscopy: Its advantages and limitations,” Acta Chimica Slovenica, **49**(09), pp. 631–642.

[126] Niegisch C., Haag S., Braun T., Holck O., and Schneider-Ramelow M., 2022,“In-situ infrared spectroscopy for chemical analysis in electronic packaging processes,” *2022 IEEE 9th Electronics System-Integration Technology Conference (ESTC)*, IEEE, pp. 336–342.

[127] Bakhtazad A., and Chowdhury S., 2019,“An evaluation of optical profilometry techniques for CMUT characterization,” Microsyst Technol, **25**(9), pp. 3627–3642.

[128] Saxena M., Eluru G., and Gorthi S. S., 2015,“Structured illumination microscopy,” Adv. Opt. Photon., **7**(2), p. 241.

A1 Methodology for Updating Efficiency Terms

This appendix explains the gradient descent method used for updating the iterated efficiency terms during post-mold cure parameter optimization, as a supplementary document to Chapter 3.4.4. **Figure A1.1** shows the extended flow chart for the machine learning model that predicts the quality terms and optimizes the post-mold cure parameters accordingly. As seen, the reference values $T_{0\,\mathrm{itr.}}$ and $t_{0\,\mathrm{itr.}}$ for iteration start points, optimization weights w_δ, w_h, w_T, w_t and iteration parameters are fed to the model by the user, along with the measured initial and desired values for warpage and reaction enthalpy.

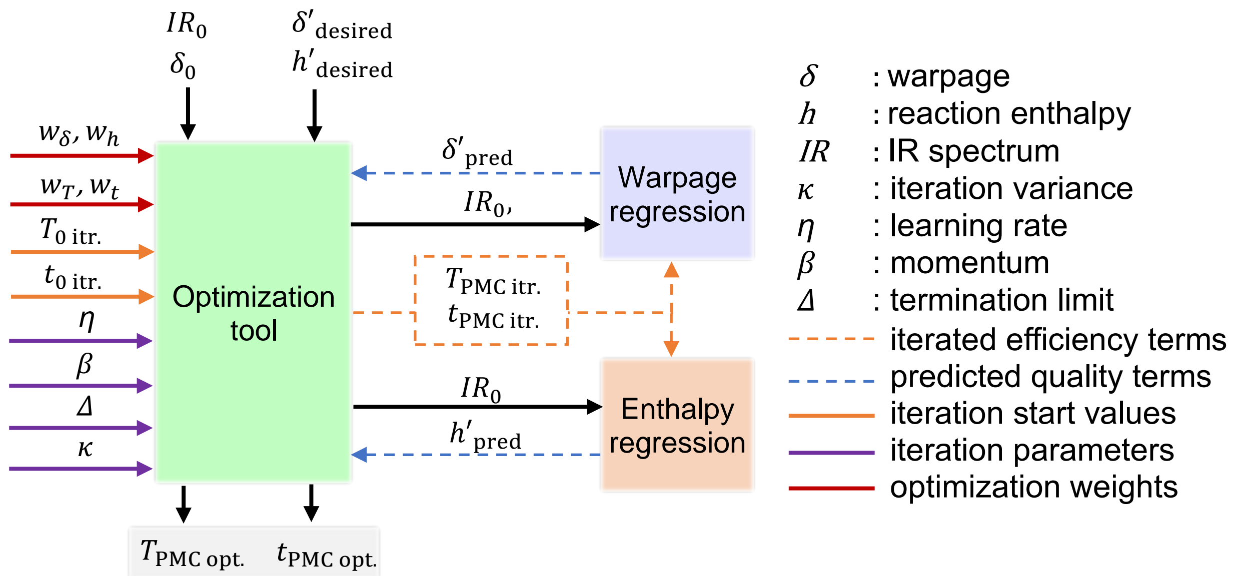

Figure A1.1: Extended flowchart for optimization tool

Figure A1.2 illustrates the steps followed during updating the iterated efficiency terms. As the first step, the final warpage and final reaction enthalpy predictions δ'_{pred} and h'_{pred} are made for the defined iteration start parameters $T_{0\,\mathrm{itr.}}$ and $t_{0\,\mathrm{itr.}}$, and the reference cost for that point is calculated using Equation (3.2). Following that, if the predicted values for warpage and reaction enthalpy are not in accordance with the desired values, an artificial experiment set is created via variating the parameters with the user defined value κ as shown in Step 2. Then, using IR_0 and δ_0, and also the artificial PMC parameters, the resultant δ' and h' values are predicted, followed by the calculation of costs for each one of these points. As next, in Step 5, a simple ANN is built that fits a relationship in between the calculated costs and the variated PMC parameters, with the goal of detecting the

changes in cost as a response to variations in the PMC parameters. After training this dummy network, the assigned weights indicate how strongly and in which direction these iterated PMC parameters affect the resultant cost. Having the assigned weights at Step 6 the iterated PMC parameters are updated in the direction of decreasing costs using the gradient descent method, which is introduced in Chapter 2.2.4 with Equation (2.10). The predictions for δ' and h' are repeated, and the new reference cost is calculated for the updated parameters. As the termination condition, the absolute difference of reference costs are compared with the termination limit Δ, which is defined as a value that indicates a stagnation in updated iterations. If the difference is below this limit, it is assumed that the predictions are converged to a value and the loop is terminated. If not, the loop is started again from Step 2, with the updated reference points.

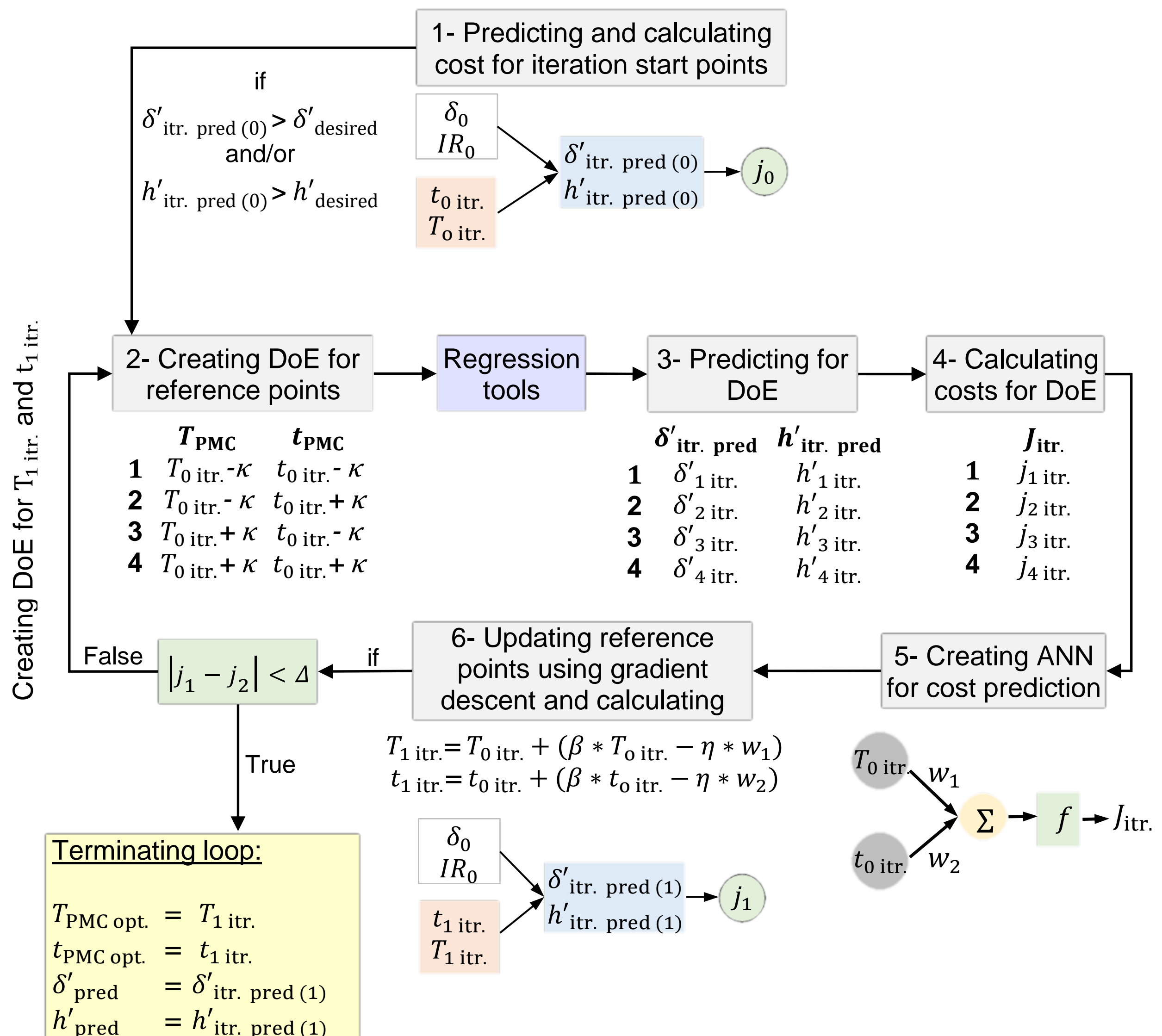

Figure A1.2: Flowchart for updating iterated efficiency terms