DISSERTATION

SYSTEM UNDERSTANDING OF HIGH PRESSURE DIE CASTING PROCESS AND DATA WITH MACHINE LEARNING APPLICATIONS

Submitted by

David J. Blondheim, Jr.

Department of Systems Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2021

Doctoral Committee:

Advisor: Charles Anderson

Steve Simske
Donald Radford
Michael Kirby

ABSTRACT

SYSTEM UNDERSTANDING OF HIGH PRESSURE DIE CASTING PROCESS AND DATA WITH MACHINE LEARNING APPLICATIONS

Die casting is a highly complex manufacturing system used to produce near net shape castings. Although the process has existed for more than hundred years, a systems engineering approach to define the process and the data die casting can generate each cycle has not been completed. Industry and academia have instead focused on a narrow scope of data deemed to be the critical parameters within die castings. With this narrow focus, most of the published research on machine learning within die casting has limited success and applicability in a production foundry. This work will investigate the die casting process from a systems engineering perspective and show meaningful ways of applying machine learning.

The die casting process meets the definition of a complex system both in technical definition and in the way that humans interact within the system. From the technical definition, the die casting system is a network structure that is adaptive and can self-organize. Die casting also has nonlinear components that make it dependent on initial conditions. An example of this complexity is seen in the stochastic nature of porosity formation, even when all key parameters are held constant. Die casting is also highly complex due to the human interactions. In manufacturing environments, human's complete visual inspection of castings to label quality results. Poor performance creates misclassification and data space overlap issues that further complicate supervised machine learning algorithms.

The best way to control a complex system is to create feedback within that system. For die casting, this feedback system will come from Industry 4.0 connections. A systems engineering approach will define the critical process and then create groups of data in a data framework. This data framework will show the data volume is several orders of magnitude larger than what is currently being used within the industry.

With an understanding of the complexity of die cast and a framework of available data, the challenge becomes identifying appropriate applications of machine learning in die casting. The argument is made, and four case studies show, unsupervised machine learning provides value by automatically monitoring the data that can be obtained and identifying anomalies within the die cast manufacturing system. This process control improvement thereby removes the noise from the system, allowing one to gain knowledge about the die casting process. In the end, the die casting industry can better understand and utilize the data it generates with machine learning.

ACKNOWLEDGEMENTS

Completing a PhD requires more than one's own fortitude (or perhaps stubbornness in my case), including the assistance of many people who came into my life even decades before I started this educational program. As many as I thank here, there are multiples more whom I do not mention by name, but appreciate their role.

First, I must thank my parents, David and Peggy Blondheim. They encouraged me in every endeavor I undertook, specifically within my education. Their requirement for me to fund my undergraduate degree is more appreciated than they know. I learned responsibility and how to work hard from my parents. For that, I am forever grateful. I hope I can teach the same to my children.

Next, I must thank Dr. Chuck Anderson for being my advisor on this PhD journey. Looking back now, I am not sure either one of us fully knew what we were getting into when a manufacturing "plant rat" approached a computer science professor with an idea to research machine learning applications in die casting. His guidance and patience was much appreciated and greatly helped me.

The rest of my PhD committee also needs to be thanked for their assistance throughout this process. Dr. Steve Simske, Dr. Donald Radford, and Dr. Michael Kirby have provided insight in classes and feedback on my work that helped shaped this final document. They have all dealt with my late-night emails and random questions during the past years.

I would have never attempted a PhD without the support and assistance from many people at Mercury Marine. First, I must thank a former Mercury employee, Mike Meyer, who hired me. Mike provided me much latitude that allowed me to combine my passions of manufacturing and data analysis and their applications within the foundry. Without his support, none of the other puzzle pieces would have connected for this PhD. Next, I need to thank Chris Drees for starting Mercury's Industrial Internet of Things (IIoT) initiative in 2017 and selecting me to lead it. It has provided me with a vehicle to accelerate the data collection within manufacturing. This thanks extends to John Buelow and his leadership and support in growing this IIoT initiative under his direction. He also provided me the independence to push the boundaries that I did. I must also thank Jerry Cegielski for the continued support in both my roles within foundry and IIoT/Connected Operations. Without his help, I would not have been able to focus on the critical projects that I have within Mercury.

This PhD would not have been possible without the assistance from four other employees at Mercury. I owe Clay Rasmussen many thanks for all his assistance with CAD and casting simulations. His skill with models created many of the figures found throughout the dissertation. A special thanks is also owed to Mark Potratz for assisting me with all the questions I had on the die casting process through the years. I will never comprehend the die casting process and equipment like him. I hope this work removes some of the "magic" associated with the industry that only people like Mark with more than 45 years of experience in die casting will ever truly know. I also must thank my technical counterparts and co-PhD students within the foundry: Alex Monroe and Adam Kopper. Their feedback and collaboration on papers, presentations, technical topics in both the foundry and machine learning, and patience with me is greatly appreciated. I would not have finished this degree without your assistance and competitive push.

Most importantly, I must thank my immediate family. This PhD has been a long journey that really has been a family event. From AJ playing Minecraft next to me while I was listening to classes at night, to the proofreading and editing Kerry has provided on all my writing, to needing to rock Teddy to sleep some nights as I ponder the topics written about here, this would not be possible without their support. This degree is as much theirs as it is mine.

Kerry, thank you for being my best friend. We will continue to overcome every challenge thrown at us. I love you so much. AJ and Teddy, the only person that can limit your potential is yourself. Never fall into that trap. You will both succeed at everything you do by hard work, dedication, and never being satisfied with good enough. Always continue to learn and challenge yourself to do better. I am so proud of both of you and love you more than you will know.

PREFACE

Data and machine learning is changing the world and improving our lives. With or without direct knowledge of its influence, it shapes what we purchase, read, and watch every day. Advancements in image recognition and natural language processing created amazing results. Deepfake videos make one question what we can really believe. Conversations are had with an electronic device sitting in the kitchen. With all these advances, applications of machine learning in manufacturing would seem to be a trivial challenge for the power of this technology. Yet, machine learning has yet to develop a strong foothold within industrial applications. Much research and academic work has shown benefits often with small, controlled studies. There are substantial challenges to overcome when applying this to complex production environments.

This work has combined my passions of the manufacturing world I have worked in for close to two decades with the use of data and machine learning. The goal is to gain an understanding of the process that has does not exist today. I want industry to recognize the complexity that die casting deserves to be classified as and then reduce that complexity to manageable parts to better learn from it. This dissertation provides a means to share this knowledge.

Each chapter of this dissertation can ideally be read by itself. Several of the chapters and case studies included are based on publications that have been completed through the years of my study. The organization of the chapters within the dissertation is a much more logical progression than the chronological order in which I studied and learned. It is said that education is not a straight line. The work that comprises this dissertation is additional proof of that.

Chapter 1 begins with an introduction to die casting and starts to touch on the complexity of the die casting process. A review of die casting process optimization literature is provided that will show gaps within the approaches used to date. Chapter 2 discusses the details of complexity theory within the die casting manufacturing system, the need for a systems engineering approach, and ultimately completes a data framework for the die casting system. As will be demonstrated, the volume and velocity of the data generated in die casting needs machine learning tools to help process and identify important information.

Chapters 3 and 4 build on the complexity of the die casting system in two different ways. Chapter 3 details a study performed on the stochastic nature of porosity formation in die castings. The work shows porosity formation was random even when variables that the industry typically controls to improve part quality were held constant. Chapter 4 discusses the human impact on classification of defects and the importance of the Critical Error Threshold (CET) of the casting process. The CET, combined with the high-quality performance of most manufacturing operations, focuses the use of machine learning to areas within die casting outside of the traditional supervised machine learning approach.

Chapter 5 entertains the topics of applications of machine learning within a production die casting process. A review of machine learning and some of the challenges associated with manufacturing are discussed prior to a review of four machine learning case studies completed at Mercury Marine as part of the PhD research. Chapter 6 concludes the dissertation with a discussion around conclusions learned and recommended areas of future studies that can be expanded from this work.

There is still much work to be done for machine learning in manufacturing to approach the same levels of adoption as seen in other industries. Manufacturing is a traditional industry and is slow to change. The complex nature of manufacturing systems makes it intrinsically harder to successfully apply machine learning with success. Non-traditional uses of machine learning will be required within the industry. Without a strong call to learn and adapt these technologies, machine learning will continue to be a dream or academic experiment in manufacturing. This work provides another voice calling for adaption and implementation of machine learning in manufacturing. The most important element to be gained is the additional insight and new knowledge about a process that has existed for more than 100 years.

-D. Blondheim, Jr.

DEDICATION

To my wife Kerry, for being my partner and best friend. Thanks for always being there for me.

To my son AJ, for re-teaching me the wondering and excitement of learning and questioning everything.

To my son Teddy, for showing me one can communicate volumes without ever speaking a word.

PS -Kerry, I promise I'm done with school now!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Die Casting Introduction

Chapter 1 describes the manufacturing process of die casting. Once the reader is introduced to the basic process, the complexity of the die casting system is explained. Next a review of die casting optimization literature is reviewed to highlight the deficiencies that exist in academic applications of machine learning to die casting process without a thorough understanding of the entire die casting system.

INTRODUCTION

High pressure die casting (HPDC or die casting) is a highly complex manufacturing process. Die casting is composed of multiple systems that control hydraulic, mechanical, and thermal processes to produce near net shape castings [1]. The design, setting, and control of these systems dictates casting quality and equipment performance. Historically, the die casting industry has collected and analyzed a fraction of the system data to control and optimize the process [2].

The North American Die Casting Association (NADCA) estimates annual sales of $8 billion for aluminum die casting in 2019. This represents more than 80% of the American Foundry Society's (AFS) forecast in all aluminum castings of $9.67 billion [3]. Current methods for control and optimization in the industry produce a median internal scrap rate of 8% of parts produced and equipment utilization of 68% [4]. Consider a foundry that has a world-class scrap rate of 2%. With a 60-second cycle time per casting, this manufacturer still produces 1.2 defective parts per production hour. Improving uptime and reducing scrap costs can create significant value within the die casting industry.

Die casting faces many challenges when it breaks from traditional data collection and moves to advanced analytics on large data sets. One of the first challenges is a fundamental understanding of how defects are formed and the classification of casting defects. The collection, processing, storage, and application of large volumes of data are all challenges. Additionally, the serialization and traceability of castings through the supply chain takes much effort. Finally, there is a cultural element to the use and acceptance of advanced analytical tools that must be overcome [5]–[9].

By applying a system engineering approach to die casting, the data generated will be orders of magnitude larger than what is traditionally collected. The industry would transition from dozens of data points saved each cycle to millions [2]. This level of data becomes overwhelming to anyone tasked with optimizing the die casting process. For the die casting system, a framework is needed for the data generation to understand the data required to solve industrial problems. The velocity and volume of data generated by die casting requires machine learning to understand, optimize, and control the die casting system.

DIE CASTING BASICS

High pressure die casting is a manufacturing process for producing near net shape metal castings, typically in large quantities. The process involves injecting liquid metal into a reusable mold at fast velocities and high pressures. Historically, die casting has been used for small to medium weight parts, typically under 20 pounds. Recent development of equipment and demand for large engine blocks [10] and structural die cast components [11] have made 40 pound or larger castings more commonplace in the industry. The die casting process can be broken down into high-level steps as seen in Table 1.

Table 1: High-level die cast process steps

	Step	Description
1	Spray Lubrication of the Die	With the die in an open position, the spray system applies die lube to the die surfaces. This die lube serves as a release agent for the aluminum on steel die and provides some thermal cooling of the die.
2	Closing of the Die	Through a mechanical or hydraulic process, the die halves close and lock under tonnage.
3	Metal Delivery	Metal is typically ladled by a robot or delivered with a metal pump into the chamber attached to the die.
4	Metal Injection	Using hydraulics, the liquid metal is pushed from the chamber into the die to form the casting.
5	Intensification Pressure	After the mold is filled, the injection system transfers to a high-pressure phase, used to squeeze the liquid metal into the mold during the solidification process.
6	Cooling/Dwell Time	Cooling systems are cycled to remove heat from the die, promoting solidification of the liquid metal to a solid casting.
7	Die Open	Mechanical/hydraulic system opens the two die halves, with the casting remaining on the moving part of the die.
8	Casting Ejection and Extraction	Ejection system within the die pushes the casting off the die surfaces to be extracted from the cell, typically by a robot.
9	Extraction Cell Casting Processing	Once out of the die, often additional validation and processing is performed on the casting prior to delivery to an operator for final inspection and packaging.

Each of these steps contains equipment beyond the die cast machine to perform specialized processes, such as metal holding, metal delivery, cooling, and extraction. These systems interface and communicate with the die cast machine to execute the steps required to produce a casting. Each die casting cycle ranges from a few seconds on small castings to two or three minutes on large castings. Cycle time is a function of the equipment size and speed, solidification time of the casting, and level of automation associated within the process. Figure 1 is an example layout of a large tonnage die casting machine installed and labeled with operational steps as seen from Table 1.

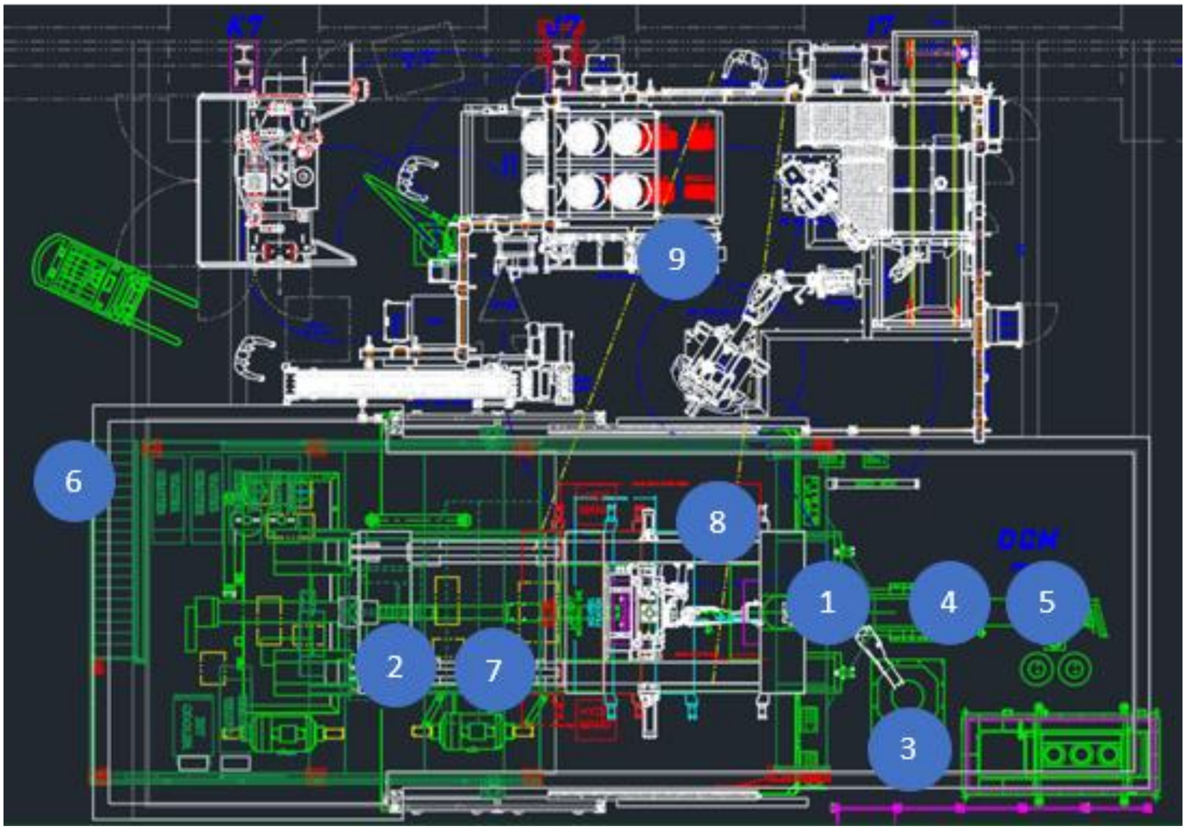


Figure 1: Example die cast cell layout
(photo permission from Mercury Marine)

Aluminum, magnesium, zinc, and zinc aluminum are the most commonly used alloys in die casting, although others including tin, brass, copper, and lead are also able to be used in certain applications [1], [12], [13]. Products from many different manufacturing industries use die castings, including automotive, agricultural machinery, recreation equipment, hand tools, home appliances, office furniture, electrical equipment, toys, aircraft, and home hardware [1], [12]. These industries utilize the advantages of die casting to create highly complex, near net shape castings. An example of a large aluminum die cast block produced by Mercury Marine can be seen in Figure 2.

Figure 2: Example V8 engine block die casting
(photo permission from Mercury Marine)

High pressure die casting provides many advantages to designers [1], [12], [13]. By using high velocities and large pressures, the die casting process can fill thin sections over long distances. This provides designers a means to a lightweight part without necessarily sacrificing strength or performance. Die casting creates some of the most complex casting designs with highly repeatable dimensional results due to the intricate steel die used to produce the casting. A wide variety of material and alloy choices are possible with die casting as previously discussed. These alloys fit many industrial needs in part design. Cycle time is another advantage as the mostly automated process can create castings much faster than any other foundry process such as sand casting, lost foam, or permanent mold. Volumes of thousands to millions are common. Other benefits include surface finish, pressure tightness, reduced machining needs, and low part cost.

As with any manufacturing process, there are trade-offs associated with using die casting. To achieve the quick cycle times and low part cost, companies must make large capital investments into equipment and tooling. This creates a high hurdle to justify designing for die casting [12]. Designs changes on a casting after the tooling is in production create expensive tool modifications and shortened tool life. In addition, die casting is a destructive process to the tooling used due to the thermal cycling. As a result, the costly dies require continued maintenance costs and eventually replacement costs [14]. Finally, designs must be created so that the casting can be ejected from the die steel. This requires additional draft on surfaces as well as prevents undercut features from being included [12], [14].

The benefits of die casting have continued to outweigh the challenges, which means die castings is a widely utilized foundry process for designing and producing components. With a foundational understanding of what die casting is, the focus shifts to the complexity of the die casting manufacturing system.

COMPLEXITY OF THE DIE CAST SYSTEM

Die casting is arguably one of the most complex manufacturing processes. It involves a phase change of metal, design geometry, hydraulics, thermal and heat transfer, die design and assembly, and hundreds of decisions to determine machine process settings [1], [14], [15]. A system engineering approach has not been published or used within the industry to understand the complexity of the process and the data it creates. This section introduces many of these systems that create this complex process with the topic of complexity further expanded in Chapter 2. These systems will be expanded in later chapters to develop a framework for die casting process and data collection.

METALLURGICAL, METAL HOLDING, AND METAL DELIVERY

A material phase change is a complex process with changes happening at an elemental level. Die casting involves alloys comprised of many different elements. Certain elements impact the solidification process differently. As an example, silicon is one of the key alloying elements in aluminum die cast alloys. Silicon (*Si*) promotes fluidity and castability of the liquid metal but also impacts the solidus temperature as it is increased [16]. As a result, a casting at a high *Si* level can create different filling and solidification processes versus one at a low level. *Si* is one of typically six to eight alloying elements that are specified in commercial aluminum die cast alloys [14]. Some elements can impact the process while others influence the mechanical properties of the material [16].

Once the alloy is created, it is held in a furnace until it is ladled into the shot system. The holding furnace is a sub-system designed to hold a designated volume of metal at a given temperature and interface with the ladle system. The furnace has a control system to maintain the required metal temperatures. A ceramic filter system is often used to help avoid oxides and other impurities in the liquid metal entering the casting. The ladling system moves the metal from the furnace to the machine for each cycle. Multiple options for metal delivery exist. Three options, as seen in Figure 3, include a dosing pump, a 2-axis metal delivery system, and a 7-axis robotic system. Additionally, a metal delivery system replenishes the metal in the holding furnace from a smelting area within the foundry. Degassing units could be installed into the furnace or in the metal replenishment system to help reduce metallurgical oxides and defects. Operators manually skim the holding furnace to help ensure clean metal is ladled into the die casting machine by physically removing the oxide film that is created at interface between the liquid metal and atmosphere.

Figure 3: Metal delivery: dosing furnace (left), 2-axis ladle (center), and 7-axis robot ladle (right) ***(photo permission from Mercury Marine)***

The ladle system has many different settings that can affect the casting process. The goal with delivering metal from the furnace to the chamber is to do it quickly to prevent temperature loss in the liquid metal but still to pour the liquid metal without turbulence to avoid trapping air in the liquid metal. Pour rates, pour angles, ladle start times, metal delivery cycle time, and ladle bucket size are some of the variables associated with the process that need to be monitored and controlled.

Metal holding and ladling are important sub-systems of the die casting process. Poor quality castings can be produced when these two systems impacted the cell's overall cycle time. Metal delivery is a discrete event. After each cycle, the level in the furnace drops down the volume of metal removed. Pending controls and timer settings, as metal is removed from the furnace, the ladle bucket can take longer to reach the liquid metal and fill the bucket. The more that was removed, the longer to fill. The longer cycle time of the ladle system could become the pacing item in the overall cell cycle time. The increased cycle time creates additional cooling in the die. This unplanned cooling could create a poor thermal condition in the tool, leading to defects within the casting.

CHAMBER AND INJECTION SYSTEM

Die casting involves a large hydraulic system used to inject the liquid metal from the chamber into the die casting mold to create the casting. This filling process is often modeled with simulation to ensure a uniform injection of the liquid metal through the die during the tool design process. A chamber holds the liquid metal delivered from the ladle to be injected into the die. This chamber length and diameter for the injection system are decided during the tool design process based on part volume and target metal pressure. A diagram of the chamber and shot rod system interfacing with the die is illustrated in Figure 4. The injection process happens in two phases [15].

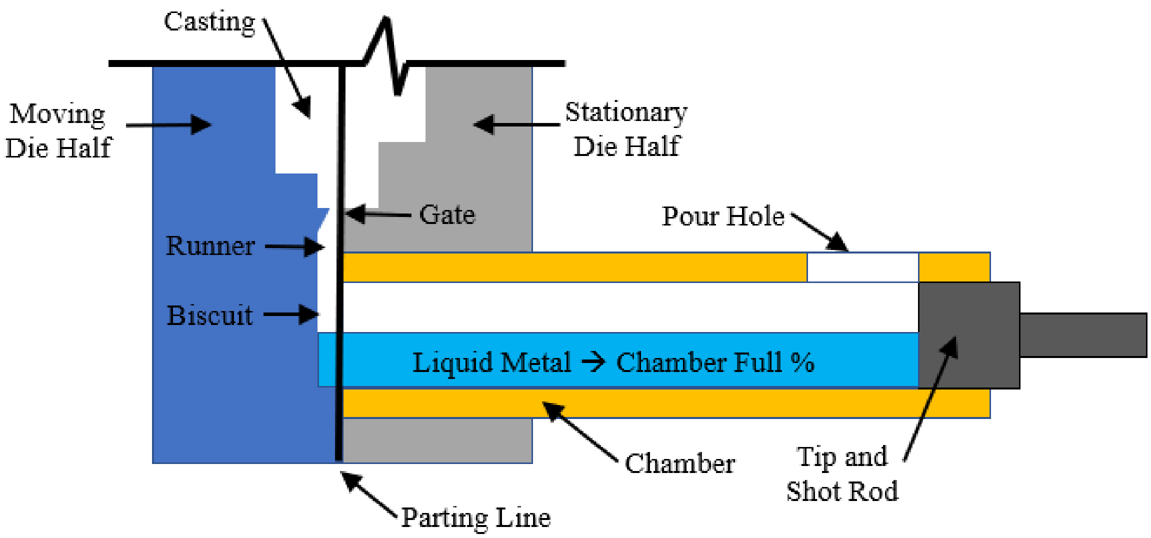


Figure 4: Die and chamber diagram

The first phase is typically called the *slow shot.* The slow shot involves moving the tip and shot rod forward to compress the partially filled chamber into a solid cylinder of liquid metal. The goal is to avoid entrapping air in the turbulence or waves in the liquid as it is injected into the casting. An acceleration in the slow shot speed is often used to reduce the likelihood of wave formation [15]. Figure 5, Figure 6, and Figure 7 are diagrams of wave formations with slow shot speeds that are respectively too slow, too fast, and correct. The second phase is typically called the *fast shot*. In the second phase, metal is pushed from the chamber into the die at an extremely fast rate. Fill times of less than 100 milliseconds are possible on castings larger than 40 pounds. The transition point between phase one and phase two is called the *start of fast shot*.

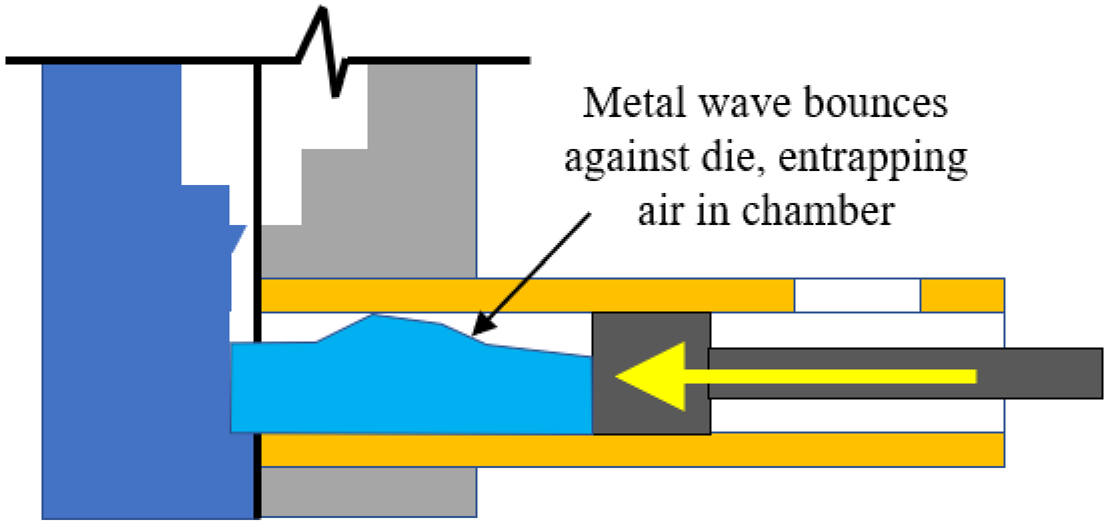


Figure 5: Waves entrap air when slow shot speed is too slow

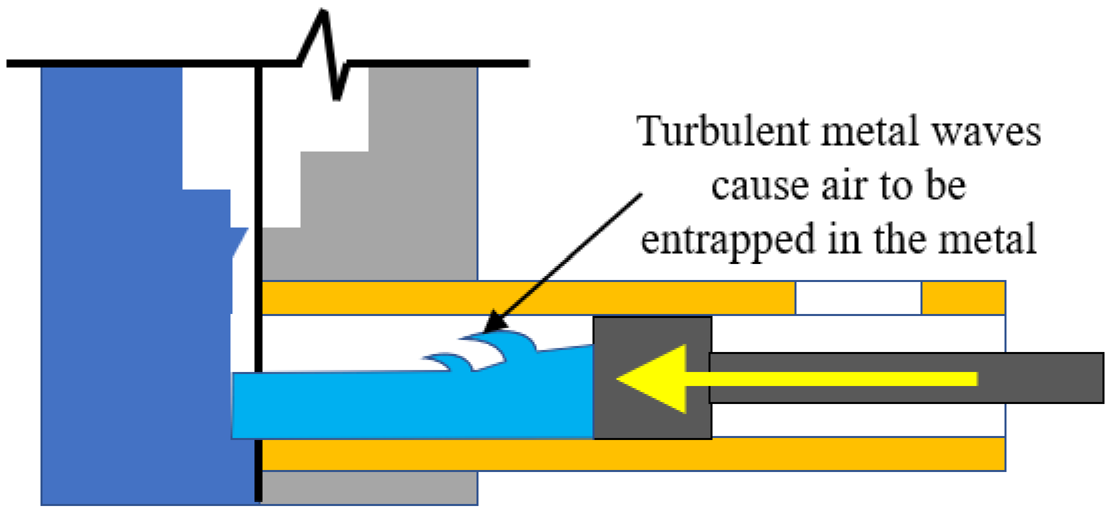


Figure 6: Turbulent waves form when slow shot speed is too fast

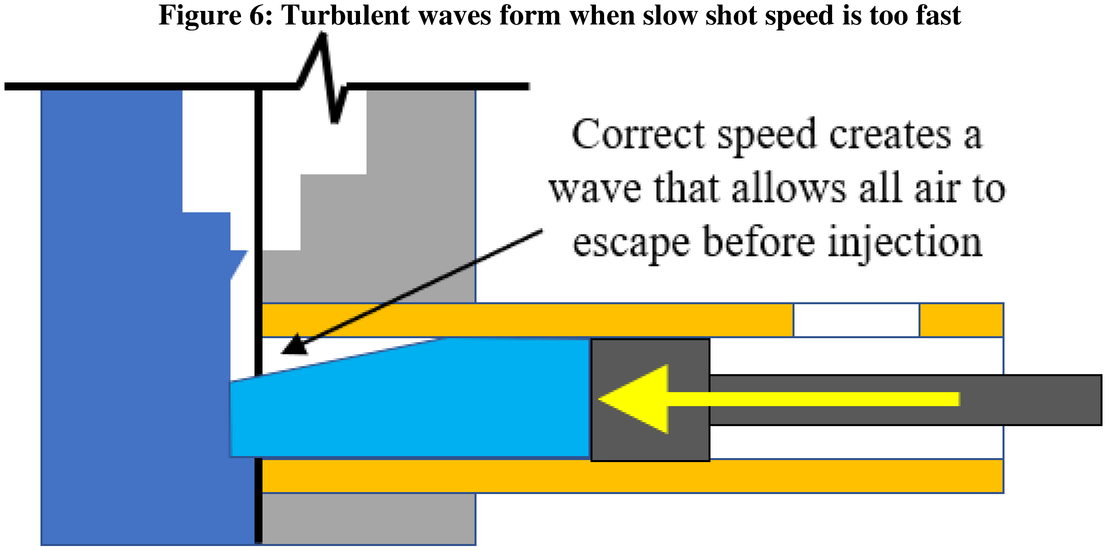


Figure 7: Correct wave formation allows all air to escape from the chamber

If properly designed, the tool will vent the air out of the mold and begin freezing the metal as the injection process is completed. The injection system then turns on an intensification process that pressurizes the remaining liquid metal in the chamber. The hydraulics of the system continue to push liquid metal into the die as the casting is solidifying. This process assists with filling shrink voids created during the phase change. All these speeds, transitions points, rates of increases, and pressures must be determined when designing the process and are subject to limitations of the equipment and the die design. There is an infinite number of options that can be selected with many of these combinations capable of making a good quality casting. The speed and pressures achieved with the injection process can be influenced by wear of the tip, cooling within the chamber, or failures within the hydraulic system, to name a few.   An example of the data typically generated during this injection and intensification process is seen in Figure 8.

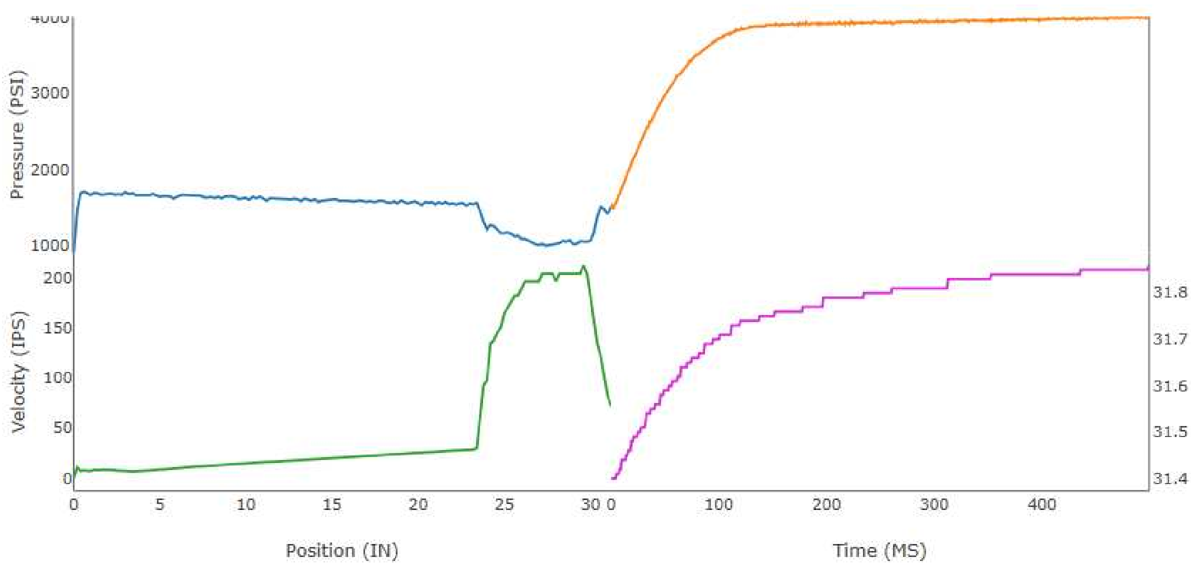


Figure 8: Example shot injection velocity and pressure graph

<u>DIE CLAMPING, SLIDE PULLS, AND EJECTION SYSTEM</u>

Hydraulics and mechanics clamp the die halves together to prevent them from separating during the injection process. Hydraulics are also used during the ejection process of the casting from the die. Decisions made during the process design can influence casting quality. For example, one decision is the clamping tonnage used to hold the die halves together. Machine capabilities and die design restrict the clamping tonnage. Four tie bars hold together the stationary and moving plates of the die cast machine. If the projected area of the casting as designed in the tool is not equally divided among the tie bars, the die may struggle to keep closed because more force is on one tie bar than the other. The die temperature increases as it absorbs the heat from repeat cycles. This growth in the tool steel can also impart more clamping force on certain parts of the die than others.

A dwell time is selected to allow the casting to solidify in the tool. A casting could explode because the walls are not strong enough to hold the high-pressure liquid metal that has yet to solidify in the casting if the dwell time is too short. Consequently, dwell time that is too long could have the casting solder or bind onto the die steel.

The machine opens the die and starts the ejection process once the dwell time is completed. The first step is to pull any die slides that formed the geometry of the castings. Multiple slides are possible in complex part designs. These slides are controlled with hydraulic cylinders. An example of a die with slides is seen in Figure 9. The order to pull the slides and the timing associated between slides must be calculated carefully. If too long, cooling of the casting may prevent the slide from pulling from the casting surface. However, pulling all the slides at the same time is not an option since the hydraulic power of the system is limited by the pump's capacity.

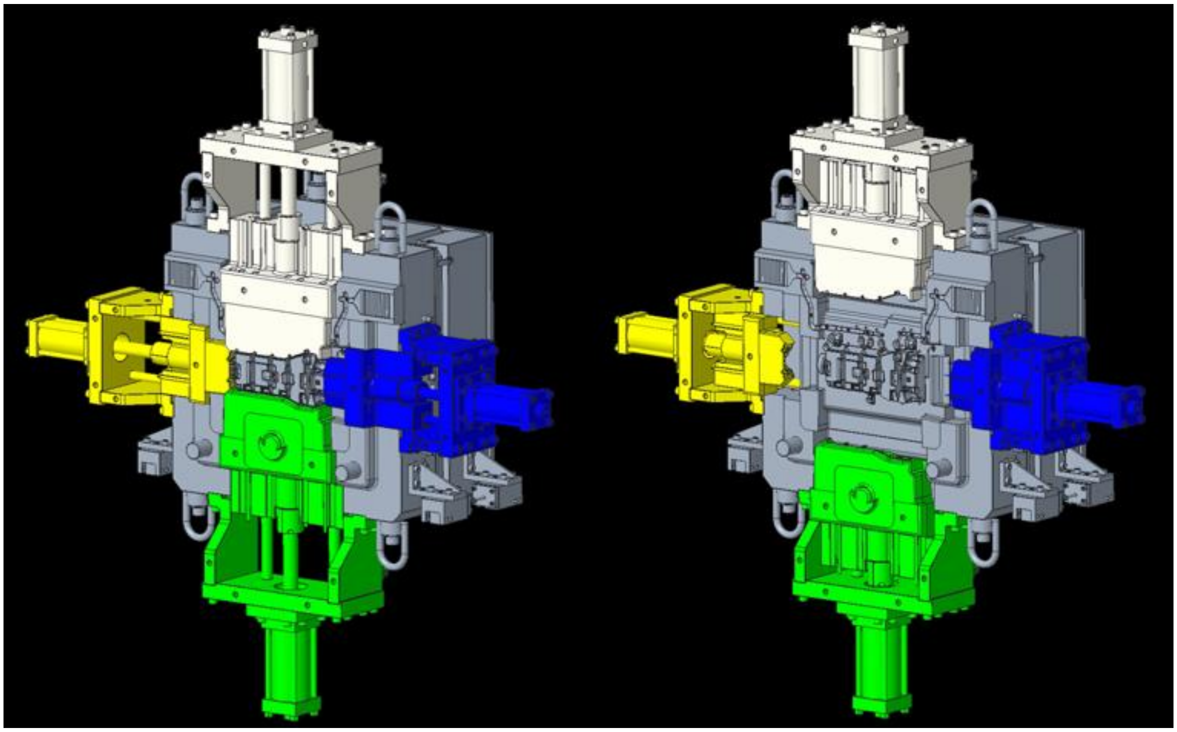


Figure 9: Die shown with hydraulic cylinders and slides in (left) and out (right) positions
(photo permission from Mercury Marine, image credit Clay Rasmussen)

Next, the ejection system in the die is hydraulically actuated. This ejection system is built into the die and includes a series of pins that push the casting out of the mold. A robot often removes the extracted casting from the cell. A hydraulic pressure must be set within the system. The die must be designed with an appropriate number of pins to overcome the force of the casting shrinking onto the die steel. An example of a die with the ejector pins in the out position is seen in Figure 10. Monitoring the pressure profile in the cylinders for ejection and slide pulls provides insight into potential tool failures or process changes that may have happened in the system.

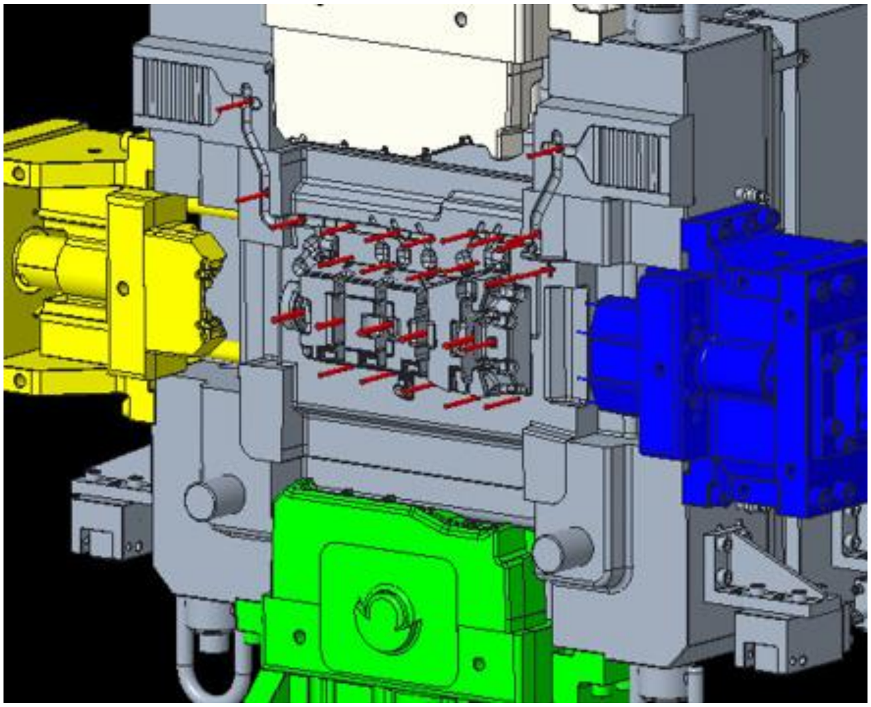


Figure 10: Die example with ejection pins in the out position
(photo permission from Mercury Marine, image credit Clay Rasmussen)

THERMAL MANAGEMENT OF TOOLING

Thermal management of the die steel is critical to casting quality and production rates. Heat is transferred from the liquid metal to the die during the solidification process. Hot water, hot oil, cold water, and high-pressure jet cooling are used to control the die temperature with passages designed into the die. The equipment used are individual sub-systems of the overall thermal management. The equipment interfaces with the die casting machine and receives on/off signals to provide thermal management at certain times in the process. Decisions must be made for temperature settings, when to start, and when to stop each of these systems. In a complex casting system, there could be a dozen thermal units, each with multiple control zones, supplying cooling media to hundreds of unique cooling passages within the die. An example of a hot water unit and a jet cool unit can be seen in Figure 11.

Figure 11: Hot water unit (left) and jet cool unit (right) in production
(photo permission from Mercury Marine)

Die spray is another method of controlling temperature in the die. Unlike the cooling lines, die spray manages temperature externally on the surface of the die. The main objective is to impart a lubricant on the surface of the die to allow for easier release of the casting during ejection. The ratio of die lube to water is variable based targeted rates and the mixing and distribution system within the facility. A spray manifold on specialized spray equipment or a multi-axis robot applies the die lube within the die casting cell.

Two approaches are often used for die spray manifolds. The first is a generic spray manifold that has only a few nozzles that typically delivery high volumes of spray and can be used on many different tools. The second approach utilizes a spray manifold with hundreds of nozzles designed specifically for a tool to pinpoint spray in critical areas. Each approach has its benefits and drawbacks in both proper lube application and long-term maintenance and performance. Examples of a 2-axis spray system and a 6-axis robot system can be seen in Figure 12. The length of time for each nozzle, nozzle direction, nozzle blockage, and movement of spray manifold all impact the amount of spray applied to the die. Flow rates and temperature of the lube also change the amount of cooling that occurs. Tracking these variables can help identify issues in the spray system and when process changes happen. Thermal couples imbedded within the die close to the cavity surface or the use of thermal image cameras on the die surface can capture die temperatures.

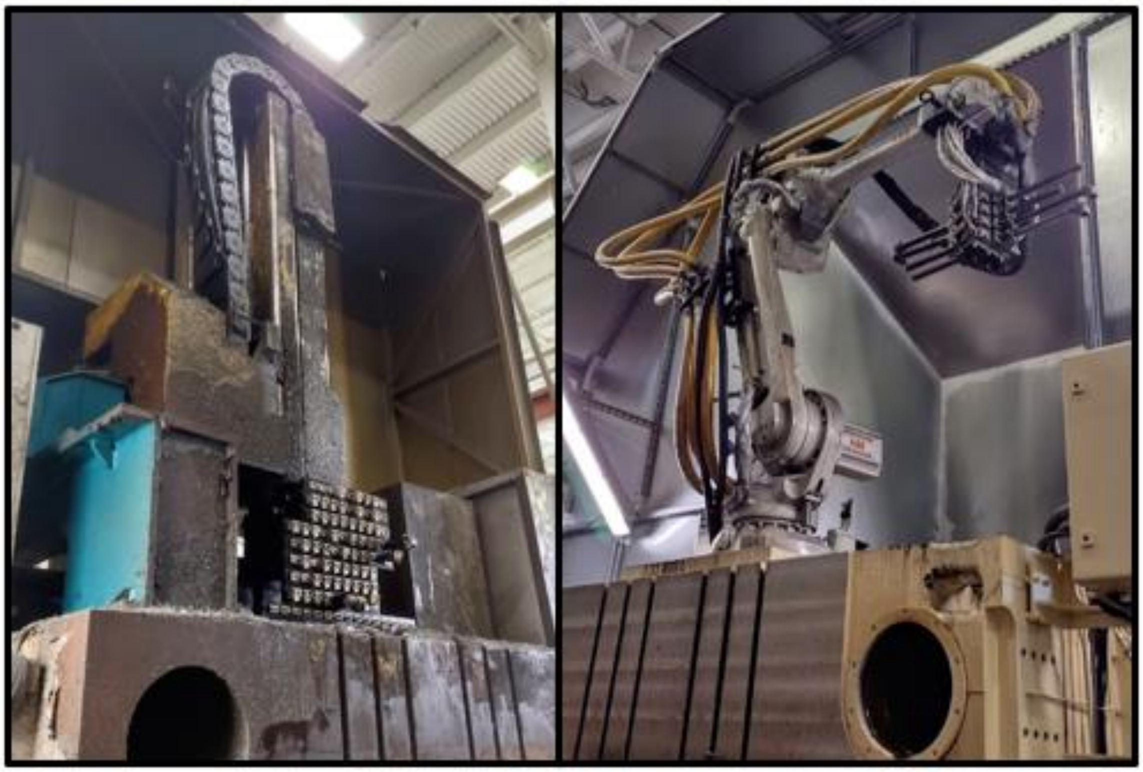


Figure 12: Spray systems: 2-axis manifold (left) and 6-axis robot (right)
(photo permission from Mercury Marine)

TOOL DESIGN, ASSEMBLY, AND SETUP

The design and intricacy of the tooling used in die casting is a large portion of the complexity in the process. A large tonnage die used for a V-block engine (as previously seen in Figure 2) can have more than 5,000 unique components required for assembly. These components must be properly designed, manufactured, and assembled into the die cast equipment for each run. Components can range from small fittings to holder block steel that weighs tons. The smallest fitting can play as important of a role as the largest piece of steel in producing a quality casting. The design of the die must balance metal flow and thermal control to produce a quality casting.

The flow of metal through the die is dictated by the injection settings, the gating system, the part design, and the venting system. Part design and geometry drives dictates many of the flow decisions, including part orientation within the die, areas that can be gates, and the last fill area of the casting. In turn, the designer must plan how the metal enters the die and how the air from the empty cavity leaves. For the liquid metal, runners are designed to transition the metal from the chamber to the casting. The metal enters the castings at the gates. Once inside, it pushes the air from the cavity out of the casting mold, through the venting system. The metal fills the casting and then flows into the overflows, which are designed to capture the initial metal front as its pushed through the mold. The metal slows in the venting system, and then is designed to solidify in the vents, which ideally has allowed all the air out but none of the liquid metal. These features can be seen in die model shown in Figure 13 and in an actual casting seen in Figure 14.

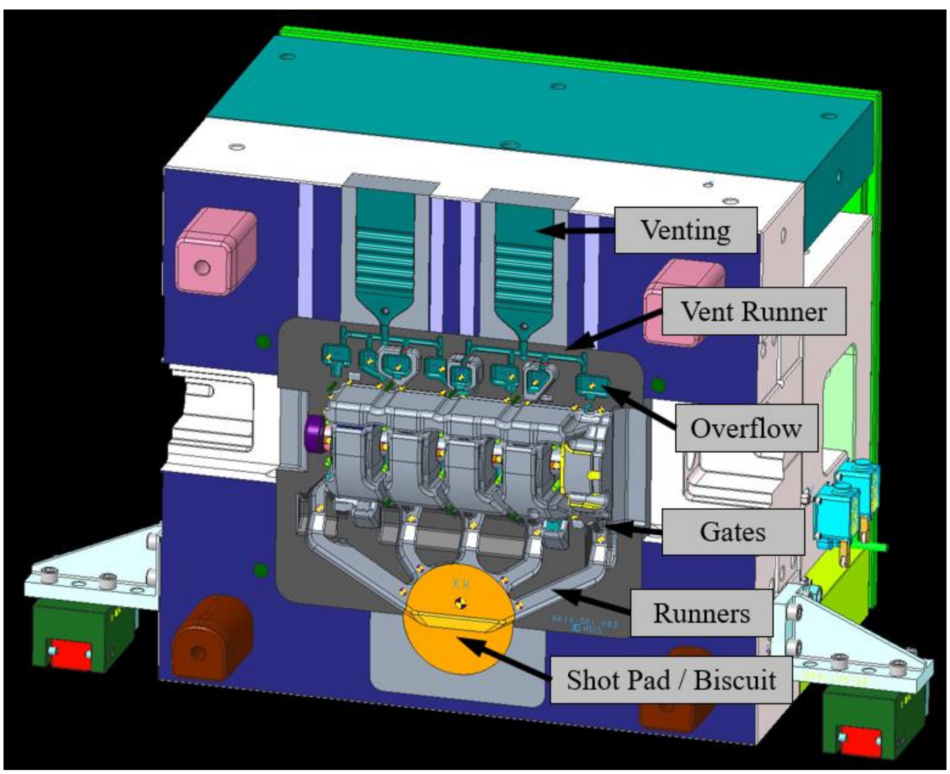


Figure 13: Key nomenclature of features within the die
(photo permission from Mercury Marine, image credit Clay Rasmussen)

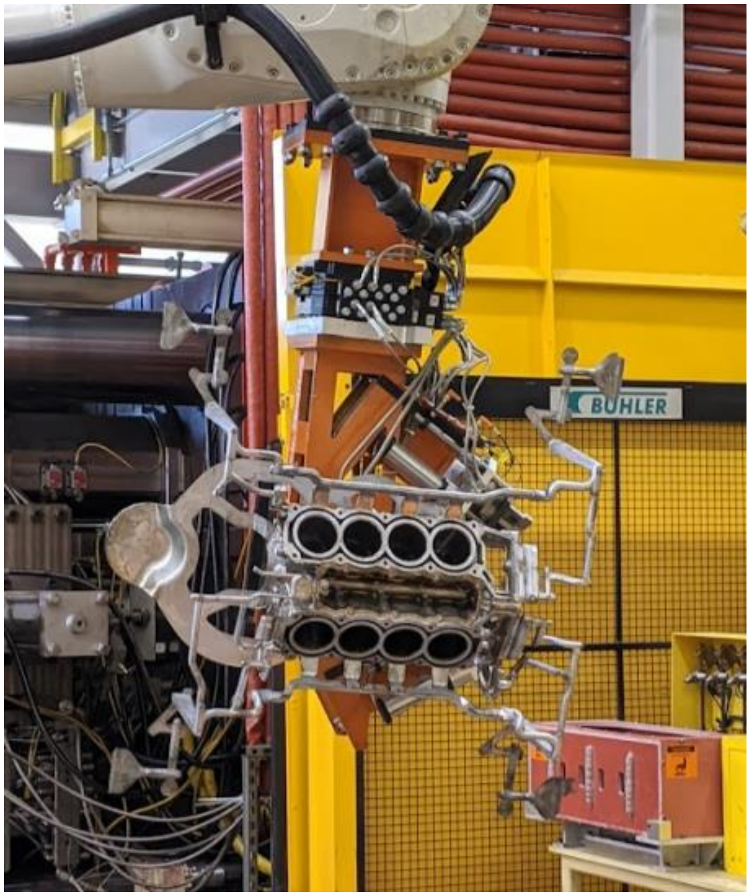


Figure 14: Biscuity, runner, gating, and venting system on casting
(photo permission from Mercury Marine)

Flow of air and metal through the casting is a complex problem. The die casting industry utilizes tools such as flow simulation to help design tooling and gain an understanding of how different gating and venting systems impact the filling process of the casting. Multiple iterations of simulations are completed to optimize the design. An example of a fluid flow simulation is seen in Figure 15. Metal is injected from the chamber through the biscuit and runner system into the casting. This injection process typically takes 50 to 100 milliseconds during the injection process.

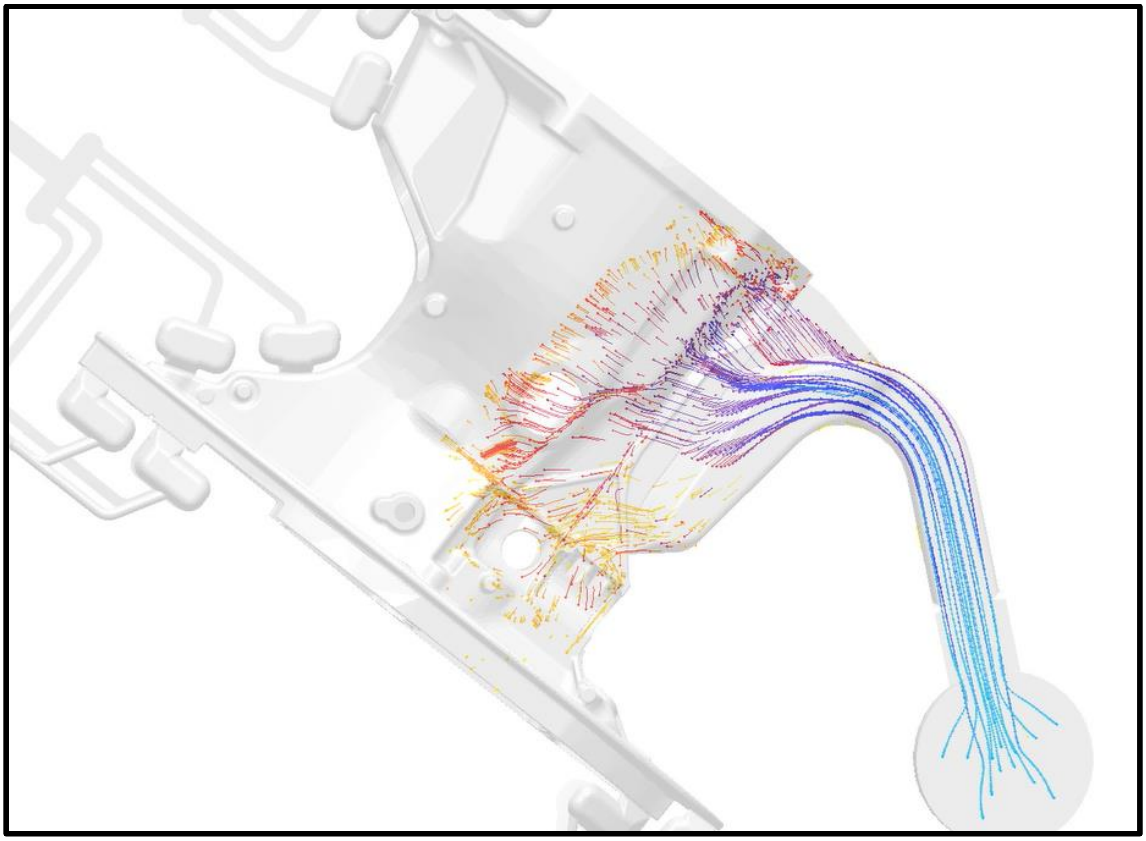


Figure 15: Simulated metal fluid flow during injection
(photo permission from Mercury Marine)

One of the critical features for design is the thermal management of the die temperature. The die casting industry uses different cooling systems to aid in this process. These systems introduce cooling media through dozens to hundreds of unique passages used to manage the thermal process of the cycle. An example of the complexity of these cooling lines can be seen in Figure 16.

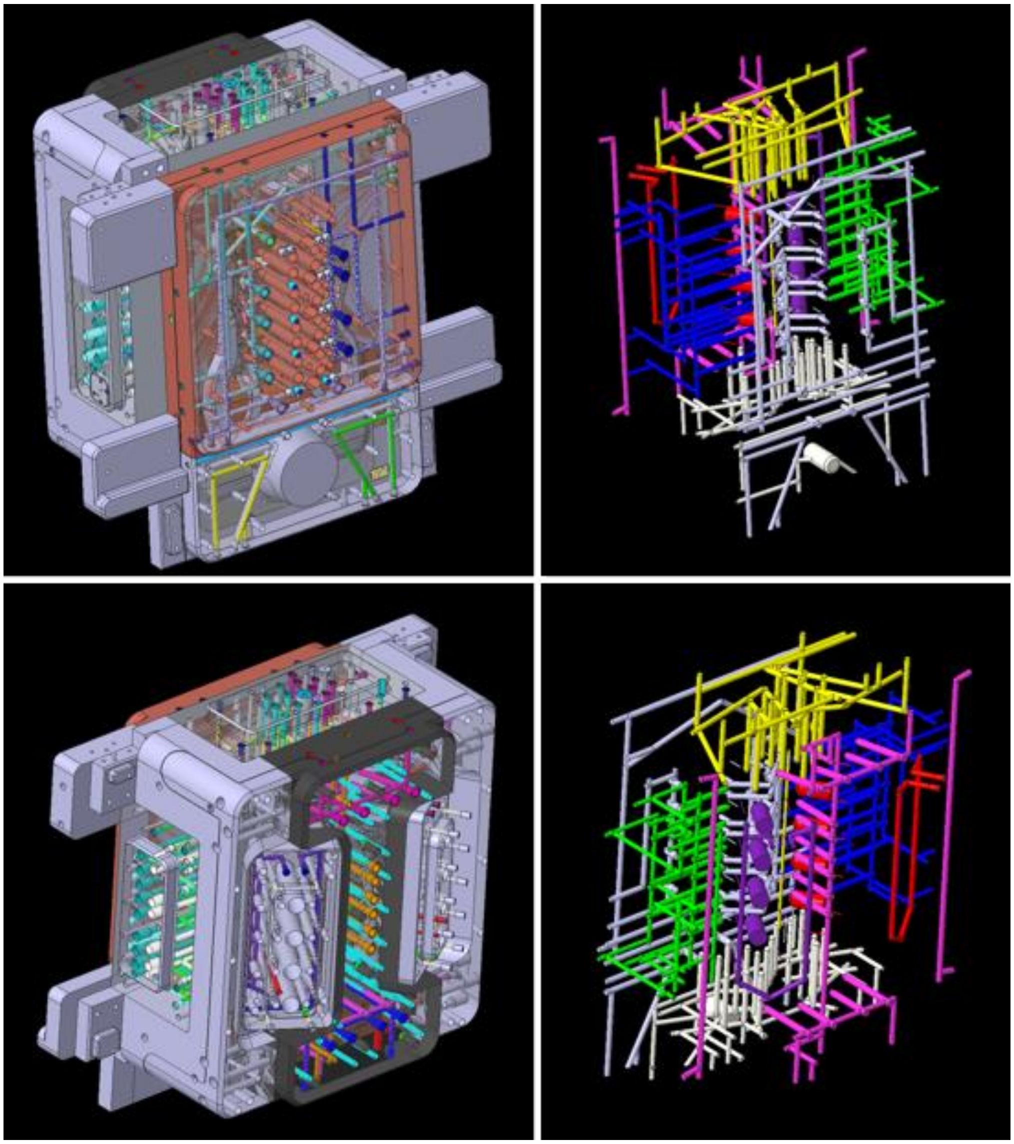


Figure 16: Complex thermal cooling lines in die design
(photo permission from Mercury Marine, image credit Clay Rasmussen)

The complexity of the die design makes it prone to human error during assembly and setup of the die. One line hooked up backward could prevent cooling from flowing through the passage, thereby affecting the thermal management and porosity formation. Systems to check the proper assembly and setup of the tool are needed. Flow rates, temperatures reductions, and pressure readings validate the consistency of thermal management of the die. Figure 17 shows an example of a portion of the thermal lines on a production tool that must be correctly installed during each setup or maintenance of the die.

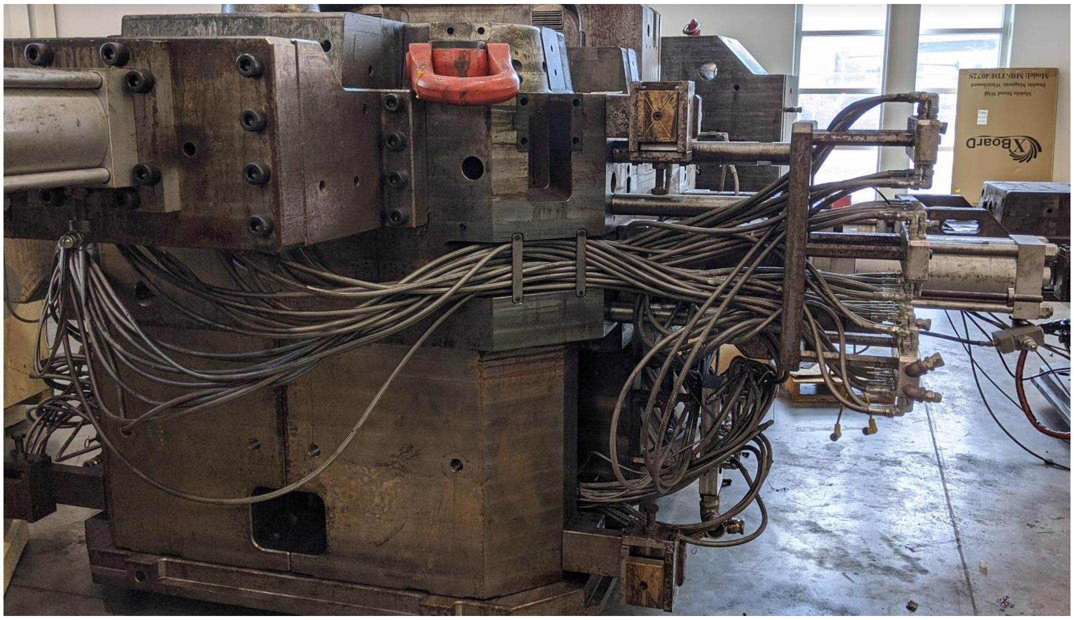


Figure 17: Example of hot water and jet cool lines on a moving half of a large die
(photo permission from Mercury Marine)

EQUIPMENT PERFORMANCE AND ENVIRONMENT

Many of the speeds, pressures, and timing of the die casting equipment depend on the performance of the motors and pumps in that system. Die casting machines typically have two to six different hydraulic pumps used to pressure all the hydraulic systems used for clamping, injection, ejection, and slide pulls. Understanding the performance on these parts of the system provides insight into possible equipment performance issues or outright failures resulting in process stops and downtime events. Temperatures, revolutions per minute (RPM), and vibration analysis are all data streams that can be generated by monitoring this equipment. Oil temperatures used in the large hydraulic systems can also be measured to understand the process.

Environmental factors also influence the die casting process. Ambient air temperature, air flow, and humidity levels can change the die temperature and associated cooling rates. Indirectly, air temperature can affect plant-wide water systems, which can change the input water temperature and recovery times within the cooling units attached to the dies.

SYSTEM COMPLEXITY CONCLUSIONS

Die casting is a complex system. As such, it deserves a methodical, system engineering approach to review and define all the processes and associated systems. The complexity and systems approach for die casting will be further discussed in Chapter 2. Dissecting the complex process with a systems approach will provide the details needed to create a better understanding of the available data. It addresses some of the shortcomings that are seen in current literature regarding die casting optimization and the use of machine learning.

PROCESS OPTIMIZATION LITERATURE REVIEW

There is a history of process optimization research for die casting. Several authors have utilized Taguchi methods to develop optimized process settings to reduce defects [17] [18]. Other authors have utilized statistical methods like regression [19] [20] and machine learning algorithms like genetic algorithm [21] or neural nets [22] to optimize parameter settings. Although there is value in these approaches and the advantages they can provide to the industry, there are also shortcomings from an industrial perspective that need to be understood and addressed.

In many publications, experimental design inputs are poorly chosen for the experiments. For example, Tsoukalas' publication uses a second stage plunger speed of 1.2, 2.5, or 3.8 m/s [21]. Given the casting volume and plunger area provided in the paper, these speeds translate to cavity fill time of 40 milliseconds to 127 milliseconds. Per industry guidelines, porosity and surface finish defects are a trade-off of fill time. Increased fill time leads to reduced porosity but poorer surface finish. Decreased fill times lead to better surface finish but poorer porosity performance [15]. It is not a surprise that the optimization performed in this publication landed exactly at the same conclusion given the range of values tested. One would reach the same setpoints as the author did without any analytics, experimentation, and optimization by using the industrial calculations published by NADCA [15]. In most of the literature reviewed, the ranges of inputs were significantly larger than what would be even considered in industry. In the end, many publications with the optimization method showcased simply matched standard industry practice.

The amount of input parameters generated by the die casting process is significantly larger than typically measured. The most widely accepted data collection system used in the industry measures the injection velocity and intensification pressure system [23]. Typically 20 to 30 inputs are calculated from the time-series data collected [2]. Much of the literature published makes initial assumptions on which inputs are the most important and only gathers data on that subset. Three to five inputs were typically used [17] [18] [21] [22]. This approach makes sense given most of this work is done experimentally with limited samples produced for research. This approach ignores potentially significant inputs in the optimization process. Han et al. was an outlier with more than 20 input variables used in the statistical analysis for optimization [20], but this too is a tiny subset of the potential data die cating can generate.

Finally, the actual experimental output differs from the different publications. Han et al [20] and Zheng et al [22] used a production intent casting design and process to complete their analysis. However, Zheng et al produced about 30 different castings for the analysis, while Han et al used 500,000 castings generated during a year of production at Fiat Chrysler Automobiles (FCA) die casting plant. Other authors produced castings, but the casting was a research design of a simple, thin box that could be produced in a small die cast machine often associated with academic research [17] [21]. These simple box designs are used as easy means of testing density, which is translated into porosity defects based on void space. This approach and casting design is not representative of typical casting geometries used in industry. Their findings have limited industrial application. The final approach used in the literature did not involve making castings at all. Instead commercially available solidification simulation software was used with the different inputs. Outputs of these simulations were quantified and used as the dependent variable to optimize [18].

The literature found regarding die casting process optimization provides useful methods for the industry but lacks the complexity of real world die casting applications for quality prediction. The optimized solutions reached often match basic recommended calculations for the industry. A need exists for a better data collection and optimization process in die casting.

CONCLUSIONS

A gap exists between theoretical solutions and actual applications for data gathering and process optimization in die casting. Closing this gap requires a systems engineering approach to the complex die casting system. Additionally, a foundational understanding of defects and classification is needed to recognize the most realistic applications machine learning.

Chapter 2 discusses the details of complexity theory within the die casting manufacturing system, the need for a systems engineering approach, and ultimately completes a data framework for the die casting system. As will be demonstrated, the volume and velocity of this data being generated in die casting needs machine learning tools to help process and identify important information.

Chapters 3 and 4 build on the complexity of the die casting system in two different ways. Chapter 3 details a study performed on the stochastic nature of porosity formation in die castings and demonstrates the variables the industry typically controls to improve part quality did not influence the random nature of the porosity formation. Chapter 4 discusses the human impact on classification of defects and the importance of the Critical Error Threshold (CET) of the casting process. The CET combined with the high-quality performance of most manufacturing operations, focuses the use of machine learning to areas within die casting outside of the traditional supervised machine learning approach.

Chapter 5 entertains the topics of applications of machine learning within a production die casting process. A review of machine learning and some of the challenges associated with manufacturing are discussed prior to a review of four machine learning case studies completed at Mercury Marine as part of the PhD research. Chapter 6 concludes the dissertation with a discussion around conclusions learned and recommended areas of future studies that can be built off this work.

Chapter 2: Die Casting Data Framework

Chapter 2 will discuss how Industry 4.0 is helping manufacturers create a cyber-physical representation of their process with data. This representation is important as die casting can be formally defined as a complex system. A systems engineering approach is needed to better understand the die casting process and the data that can be generated within each cycle. A data framework for die casting is then presented. This framework shows the data that exists in die casting to be several orders of magnitude larger than what is typically used within the industry.

INTRODUCTION

Industry 4.0 and Machine Learning (ML) have become important aspects of manufacturing in the 21st century [5], [24]–[29]. As outlined by Hoyer *et al* [30], Industry 4.0 relevance, potential, impact, and economic benefit are heavily debated topics. Companies often lack a clear vision as to what Industry 4.0 is, which furthers the debate on the topic [30]. The Industry 4.0 concept was introduced in Germany in 2011[30]–[33]. Although a decade has passed, many companies struggle to understand and apply Industry 4.0. A 2016 McKinsey study showed 6 out of 10 manufacturers face implementation barriers that are so strong that they achieved limited to no progress on their Industry 4.0 initiatives [34]. These barriers can include financial issues, organizational challenges, lack of employee skillsets, and general resistance to change [24]. In 2018, a study of German industry showed that only 9% of companies were able to implement a comprehensive Industry 4.0 approach within their organization [35]. This value actually decreased to 8% in the 2019 study [36]. There is a focus on the technological aspects of Industry 4.0 but not the conceptual use of Industry 4.0. The conceptual context of Industry 4.0 is when technologies are used to create connected business environments for improved, efficient decision making [30]. Instead, many companies adopting Industry 4.0 have a specific project focus and have yet to consider Industry 4.0 on a broader, organizational scale [30], [35].

A systems approach is needed due to the complexity of both Industry 4.0 and manufacturing environments [25], [30], [32], [33], [37]–[39]. The focus of this work is to gain a better understanding of the die casting manufacturing system through data and analytics. The die casting process, the type of data that exist in the process, and the scale of data the process generates can be defined in a data framework through a systems approach. This framework creates a roadmap for the die casting industry to reduce and monitor the complexity within the system. The framework, combined with Industry 4.0 data collection and machine learning analytics, creates the opportunity to solve die casting's critical problems in quality and uptime.

As will be demonstrated, the velocity and volume of data generated within a die casting system is big. This data is not the normal "big data" associated with the large number of rows of data. Instead, the die casting process generates an extremely wide dataset for each cycle. This introduces a host of challenges for machine learning applications. Demonstrating the ability to create value in these datasets becomes the final challenge of this work.

Data creates the connection between the physical and cyber worlds. When applied to manufacturing, data provides endless opportunities to better understand and improve. Before the data framework is defined, a review of the complexity of manufacturing and the need for a systems engineering approach is discussed.

COMPLEXITY AND DIE CASTING

The term "complexity" is readily used in die casting to explain the system. NADCA publications on fluid flow [40], part design [41], conformal cooling [42], die design and construction [41], [43] , heat transfer [44], and the overall die casting process [9], [45], [46] reference the complex nature of an individual technology or sub-system. The different systems, environment, technologies, and people all interact to create a complex process.

As known and used as the word "complex" is within the industry, there are no publications to be found connecting the die casting process with the complexity theory of systems. A thorough review of literature was completed with several discussions on manufacturing systems and plants but no publications on complexity theory applied to the die casting process. A brief review of complexity theory is needed to help understand the need for a systems engineering approach to solve complex problems like die casting.

Like many different scientific terms, “complexity” and “complex system” have no universally agreed upon definition [47]–[49]. In 1978, Vemuri described complex systems as those being large structures that adapt through time and contain a behavioral element [50]. These definitional elements have been built on throughout numerous publications. Mitchell [47] proposed a definition of a complex system as “a system in which large networks of components with no central control and simple rules of operation give rise to complex collective behavior, sophisticated information processing, and adaptation via learning or evolution.” Johnson [48] described many key components of a complex system that include many interacting objects, objects have memory or feedback, objects can adapt their strategies, the system is open and influenced by its environment, the system appears to be alive, the system exhibits emergent phenomena, emergent phenomena arise in the absence of any sort of central controller, and the system shows a complicated mix of order and disordered behavior. Sillitto [49] breaks complexity down into 2-by-2 matrix, where there is complexity to do with the problem/solution on one axis and the subjective/objective complexity on the other axis, thereby creating multiple types of complexity. Finally, Warren [51] proposed several theoretical foundations to complexity that include: dynamics, nonlinearity, nonlinear dynamics, and connections/emergence.

Not looking to enter the fray of definition discussion, the publications reviewed highlight many key themes that are appropriate to use when considering the die casting system. Four components will be used to describe the complex die casting system:

- Network Structure
- Adaptive
- Self-Organizing
- Nonlinear System

NETWORK STRUCTURE

Complex systems are large, interconnected network structures. Many interconnected networks comprise a die casting foundry and allow it to function. Equipment, casting design, die cast dies, priorities, utilities, metal delivery, maintenance, humans (operators/technicians/management) all are unique systems that interact and make decisions that impact the quality and output within the foundry. Investigating an individual cell simplifies the complexity of a the overall die casting system. However, this individual cell remains a complex system, given the interactions and effects that environment (factory) has on that individual cell.

Here is a theoretical example of how the many systems in die casting interact with an individual cell. Imagine the individual die casting cell is at the end of the line for plant air, so its pressure varies in the cell based on usage elsewhere. This means the flow rate for the die lube being sprayed is inconsistent, changing the heat removal rate based on the amount of die lube applied each cycle. Schedule priorities changed within a shift, so the cell that was idle needs to start producing product. Warmup-cycles begin on a cold die, and production starts. The temperature of the air, the production pace of the operator, and the number of stops caused by maintenance issues or breaks all affect the die thermal temperature. The cell stops because the metal-hauler did not arrive in time for a metal delivery. Not only does the operator need to reset all the robots within the cell due to this unexpected stoppage, but the die was open, which allowed additional heat to escape. This heat loss was especially significant, because the operator left the door next to the machine open, since it was hot in the plant and the cool fall breeze felt good. The cooler weather also impacts the water cooling towers for the cold-water system within the die. Consider an overnight operator who started the cell and began the shift struggling with casting quality issues. The frost that occurred over the weekend also reduced the cooling water temperature lower than normal. As the machine cycles through the struggles of startup, the water temperature warms and the quality issues subsided. However, this change did not happen before the operator made some adjustments to the machine trying to fix the issue, thereby altering the process for future cycles.

Die casting is a network of connected systems of equipment, humans, environment, and processes. Individually, these processes could easily be broken down and modeled precisely. This modeling becomes inherently more complicated as the network of interactions is considered between the systems.

ADAPTIVE

A complex system changes, adapts, and evolves over time. It truly is a dynamic system. Die casting is also a highly dynamic system. Process settings, die construction, equipment wear, and hardware updates are a portion of the system that adapt, evolve, and change.

As will be discussed, there are more than 100 individual settings that must be entered into a typical die cast machine for it to have enough information to cycle. Each of these settings controls a portion of the overall cycle. The timers control the overall cycle time and the amount of cooling used within the die. Velocity set points control the fast and slow shot speeds of the injection. Slide pull sequencing controls the order of movement of slides prior to ejection. Individually, certain settings could have a larger impact on part quality than others. However, the interaction of all of them create the overall cycle time, which manages the thermal process within the die. In foundry environments, changes to these settings occur. From a difference of opinion between operators on which settings the machine runs best to quality issues that occur due to specification changes, the process will be modified through time. Much like evolution, some of these changes may improve the process, while others may not. Changes that occur are sometimes immediately seen in the visual casting quality from the die casting machine. However, many quality issues are discovered weeks or months after the change once the casting is machined in post processing. Because humans help develop and control the settings, humans also change these settings to make the process better.

The die casting die, another node of the network, evolves and adapts through time via process and human interaction. The die casting process is a violent process caused by the injection of liquid metal at high flow rates into a steel mold. The liquid aluminum chemically prefers to create an intermetallic bond with the steel. A die will see repair work throughout its life. Areas of the steel will get soft and piece out, requiring the tool to be welded, heat treated, and re-cut to size. The gates see extremely high flow rates as the metal is injected. Through time, these gates will wear larger. The area where the cavity was originally machined may differ significantly halfway through its life without any modifications. This degradation can change the process slowly as technicians try to counter the changes that the machine sees due to this. Humans also cause the die to change. A part revision will create the need to modify the die steel to accommodate the design change. A quality problem uncovered after multi-year testing may drive additional cooling lines to be added to improve the solidification. The modifications led by humans are often significant and a step change in performance of the die from one run to the next. This information is important to monitor within the system.

The equipment used within die casting also evolves through time caused by general use and therefore wear. Valves and fittings in hydraulic systems will wear and fail. Solenoids, cylinder seals, bushings, motors, and pumps have a given life before they stop functioning as intended. Although some will dramatically fail and stop production on the entire machine, others will fail slowly. This may cause issues with machine performance for weeks or months before troubleshooting uncovers the failure. Replacing the component alters the performance of the system which can change the process performance. Typically, evolution brings the ideology of steady and slow improvement to mind. In die casting, the evolution that occurs is a steady decline in performance that needs to be detected and replaced. This creates a complex system.

When a component fails, it's replaced with a new component that changes the performance of the system. Not all equipment changes are done based on maintenance replacements. Die casting technology has evolved through time. This affects the hardware and the materials used within the process. As an example, chemicals in die lube from 50 years ago have been banned and are illegal due to environmental and health concerns. The chemistry of the lube has evolved, so spray processes used in the past become obsolete and need to be completely redone. Hardware has the same evolution. Spray manifolds on booms that move only in two directions are being replaced by 6-axis robots with fewer nozzles that are highly programmable. When implemented, the technology shift fundamentally changes the spray process, thereby altering the thermals of the die. Technology changes in today's Industry 4.0 world also typically provide more options for control and monitoring as compared to the past. Spray is one example. Many other pieces of hardware often change within a die casting cell through the years. Thermal units, robots, dies, hydraulic components, and even the oils and lubes used will all evolve.

All these evolutions and changes make the entire die cast system highly complex. The information from the machine this production run may or may not be fundamentally different due to changes since the last run. The number of components or sub-systems involved add further complexity to this dynamic system. In the end, the process will change through time, which means the data will change as well. Understanding this dynamic, complex system is important. A detailed record and understanding of the data created from system settings to output performance is useful to reduce the complex system into more manageable pieces.

SELF-ORGANIZING

Complex systems often experience self-organization within the system. Self-organization can be defined as the "emergence of global structure out of local interactions" [52]. On the surface, this self-organization may seem counter to the top-down management style of many manufacturing organizations. The plant manager or supervisor makes a decision, and the staff blindly executes. This, of course, is not true in all plants. Even in plants where this façade exists, a level of self-organization will still occur. Technology, process, or system improvements tested in one area can be adapted throughout a plant with no formal directive. One operator talks to the next, and the network within the system can pass information and organize itself to a different state. Ideally, feedback loops within the system drive the system to an improved state with less disorder.

My experience has shown self-organizing to be alive and well within a die cast foundry. Most foundries have top-down directives of setups and schedules. However, many decisions are made at the local level and drive this self-organization. A theoretical example would be a die casting technician making a quality improvement by increasing the slow shot speed on a given casting. This technician talks to another during a break about the improvement just experienced. The second technician struggling on a different job, tries a similar approach that is successful as well. These two examples with positive feedback loops drive the change to other areas. The progress continues to spread at the local level until all parts are adjusted. No one directed this change, but it evolves through the network and improves the overall system. This is just one example within die casting. Other examples where self-organization likely occurs, and the author expeirned, are changes within tool design, how equipment is maintained, technology improvements (spray, chambers, extraction), and process approaches.

In the long-run, these self-organization systems drive improvements and reduce disorder within the system. Occasionally the feedback loop may be extremely slow. A change could be made and spread locally throughout the organization only to learn there is an issue that increased scrap rates or downtime on a machine. The speed of undoing this change will depend on the severity and scale of the issue. A top-down directive could quickly undo the process. If not fully understood by management, the feedback loop, will eventually drive the individual to make changes to bring the system back to its previous state. The die casting system shows self-organization which makes it complex.

NONLINEAR SYSTEM

The final topic used for complexity is nonlinear systems. In a nonlinear system, the effects of an input are not proportional to the size of the output. This can be seen in both positive and negative feedback loops. A positive feedback loop in a nonlinear system can be extremely sensitive to its initial conditions. Wild changes can occur within these sensitive nonlinear systems, making prediction impossible. The concept of mathematical chaos is driven from these types of systems. That is the challenge of complex systems: they exist between simple, orderly systems that are fully predictable and chaotic systems that are completely unpredictable. A complex system will show features of both given their nonlinear components.

Fundamentally, die casting has several non-linear features. The basic process of die casting is fluid flow and heat transfer. The nonlinear Naiver-Stokes equations model fluid flow within a three-dimensional space. The injection process of molten aluminum into the die presents countless opportunities for varying initial conditions of tool temperatures, surface conditions, and timing when comparing cycle-to-cycle. Heat transfer within the die casting system is the other key process involving nonlinear systems. The radiation equation has temperature of the environment and feature both raised to the 4^{th} power [53]. Any variation in initial conditions between cycles could have meaningful differences within the process shot-to-shot. Beyond radiation, many of the conductive and convective heat transfer equations are simplified with a static heat transfer coefficient. However, the thermal conductivity is a function with temperature thus making the traditionally linear heat transfer equations now nonlinear. Beyond heat transfer within the die, the energy and heat released with the phase change of the metal is nonlinear [53]. The latent heat required at phase change is not linear during these transitions. Again, depending on stating conditions, the amount of heat removed from a die can vary with different initial conditions.

Die casting is a complex process because the nonlinearity of fluid flow and heat transfer associated with the process. The die casting process is sensitive to its initial conditions. With data capture, these initial conditions can be collected and stored to help analyze what impact they may have to the die casting system and how one can help reduce the chaos associated with die casting.

COMPLEXITY CONCLUSIONS

Die casting is a complex system. Many different nodes within a network of interactions comprise the die casting system. Die casting adapts and evolves from the feedback within the system. Processes will self-organize based on these best practices experienced in different areas of the plant. Finally, die casting is nonlinear and outcomes will be dependent upon its initial conditions. Figure 18 is an overview diagram of these components in the complexity of the die casting system.

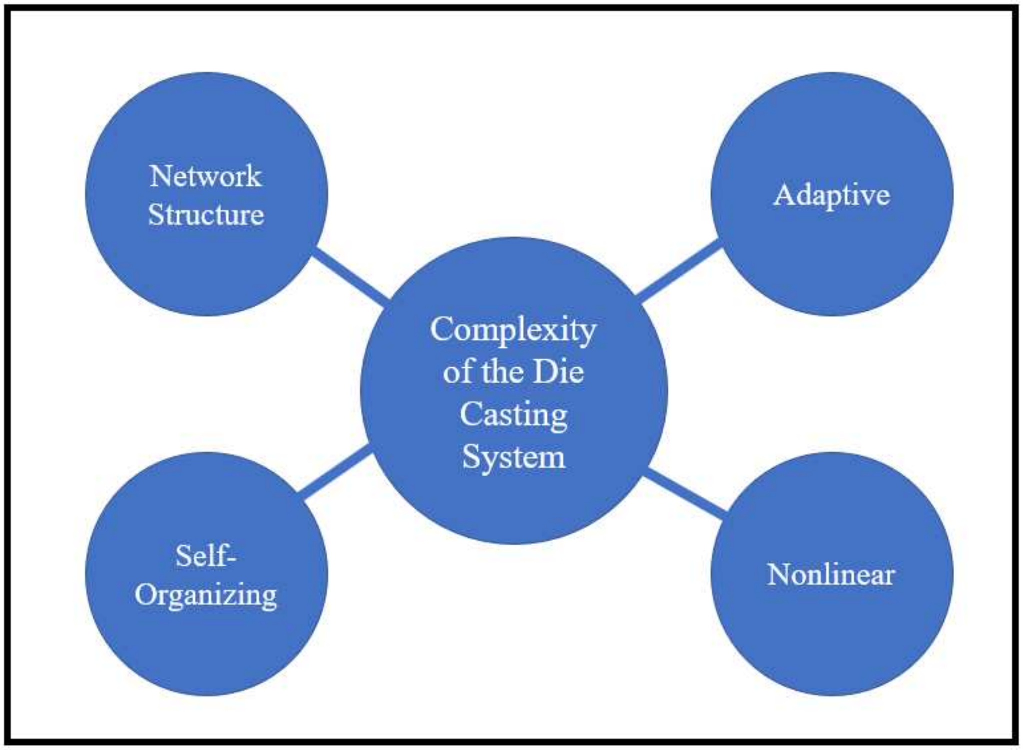


Figure 18: Complexity in Die Casting

Comprehension and appreciation of the complex nature of die castings is the first step to improved process knowledge. The next steps are utilizing the tools needed to analyze a complex system. Systems engineering and a systems approach will dissect the complex nature of die casting into manageable pieces that can be analyzed and understood. The next section discussions this systems approach needed to create a die casting data framework.

SYSTEMS APPROACH

Systems engineering (SE) is a disciple designed to address highly complex systems [54]–[57]. The toolset that exists to break down and evaluate systems creates improvement and understanding of that system. Both Industry 4.0 and die casting represent two unique complex systems. Logic would initially lead one to believe that the combination of two complex systems would create a larger, even more complex system. Although this may be the case, an argument can be made the addition could reduce the overall complexity. The complexity of one system can reduce the complexity of another [58] by providing the feedback needed to better control the original system. Before that argument can be made, a background of the systems engineering is needed.

SYSTEMS ENGINEERING APPROACHES

Traditionally, systems engineering (SE) is an approach used to engineer and implement complex systems [54]–[57]. Each textbook published has its own systems engineering definition. Many share traditional SE concepts of complex systems, life-cycle analysis, engineered systems, and risk management. The words of each definition may be different, but the goal is the same. For simplicity, the author will reference the definition provided by the International Council on Systems Engineering (INCOSE), a not-for-profit membership organization and professional society for systems engineering :

> "Systems Engineering is a transdisciplinary and integrative approach to enable the successful realization, use, and retirement of engineered systems, using systems principles and concepts, and scientific, technological, and management methods." [59]

One concept in most definitions of systems engineering is the system lifecycle. SE focuses on the design, development, testing, implementation, and end-of-life of a system [54]. Methodologies exist that provide an approach for engineering and implementing complex systems. One of the most used in SE is the "V" diagram. This diagram shows the decomposition of requirements during design and the integration during the validation. The "V" diagram stresses the connections between the requirements during development and validation. An example of this diagram can be seen in Figure 19.

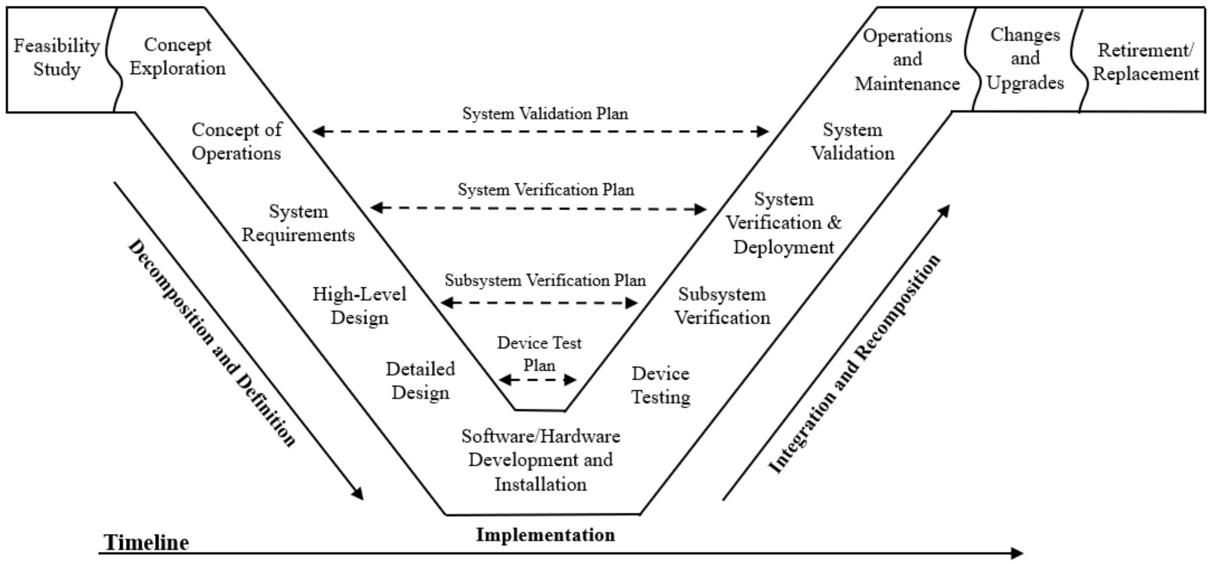


Figure 19: Systems engineering "V" Diagram based on [60]

Most frameworks provided within the SE disciple follow this same lifecycle approach. This is extremely useful when designing a new system, but it is not particularly useful when looking to apply a SE methodology to an already existing system, such as a die casting manufacturing system. The focus of manufacturing systems is realistically the long-term application of "operations and maintenance" located at the top of the "V". If the entire manufacturing system was designed with a systems approach, the systems approach would have already been considered. Unfortunately, most manufacturing systems are not designed from a SE perspective. They are complex systems that have been created and evolved over time. This means one must approach manufacturing systems from an SE perspective differently. Instead, two other concepts from SE will be utilized to help gain an improved system understanding for die casting. These tools include a decomposition of the system hierarchy and a system context diagram.

Decomposition of a system is a critical component to the systems engineering approach. The entire left side of the "V" diagram involves breaking down the requirements and components of a complex system into manageable parts. This is fundamentally what a SE helps accomplish. The dissection of systems into smaller elements makes them easier to understand, design, and ultimately control. The framework typically used for this decomposition of the hierarchical nature of systems is seen in Figure 20.

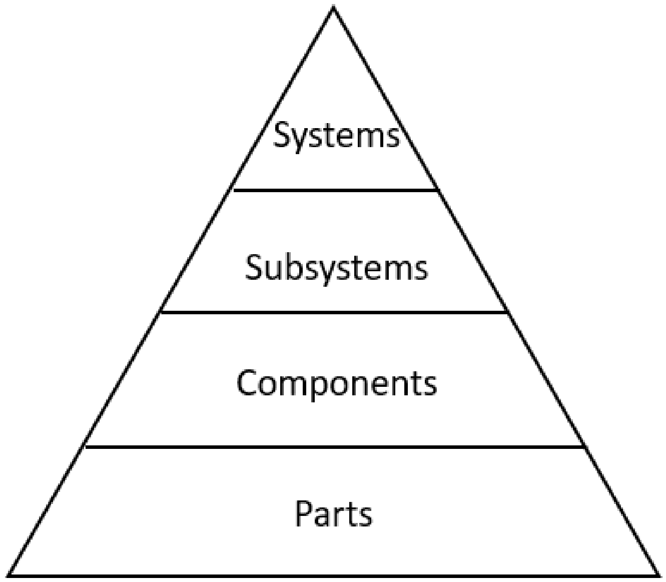


Figure 20: Decomposition of hierarchical systems

A systems context diagram shows a system's interactions. Typically, the diagram shows the interactions with external entities [54]; however, defining external entities versus internal entities depends on how the system is defined. The approach used to create Figure 21 was based on the key process interactions that exist within die casting to the networks of systems involved. An explanation of what a die casting system is must be defined. For most die casting data frameworks, a die casting system will focus on the die casting process within one cell. As mentioned earlier when describing complex systems, the entire foundry likely could be defined as the die-casting system. The network, evolution, and self-organizing nature takes place typically across cells. However, to create a useful data framework that will benefit the industry, the focus will be on one cell. As a result, Figure 21 is based on this goal with the data framework in mind.

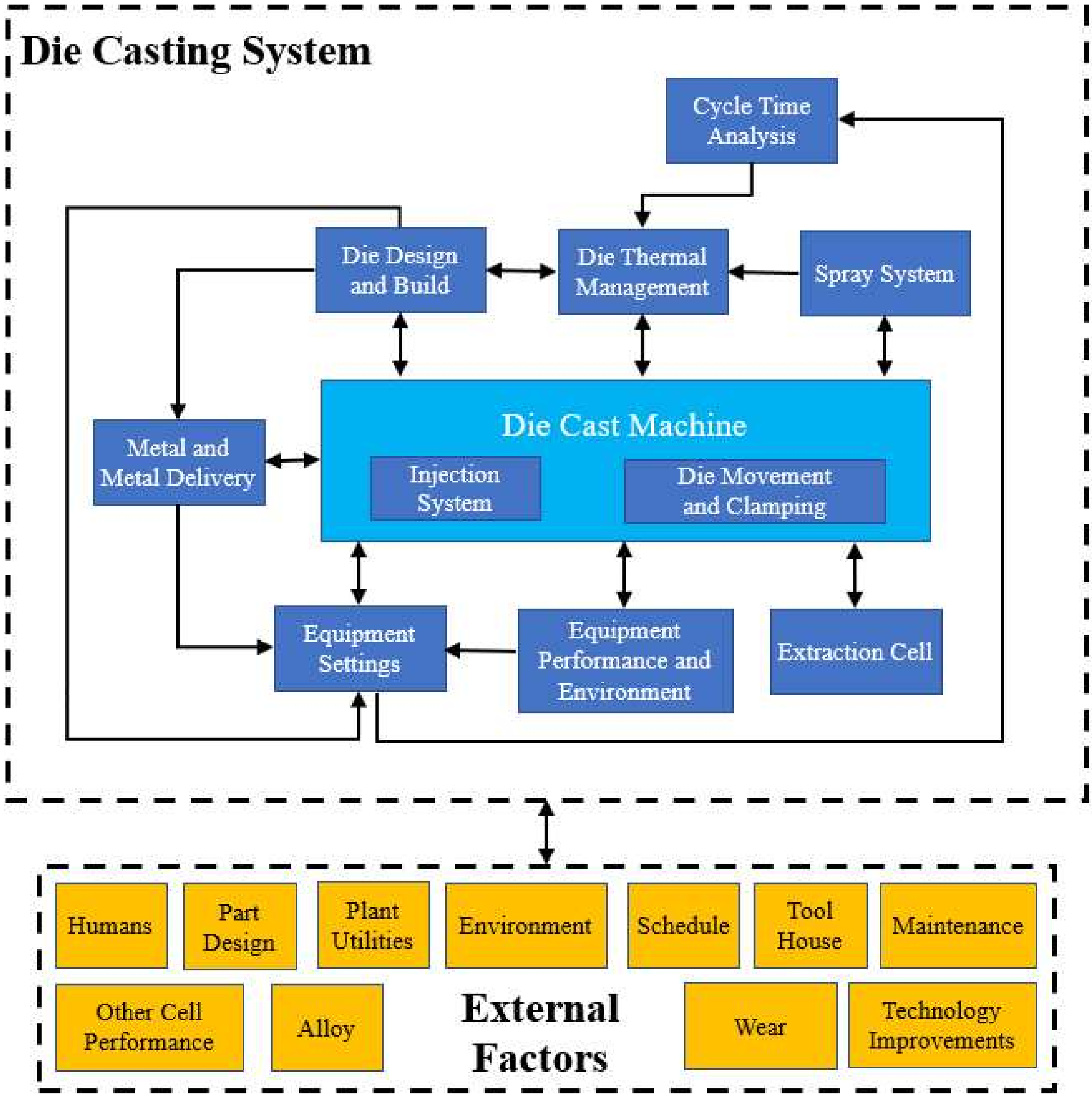


Figure 21: Die casting context diagram

As the context diagram shows, many different elements comprise the die casting system. Defining these as subsystems or another entire system is difficult given the complexity and nature of work involved. For example, "Tooling Management" associated with the die design, build, maintenance, and scheduling could be classified as a system in and of itself, which is integrated into the die casting system. To help illustrate this example, Figure 22 was created to show the detail involved within management of the die casting tooling. Similar argument can be made for metal management; plant utility interaction with multiple subsystems like spray, thermal management, and the equipment; and equipment maintenance systems. The integration of systems and the concept of system-of-systems is important aspect of SE to review as a data framework associated with the die casting system is constructed.

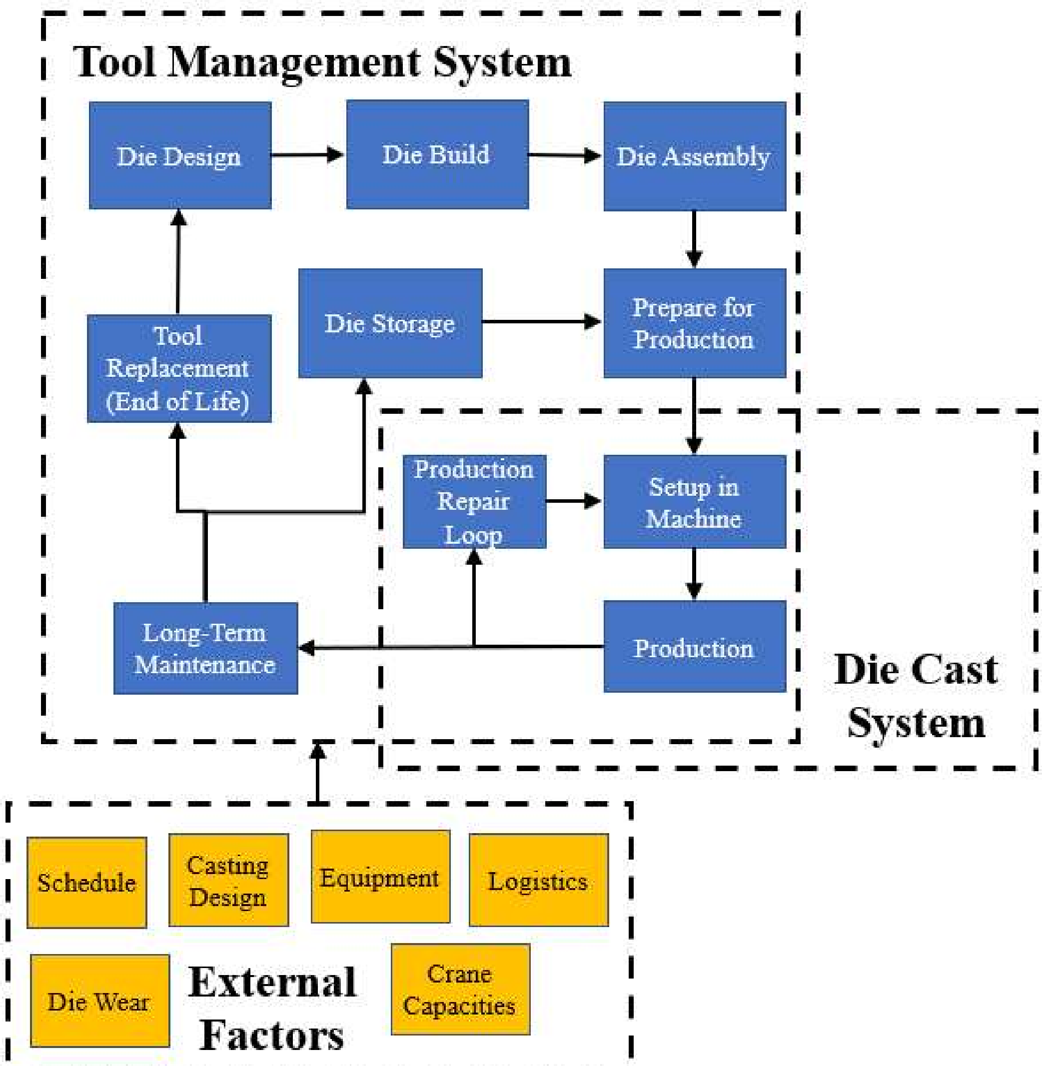


Figure 22: Tooling Management system diagram

A 2019 publication defines systems-of-systems (SoS) as an approach to understanding highly complex, multi-system technologies that interact and work together for a desired outcome [61]. Similarly, a recent thesis created a framework for distinguishing between a complex system versus a system-of-systems, while acknowledging that no universally accepted definition exists for SoS [62]. By these definitions, the die casting manufacturing process could qualify for a system-of-systems description.

These SoS definitions have grown in complexity from the original publications regarding system-of-systems.

Using the SoS terminology for this research came from the original definition of system-of-systems. This concept of "system-of-systems" was first used by Boulding in 1956 [63]. He used SoS to build an argument that multi-disciplinary work could build a better system than its individual parts. Both Boulding and Ackoff [64] argued for the need to develop a framework of the system to help lead to better comprehension and future research opportunities. In the end, that is ultimately the goal of this research: develop a framework defining the multiple systems and subsystems in die casting and the data they generate. This data framework helps create a foundation for machine learning applications within die casting from a systems perspective. Similar to how the periodic table was developed and openings were left for elements not yet discovered [63], this die casting framework provides a structure for the industry to understand the process that evolves with new learnings and technological advancements.

REDUCING COMPLEXITY WITH FEEDBACK

The goal of systems engineering is to reduce complexity within a system. Complex systems maintain equilibrium with feedback control [65]. Feedback is a fundamental systems engineering principle and is an essential component when working with complex systems [66]. Typically, one thinks of feedback control as the process used to monitor and control equipment. Cruise-control on a vehicle is an often-cited example of a feedback control application. The goal with feedback is for the controller in the system to adjust inputs based on results from outputs to maintain a targeted output. Traditional control systems would see feedback control as the cruise-control adjusting the fuel going into the engine based on the reduction of speed going up a hill. Alternatively, feedback within a biological system is a plant wilting to conserve water when the plant's feedback system shows a shortage. The cruise control system can be modeled mathematically and executed in the controller and sensors that comprise the system. The feedback process in biological systems can be represented mathematically but is biologically controlled. Both types of feedback are important in die casting.

Feedback control already exists in much of the die casting equipment. Examples of control systems include the feedback loop on the injection system, the temperature control on thermal units to maintain a set water temperature, and the chiller on the hydraulic unit of the die casting cell starting once the hydraulic oil reaches a certain temperature. In these cases, sensors take measurements of the system. With this information, the controller of the system initiates a response within the system. Other feedback systems beyond equipment exist in die casting as well. Typically, quality of the casting is a feedback to the process. If an operator sees a casting defect on the parts coming out of the cell, he or she will likely make an adjustment to the process to help eliminate it. Sometimes, this feedback loop has a large time delay due to additional post processing, so the castings with the quality issue may be discovered long after the production run is done. The feedback would be used for an improvement in a future production run. Both systems are close-looped feedback. A generic diagram of close-looped feedback can be seen in Figure 23.

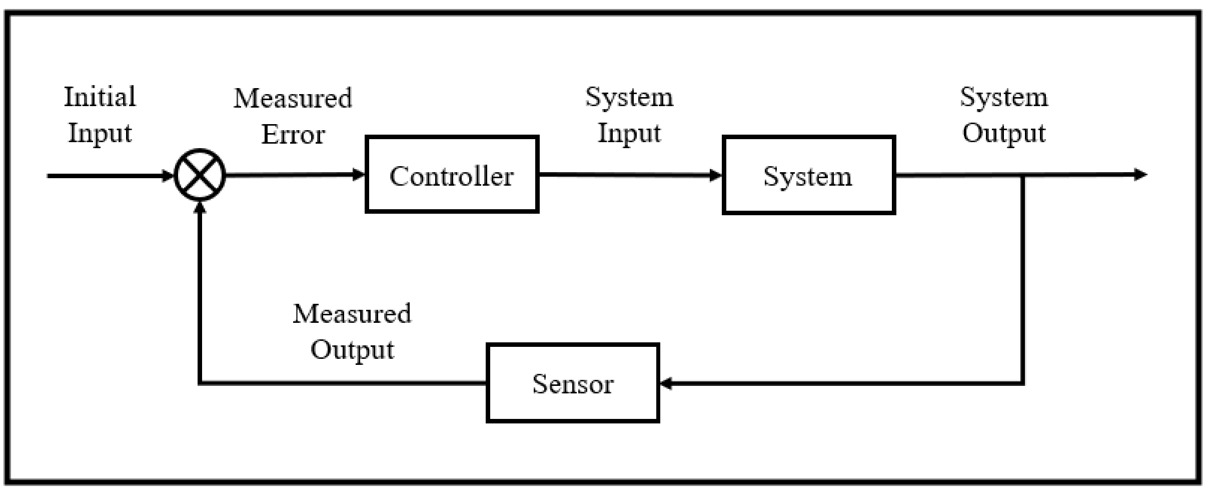


Figure 23: Generic closed-loop feedback system

Review of the close-loop diagram shows several critical pieces of data needed to help with the control. The initial input into the system needs to be understood. In die casting, and in the upcoming data framework, this comes in two different forms. The first is setting information associated with the equipment. At the start of injection process, the hydraulic system is stopped. The system initiates and says the pressure must be increased in the cylinder since its current measurement is at zero. The controller opens the valve allowing hydraulic oil to enter the cylinder, raising the pressure. These initial settings are critical for a complete system understanding of how the equipment is performing. The other critical data is the sensor measurement in the feedback loop. This tells the story of the output of the system and helps explain the actual process that was performed. The setting input parameters and the measured output values are two of the four types of data that will be discussed in the next section.

At the component level within the process, the feedback system is straightforward. When looking at the system of die casting, this feedback system becomes fuzzy and, at times, conflicting. The need to hit production schedules may override the desire to perform needed maintenance to fix a leaking hydraulic hose creating a small puddle on the floor. Individual systems may be monitored and controlled, such as temperature on a hot water thermal unit, but there is often no measurement or feedback system to ensure the water is properly setup and flowing as needed. The system relies on humans to perform many of these tasks. The casting quality is then used as feedback to the system, but as will be shown, it can take days, weeks, or even months for that feedback.

The lack of complete system monitoring of the die casting process creates many of the issues that cause poor quality and downtime within die casting. Feedback loops monitor and control the system. Data collection associated with Industry 4.0 is creating the opportunity to get this feedback in die casting. As stated previously, Industry 4.0 is a complex system. Most companies that implement it do so with point solutions. Limited companies are attempting system-wide Industry 4.0 implementations. However, the benefit of having a complete picture of input settings and output variables collected, stored, and monitored outweigh the added complexity of implementation. Industry 4.0 will reduce the complexity of the die casting system by providing feedback on the entire die casting system. Industry 4.0 is a worthy endeavor for the die casting industry.

SYSTEMS CONCLUSIONS

Die casting, a complex manufacturing system, contains the four key components of complex systems. Die casting is a network structure. It is a dynamic system, able to adapt and evolve through time. Foundries will self-organize and improve through time without top-down directives. Finally, many of the processes for fluid flow and heat transfer are nonlinear, and small changes with the input can result in significantly different outputs.

Defining die casting as complex from a systems perspective provides a toolset from the systems engineering discipline to understanding the system. Traditional SE tools like a "V" diagram are useful in designing and developing new systems but require an altered approach when applied to manufacturing systems. Utilizing tools like the decomposition of the die casting system and context diagrams allow one to apply the SE toolset to an "in-production" system compared to the traditional new system design project.

A statement from Schneider *et al* [58] summaries the need within die casting well: "a system can only control something to the extent that it possesses sufficient requisite variety to form a representation of that thing." Without full understanding the process and data being generated, control of the system is not possible. A systems analysis and resulting data framework allow for better understanding and control of the die casting system. The data can be used to provide feedback for this control. The feedback and control can remove the variation that naturally occurs in the process that would otherwise be unknown. Reduction of variation helps create a consistent thermal cycle within the die by eliminating process variation and improving equipment uptime. These reductions improve casting quality and equipment utilization, two struggles in the die casting industry.

In the end, the goal of the data framework provides a means to improve the die casting system by better defining it. This definition will improve the complex system in four categories: performance, stability, robustness, and flexibility [65]. The first will be to improve the performance of the system to a cost constraint. Die casting achieves better performance with increased production throughout, quality improvements, and downtime elimination. The second category is to increase the stability of the die casting system. Stability is defined in the mathematical sense. As such, a data framework will provide a better understanding of initial conditions and how the nonlinearity of die casting is influenced by these conditions. The third category is robustness of the system to work across a wide range of operating conditions. Being able to develop a more robust process in manufacturing starts by understanding where the operating conditions are today. Having data gives the industry this knowledge. The last improvement to be made to the system is with flexibility. The framework provides insight into the type of data collected and the opportunities that exist from the current system and subsystems in die casting. As technology or equipment change, the framework needs to flex along with the improvements to still provide guidance on critical processes and data. Improving the performance, stability, robustness, and flexibility of the die casting system is critical to improve the output of the system. The next section reviews and provides the process and data framework within die casting to help achieve these goals.

DATA FRAMEWORK

By utilizing a systems approach, the different types of data from the die casting system can be defined and a framework built around the process. This section will provide the framework for die casting data and tables containing the data that can be generated, collected, and stored to aid in the feedback and control of the process.

It is important to realize the exact number of variables and the data generated within die casting is an estimate only. The die casting industry continues to evolve. Technology continues to evolve the die cast equipment (3-plate machines versus 2-plate machines) and demands for products change the size of machines used. In 2014, a NADCA survey showed the vast majority of North American die cast machines were 400 to 1,500 tons, with less than 50 machines being larger than 2,500 tons [4]. In the last several years, record-setting die cast machines have been put into production. It started in 2017 when a 4,500-ton die cast machine was implemented for production of engine blocks [67]. Then, in 2018 a 6,100-ton machine was announced and started production in 2020 [68]. Most recently, a 8,000 ton machine was announced in 2021 [69]. Directionally, the framework will capture much of the data, although it must be understood this will not be all encompassing and the data available will change through time. The goal here is to show the complexity, depth, and detail associated with possible data collection in die casting. With this volume of data, analytical tools like machine learning will be required to process the data and control the process.

Different types of data can be generated from the die casting process. The die casting industry has popularized descriptive statistics of time-series data. Average fast shot velocity, average slow shot velocity, and average intensification pressure are all values that are calculated from the time-series data collected during injection. This data has existed the longest within the industry and is often the focus of research. As this section will show, these statistics are a small subset of data that exists within the die casting process. A system-wide review of the types of data is needed to truly gain a better understanding of the process.

To understand the need for different types of data, one must consider the types of problems trying to be solved. Costs associated with scrap castings, quality, or the prediction of defects within the casting, is important. A linear thinking approach to this problem may focus on the injection data given the focus and publication on the topic. Casting quality, however, is comprised of much more than the flow of metal into the die as system thinking shows. A poor die design will create poor quality regardless of the injection parameters. The placement of cooling lines within the die will directly affect the die surface temperature, altering the thermal neutral axis and where porosity is likely to form. If these cooling lines are blocked due to hard water deposits, the flow rate and calcified lining will not remove the same amount of heat as a clean die. A spray cycle with a leaking nozzle can leave excess liquid in the die cavity that will create gas porosity in the casting. The settings associated with the machine dictate the machine's movements and thereby the cycle time of the different sub-systems. Wear or failure of pumps or motors may prevent the machine from performing properly. These are just a few examples of why system thinking is needed to optimize the die casting process. Collecting data and focusing on only one sub-system of the process will cause frustration and a lack of progress in improvement.

There are five unique groups of data types identified in the die casting process. Each type of data serves a different purpose. Many of these groups interact throughout the process. Some of the data comes from physical dimensions of tools within the process, while others are settings for the PLC within the die cast machine to use for executing the process. All the data types are important for a systems understanding of die casting. These five groups are seen in Table 2.

Table 2: Data types found within die casting

Data Types
Design Parameter Data
Input Settings Data
Output – Discrete Data
Output – Time-Series Data
Cycle Time Analysis Data

The part design, and corresponding die based on the part design, controls the casting quality. The die is often thought of as a fixed component, but it will change through time. From wear associated with metal flowing through gates to casting design changes implemented by product design groups, the tool is modified throughout time. As a result, these modifications are critical to track. For example, imagine all the process data was stored for the 20,000 cycles that were put on the die to date. Feedback from machining of the casting shows a higher than anticipated scrap rate for a porosity found in a cast feature. To address this concern, an additional water line was added to the tool and a gate area was increased to provide better intensified pressure to the area. The two changes eliminated the scrap issue, but a problem exists in the data. The additional water line means additional water flow rates for the cooling lines. The larger gate area will impact the fast shot velocity, as there is less restriction when trying to inject the liquid metal. If the data associated with this change is not recorded, then one would look back and say the scrap problem was addressed by increasing water flow and changing the fast shot speed. The reality is these items changed because of modifications to the design of the die. This is an important distinction that can be lost in the complexity of the die casting process and highlights the need for good data tracking on design parameters.

Input settings for the die cast machine are also a critical data source. Input settings include all the values and parameters that must be entered into the equipment for it to function. They involve timers, target values, limits, sequencies, speeds, and pressures for the machine. This data must be referenced as changes within the process outputs are seen. For example, if a change in slow shot speed it identified as an output, the first review should be to see if changes to the input settings caused the change in the output. Did the operator change the target velocity for that zone? Did the start point of the velocity average window change? Did the speed control get changed from an open loop to a closed loop system? Ideally, these values should remain fixed, but in practice any of those changes could happen by a technician. Without ruling those changes out, time will be wasted investigating other potential failures such as verifying the hydraulic valving or testing for leaks in a cylinder seals. Additionally, if one verifies none of these settings changed, yet the output has changed, the data points to a potential machine failure to address. Input data to the process provides the backbone for examining the outputs that are gathered during the casting process. This is critical data that must be stored during the cycle.

Output data is typically the focus of data collection efforts. Understanding how the machine performs is needed to validate process control. Output data exists in two forms – discrete data and time-series data. Discrete data is individual readings of a specific output. The furnace temperature at ladling is an example of a discrete data point. Time-series data is a variable that changes through time. Velocity during the injection process is an example of time-series data. Many time-series data sources exist within die casting including shot velocity, cylinder pressures, water flow rates, die lube spray flow rates, and pressures of slide pulls. Time-series data is often simplified with descriptive statistics like averages across periods of time. These single data points are easier to visualize trends cycle-to-cycle. Output data is sometimes measured back to input data such as limits that are expected. If an output crosses a limit threshold, the casting could be deemed as scrap by the machine. Output data of both the process and the machine need to be collected. Output data on the process can assist with optimizing the casting process. Output data from the machine can provide insight into predictive maintenance and downtime avoidance. Discrete and time-series outputs comprise the largest share of the data generated by the die casting process.

The last data type to consider is cycle time data. Cycle time data refers to the time difference between multiple steps within the die casting process. The part-to-part cycle time is important in the die casting process for both financial and quality reasons. Cycle time of the equipment determines the cost of the casting. Improving cycle time reduces costs. Additionally, because die casting is a thermal process associated with transforming liquid metal to a solid casting, the cycle time also affects the heat transfer within the process. Troubleshooting and optimization of the die casting process comes a thorough understanding of the process cycle time. Optimization must consider quality changes from thermal cycle adjustments balanced with financial improvements form cycle time reductions. As will be discussed, this is done by collecting time data associated with each movement and process within the die casting system.

The interaction and complexity of die casting make all these data types important. Defining the types of data also creates a better understanding of the potential data to be gathered during a system review of die casting. These types also help shape some of the groupings that will be presented. The goal for the remainder of the chapter is to define many of the critical systems, inputs, variables, and processes, so a detailed data framework for each of these groups can be created. It must be noted the variety associated with die casting, from the equipment used to the parts produced, drives significant variation in the amount of data that can be generated. The framework provided is meant to be a guide to foster the systems approach to thinking about data within die casting. Exact values and number of data points will be different from machine to machine and foundry to foundry. It is not the intention to create a bible of the data, but a framework that can be expanded and modified as needed with changes in technology within die casting. The result of this system review is a data set significantly larger than the die casting industry has considered in the past. With this realization, novel approaches such as machine learning must be utilized to analyze and gain insight on the data that is generated every cycle.

The framework groups, per Table 3, represent the key sub-systems associated with the overall die casting system. These were initially referenced previously in Figure 21. Analyzation of the data associated with each of these sub-systems can allow a complete framework of data to be created. Each framework group will be detailed in separate sections. The section will conclude with a table that will summarize all the data associated within the framework group. The result will be a full picture of the data that can be collected and stored to understand and optimize the die casting process.

Table 3: Die casting data framework groups

Data Framework Groups
Die Design and Build
Die Cast Equipment Settings
Injection System
Die Movement and Clamping
Equipment Performance and Environment
Metal and Metal Delivery System
Thermal Die Management Systems
Spray System
Cycle Time Analysis
Extraction Cell

DIE DESIGN AND BUILD

The design and build of a die casting die involve hundreds of different decision points. The goal is to design a die capable of produce a high-quality casting with a reliable tool. Industry standards provided by NADCA [15] or die manufacturers often guide the die design process.

The process of die creation begins with the design of the desired casting. Geometric, dimensional, surface finish, and porosity requirements must all be understood before selecting die design features. The geometry of the finished part dictates orientation of the part in the die, number of slides, likelihood of shrink porosity, flow distances across the part, avoidance of unbalanced filling, and best areas to gate into the casting. Dimensional requirements drive decisions such as shrink factor used, selection of which part of the casting goes into the stationary versus moving half, variation in tolerances created by slides, and tolerances between parting line halves. Surface finish and porosity requirements are opposing needs. High surface finish parts require fast fill times of the cavity, but this increases the likelihood of gas porosity trapped in the casting. If porosity and pressure tightness of a machined casting are required, a slower fill time is required. This tradeoff is controlled by the selection of the percent solids allowed during injection. As seen in Table 4 [15], the range of percent solids selected affects the surface finish and porosity of the final casting. This percent solid selection is a critical parameter for calculating the fill time of the cavity as seen in Equation 1 [15].

Table 4: Percent solids design selection for different quality requirements [15]

Percent Solids	Quality Requirement
0% to 10%	Highly Decorative
10% to 20%	Decorative
20% to 30%	Some Porosity and Knit Lines
30% to 40%	Low Porosity, Poorer Surface Finish
40% to 50%	Lowest Porosity for Machined Pressure Tight Castings

Equation 1 [15]

$$t = k * c_t * \left(\frac{T_i - T_f + (S * Z)}{T_f - T_d}\right)$$

Where:
t = ideal fill time (seconds)
k = solidification constant (seconds/inch)
$T_i =$ **metal temperature at injection (°F)**
$T_f =$ **minimum flow temperature of alloy (°F)**
S = Percent Solids Allowed (fraction x 100, ie: 50% = 50)
Z = Latent Heat Constant per percent solids (°F/%)
$T_d =$ **die temperature at injection (°F)**
$c_t =$ **minimum or average casting thickness (inch)**

For aluminum die casting on typical H-13 die steel, the solidification constant (k) is 0.866. Minimum flow temperatures (T_f) for aluminum alloys range from 1060 to 1100 °F. The latent heat constant (Z) is 6.8 for typical die cast aluminum alloys. This information and additional values for other die cast alloys can be found in different NADCA publications **[15], [70]**.

With fill time determined, the gate area for the casting to be filled through can be calculated based on a series of equations and decisions. Initially, gate thickness (G_t) must be determined. This thickness will impact the overall length of gate needed but also plays into the trimming and removal of the gates from the final casting. The steps used to determine the gate area will likely be iterative during the design process. With the gate thickness determined, the first step is to determine the atomized velocity (A_v) of the liquid metal injected into the die. This is determined in Equation 2 [15]. The J factor constant for aluminum is 400 [15].

Equation 2 [15]

$$A_v > \left(\frac{J}{G_t} * \rho_l\right)^{0.588}$$

Where:
$A_v =$ **atomization velocity (in/sec)**
J = J factor constant
$G_t =$ **gate thickness (inches)**
$\rho_l =$ **liquid metal density**

The minimum atomization velocity is then used to calculate the minimum gate velocity ($G_{V\text{-}min}$) which includes a flow distance factor (FDF) based on how far metal must flow through the casting to ensure the atomization of the liquid throughout the filling cycle. The minimum gate velocity is found in Equation 3. The FDF factor is based on NADCA-supplied tables as shown in Table 5 [15].

Equation 3 [15]

$$G_V = A_v * FDF$$

Where:
$G_V =$ **minimum gate velocity (in/sec)**
$A_v =$ **atomization velocity (in/sec)**
FDF = Flow Distance Factor per table

Table 5: Flow Distance Factor (FDF) for minimum gate velocity equation [15]

Flow Distance (inches)	Flow Distance Factor (FDF)
0 to 2	1.25
2 to 5	1.50
5 to 9	1.75
9 to 14	2.00
14 to 20	2.25
20+	2.50

Gate area is then calculated based on minimum gate velocity and the flow rate (Q) required to fill the casting. This flow rate is calculated in Equation 4 [15] based on the fill time calculated in Equation 1 and the casting volume (V). The gate area is then calculated based on the flow rate and the minimum gate velocity as shown in Equation 5 [15].

Equation 4 [15]:

$$Q = \frac{V}{t}$$

Where:
Q = flow rate (in^3/sec)
V = casting volume (in^3)
t = fill time (sec)

Equation 5 [15]:

$$G_A = \frac{Q}{G_V}$$

Where:
$G_A =$ gate area (in^2)
Q = flow rate (in^3/sec)
$G_V =$ minimum gate velocity (in/sec)

With the gate area calculated and the gate thickness selected to find the atomization velocity in Equation 2, the gate length (G_l) is calculated in Equation 6 [15]. The gate length will impact the number and distribution of gates selected to feed into the casting. Based on the casting design, finding enough length may require certain part orientations within the die.

Equation 6 [15]:

$$G_l = \frac{G_A}{G_t}$$

Where:
$G_l =$ gate length (in)
$G_A =$ gate area (in^2)
$G_t =$ gate thickness (in)

The equations above walk through an example to find the minimum velocities and flows needed for die casting. The design of the die casting machine determines the range of these values. The hydraulic system, including the injection cylinder piston diameter, injection cylinder rod diameter, and working hydraulic pressure, affects the metal pressure during injection. The minimum metal pressure is determined based on the minimum gate velocity previously calculated. The maximum metal pressure is based on a maximum gate velocity. Gate velocities must be balanced to ensure the atomization of the liquid metal. But if they are too large, the liquid metal will start to erode or washout the gates in the tool. This washout will affect the gate area and therefore the other injection parameters. This erosion also will lead to downtime for die repair. The operating window must be determined based on these trade-offs and limitations. The industry has a process called PQ^2 used to help select values within an operating window of the machine and the requirements for filling the die. An example of this will not be reviewed here. Instead, readers are referenced to NADCA publications for detailed examples of PQ^2 calculations [70]. The important point is the die design must also take into effect the die cast machine. Die cast machines may have different features in the hydraulic injection system, which lead to different injection parameters.

The cold chamber of the die is designed based on a tip size and a chamber length. The target metal pressure for the casting and the PQ^2 process also help the designer determine the tip diameter to be used in the cold chamber. The tip diameter and injection velocity determine the flow rate of metal filling the casting. Again, trade-offs must be considered. A high metal pressure helps reduce porosity found within the casting, but it can also lead to flashing of the die. Flashing of the die occurs when the metal pressure becomes larger than the force holding the die closed, causing die components to separate and liquid metal to fill the areas. If the die is poorly constructed or parameters are selected that create flashing, the die will likely become damaged and require expensive repair and downtime of the die casting machine. The chamber length is determined by the combination of tip size and shot volume. The shot volume is the total amount of metal poured into the chamber. This volume of metal must include the casting, gating, runners, overflows, venting runners, and biscuit size.

The designer must make additional decisions on overflows, venting, and vacuum to ensure high quality within the final casting. Overflows are attached to the casting on the outbound side of the metal flow. They are designed to be in areas of the casting that are the last to fill. Overflows provide a means to ensure there is enough metal in the shot to fully fill the casting and allow the leading edge of metal to fully leave the final casting. This leading edge often collects any remaining die lubricant or impurities within die that need to be flushed out of the casting. These overflows are not part of the final casting. To reduce remelting costs, overflows should be small and only used where necessary. Attached to the overflows are vent runners and venting or vacuum. Venting and vacuum are important as the air within the cavity needs to escape while the die is being filled. If the air does not escape, then it will be entrapped in the casting leading to porosity. A vent allows the air to be pushed out naturally as the metal is injected.

A vacuum system creates a vacuum within the cavity once the slow shot is started to remove air from the cavity within the die. Additionally, the vent/vacuum and runner system must also slow down and cool the metal, so it will solidify by the time it reaches the vent at the edge of the die or vacuum block. If the liquid does not solidify, it will spit out of a vent thereby reducing the metal pressure within the die. The solidification of the metal in the vent system allows the intensification process of the liquid metal to build pressure within the die, thereby feeding voids created during solidification of the casting. Vent area is calculated based on the gate area, gate velocity, and vent velocity. The speed of sound in air controls the vent velocity. Typically values of 8,000 to 10,000 inches/second are used for die casting [15]. Equation 7 [15] shows the calculation for vent area for the tooling.

Equation 7 [15]:

$$V_A = \frac{G_A * G_V}{V_V}$$

Where:
$V_A =$ **vent area (in²)**
$G_A =$ **gate area (in²)**
$G_V =$ **gate velocity (in/sec)**
$V_V =$ **vent veolcity (in/sec)**

The casting is removed from the die with ejector pins. Ejector pins are built into the die and placed on non-critical casting surfaces and on the gate running and venting system to push the solidified casting off the die steel at the end of the cycle. The number, size, and location of these pins impacts how well the casting is ejected from the die. These pins are attached to an ejection plate, which is connected to the ejection system on the die cast machine. Many die casting machines have a hydraulic ejection cylinder that pushes the ejection plate and pins forward. If improperly sized or positioned, the casting could fail to eject from the die, creating downtime issues and tooling repair. If these problems arise, it is not uncommon to upsize or add additional ejector pins to assist with the ejection process. These variables are important to understand and track along the life cycle of a die as they can influence the data collected within the process.

Like ejector pins, the thermal management system of the die casting die involves many up-front decisions. The tool designer must determine the location, size, and number of thermal lines meant to remove heat and maintain the die at a specific temperature. As the liquid metal solidifies, heat is transferred into the die. The goal of a properly designed thermal system is to remove the same amount of heat as introduced with each cycle, thereby allowing the die to be at a consistent temperature. Heat transfer rates for solidification are based on the liquid metal of the casting and the die temperature. Changes in the die temperature, due to improper cooling cycle-to-cycle change the solidification time. The changes in die temperature also change where porosity and shrink voids are found within the casting. The design of the thermal system is a complicated process given the complex shapes of castings and type of thermal management systems. Internal die temperatures can be controlled with cold water, hot water, hot oil, and high-pressure cooling. Manual calculations and zoning of heat removal are recommended through the NADCA publication entitled *Thermal Design & Control of Die Casting Dies* [71]. Often, however, computer simulations are utilized to finalize the design of internal thermal management of the die [1], [15], [71]. An example of a die design with complex thermal lines is seen in Chapter 1's Figure 16. Like ejector pins, changes to this system are unlikely, but still can be performed to address specific quality issues. It is important to have detailed understanding of the variables associated with the thermal management captured initially, as well as with any changes made in the die design.

There are many different decisions that must be made during die design. These decisions represent fixed values at a given time. Cycle-to-cycle, these values do not change. Run-to-run, however, the values could change if the die is modified. When the tool has to change the gating area or add a cooling line, it is fundamentally different in production than before. This could cause significant differences in the data collected from the process. A changed gate area may mean a different flow rate during injection. A cooling line modification will change the die temperature and flow rate of water through the die. These details need to be recorded, so a technician troubleshooting variation seen during production will know if the die has changed or if something else in the system must be investigated. A change in water flow rates may mean an additional line was added, or it could mean a leak is happening in the die assembly. Without this data readily available, the troubleshooting process substantially increases and leads to more equipment downtime. These are important details to understand and therefore need to be recorded. Table 6 details the critical die data.

Table 6: Die design data

Die Design and Build Data Overview			**Total Data**		66	247	468
Item	**Collection Frequency**	**Collection Time**	**Data Volume**	**Number of Instances**	**# of Data Points**		
					Low	**Medium**	**High**
Tip Diameter	Fixed	On-Change	1	1	1	1	1
Chamber Length	Fixed	On-Change	1	1	1	1	1
Gates							
Length	Fixed	On-Change	1	1 to 10	1	5	10
Width	Fixed	On-Change	1	1 to 10	1	5	10
Locations	Fixed	On-Change	1	1 to 10	1	5	10
Pour Volume							
Casting Volume	Fixed	On-Change	1	1	1	1	1
Runner/Gating Volume	Fixed	On-Change	1	1	1	1	1
Venting/Overflow Volume	Fixed	On-Change	1	1	1	1	1
Biscuit Volume	Fixed	On-Change	1	1	1	1	1
Overflows							
Land Width	Fixed	On-Change	1	1 to 10	1	5	10
Land Thickness	Fixed	On-Change	1	1 to 10	1	5	10
Overflow Volume	Fixed	On-Change	1	1 to 10	1	5	10
Overflow Locations	Fixed	On-Change	1	1 to 10	1	5	10
Venting/Vacuum							
Length	Fixed	On-Change	1	1 to 4	1	2	4
Width	Fixed	On-Change	1	1 to 4	1	2	4
Locations	Fixed	On-Change	1	1 to 4	1	2	4
Ejector Pins							
Diameter	Fixed	On-Change	1	10 to 40	10	25	40
Locations	Fixed	On-Change	1	10 to 40	10	25	40
Thermal Lines							
Type (water, jet cool, etc)	Fixed	On-Change	1	10 to 100	10	50	100
Size (diameter)	Fixed	On-Change	1	10 to 100	10	50	100
Locations	Fixed	On-Change	1	10 to 100	10	50	100

<u>DIE CAST EQUIPMENT SETTINGS</u>

Like the die design, programming the die cast equipment requires many initial decisions. These settings control the equipment's timing and performance during the die casting process. There is a wide range of settings that must be programmed that affect the process, such as injection parameters, clamp tonnage, timing of movements, and pressures. Many of these settings can become result measurements collected and stored during the process. This allows a technician to understand if the machine's performance is degrading or if at some point a setting was changed and the machine responded accordingly. This section will focus on the items that are typically programmed through the die cast machine itself. Settings associated with sub-systems like spray systems, ladle systems, and thermal management units will be discussed in this section.

Data such as program number, tool and part identification, and revision are basic data that will allow a user to sort and understand changes that happen through time. This type of data is the first that should be considered for storage when programming a die casting machine. Additionally, equipment identification, typically a machine number, should be stored especially in foundries with multiple die cast cells.

Metal injection is controlled by a hydraulic shot system on a die cast machine. This hydraulic system is built with hydraulic pumps and accumulators, so the liquid metal can be injected in the die in a manner to create quality castings. The injection process contains three main phases: slow shot phase, fast shot phase, and intensification phase.

The slow shot phase is the first phase of the injection process. After the ladle system has poured metal into the chamber, the hydraulics of the shot end are triggered to start to move the chamber tip forward within the chamber. The chamber will be partially full based on the chamber dimensions and the volume of metal required for the casting. The goal is to compress the liquid metal in the chamber until there is no remaining air in the chamber with only metal remaining. This must be done without creating turbulence in the metal. Too fast of a slow shot, will cause the wave of liquid to crest and entrap air in the metal. Too slow will cause the wave to hit the runner system and bounce back to the tip, preventing the gas from exiting the cavity during injection. Much research was completed in the 1990s [72]–[77] to develop a critical slow shot velocity calculation seen in Equation 8 based on the percentage of fill of the chamber initially seen in Equation 9 [70].

Equation 8 [70]:

$$v_{CSS} = c_{cc}\left[\frac{100\% - V_f}{100\%}\right]\sqrt{d_{tip}}$$

Where:
$v_{CSS} =$ **critical slow shot velocity (in/sec)**
$c_{cc} =$ **curve fitted constant (22.8 $\sqrt{\text{in}}$/sec)**
$V_f =$ **volume fraction of shot sleeve as initially filled (%)**
$d_{tip} =$ **shot tip diameter (in)**

Equation 9 [70]:

$$V_f = \left(\frac{V_{ladled}}{0.25 * \pi * {d_{tip}}^2 * l_c}\right) * 100\%$$

Where:

$V_f =$ **volume fraction of shot sleeve as initially filled (%)**
V_{ladled} **= volume of metal ladled into chamber (in^3)**
$d_{tip} =$ **shot tip diameter (in)**
$l_c =$ **length of chamber (in)**

At the chamber full position, often call the start of fast shot, the hydraulic system opens the accumulators and transitions from the slow shot used to ready the metal in the chamber to the fast shot phase. In this phase, the metal is injected from the chamber, through the runners, gates, casting, overflows, and vent runners, until it hits the vents and ideally freezes, which allows back pressure to help with intensification, the final phase. The fast shot velocity is based on the flow rate of metal and the tip diameter as seen in Equation 10 [15].

Equation 10 [15]:

$$v_{fs} = \frac{Q}{0.25 * \pi * {d_{tip}}^2}$$

Where:

$v_{fs} =$ **fast shot velocity (in/sec)**
Q = flow rate (in^3/sec)
$d_{tip} =$ **shot tip diameter (in)**

The calculations for slow shot and fast shot velocities provide targets for the machine to try hitting. These velocities, the transition point, and accelerations are programmed into the die casting machine typically with three data set points. The first set point is the location along the length of the chamber. Typically, this is mapped out as the x-axis on a graph. The second set point is a velocity. Velocity is controlled either and open or close-loop hydraulic system. Open-loop command will adjust the hydraulic valve without a feedback loop. Closed-loop will monitor the pressure and adjust the hydraulic valve accordingly. On some equipment, both types of control can be used throughout the injection cycle. The third data point is a command letting the machine know if the second set point is to be ran under open or closed loop control. Typical die cast machines allow 20 to 30 different set points to be programmed. Figure 24 is an example of the set points and corresponding shot profile from a Buhler Prince die casting machine.

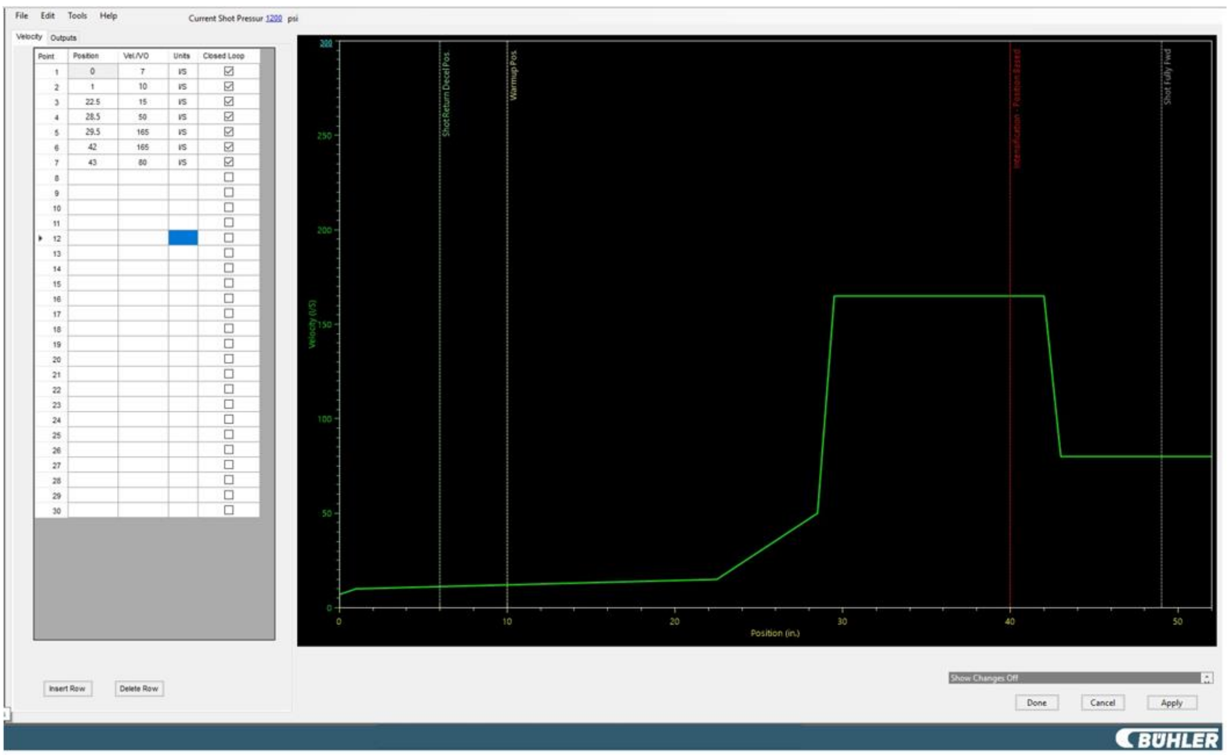


Figure 24: Example programmed set points for shot profile

Although the injection process is important, there are dozens of additional settings for the balance of the die casting cycle that also must be decided and entered. After injection comes the intensification of the casting. Hydraulic pressure level within the shot intensification system must be selected, which combined with the tip diameter and rod diameter of the injection system, determines the final metal pressure seen in the cavity. This pressure must be counteracted by the tie-bar and clamping system holding the die closed. A target tonnage is entered for the tie bars for the system. Die height, or physical space between die casting platens is selected, which directly impacts the tonnage seen on the tie-bars.

Settings must also be selected for the die interactions. In a die casting machine tow areas typically set are hydraulic systems that control the core or slide pulls and the cold-water control. The sequencing of the slide or core pulls must be determined for both the inward motion as the die closes and outward as the die opens. The out motion, or pull sequence, can be important from a casting quality standpoint, especially on large dies with cylinders that require significant hydraulic requirements. Depending on the hydraulic system on the die casting machine and the cylinder size, the system may not be able to pull all slides at once with adequate pressure resulting in stuck slides and downtime for the equipment. Sequencing becomes critical to ensure slides are pulled fully before the hydraulic system starts the next sequence. This affects cycle time of the process, as ejection will not happen until all slides are cleared. Cold water is also controlled by the die casting machine. Water valves are controlled with a solenoid based on settings programmed into the machine. Different quantities of values can be programmed, so decisions include which ones to turn on, how long to delay before start, and how long the cold water will be on.

Numerous factors control the die casting cycle time. Since the die cast machine controls the signals going into secondary systems, timers start the ladle process, the spray system, the vacuum process, and the extraction of the die cast machine. Additionally, the dwell timer must be set to determine the time the die remains closed to allow the liquid metal to solidify. Timer settings also control movements within the machine. Delays for die opening, mid-open switch, ejection, slide pulls, and shot tip retraction all impact overall cycle time of the process and therefore need to be tracked and recorded.

Beyond timers, die cast machines may also contain setup variables for different processes that occur within the cycle. For example, some machines allow for warm-up shots to be defined, so when the machine is not running for a set time period, it forces the warm-up process to help protect the die from cooling down too much before having the hot metal injected. These settings can include a series of time delays and number of cycles to run this warmup cycle. The machine will then automatically flag the extraction robot on the warm-up process and treat the ejected parts as scrap versus good production. Additionally, tip lubrication is used to help with the tip/chamber system. Often this lube is controlled by the die cast machine, so it can be pulsed on the tip or in the chamber as the machine is cycling. These requirements likely vary by machine manufacturer and tip supplier. An example machine may have up to 6 different tip lube pulses, specifying the location within the chamber to fire the pulse, and defining if that is on the forward injection stroke or the return stroke. This example provides 18 different settings that must be selected for pulsing tip lube.

Finally, during the setup process limits or windows are selected to create flags and warnings within the die casting machine. For example, the technician must select the dimensions to calculate the average fast shot speed. Typically, two different length dimensions are entered based on tip position in the chamber. This represents the window associated with the calculated average. This data becomes important when changes are made. If average trends are monitored and a pattern emerges where there is an increase, the cause of the increase becomes important. If a range of where this average is calculated changed, it may be the cause of the trend change versus a maintenance issue within the machine. Limits are also defined for critical process parameters, so if the average fast shot speed is below a threshold, the die cast system will automatically define that casting as scrap and pass that information to the extraction robot to scrap the casting versus allowing it in the production population. Again, operator changes of these limits could directly affect the immediate and historical output, thereby making this information important.

Like die design, die casting equipment settings involve numerous decisions that must be made. These settings are valuable data to the overall die casting process. Since operators and technicians set these, the parameters can be adjusted by those same people. Even though it is discouraged in a production environment, it is not uncommon for operators on different shifts to have preferred settings where they feel the process runs best. It is important to collect the data from these settings to help understand changes that occur in the output variables within the process. For example, is the die hotter due to a broken water line within the die, or did an operator make a change in the water timer reducing the flow? Is the fast shot speed slower because of a valve issue or binding tip in the chamber, or did the operator change the average window or purposefully turn down the speed? The settings provide important details to help rule out possible changes in the process. This makes them important data points to collect and monitor within the die casting system. Table 7 provides a summary of the data totals for the die casting equipment settings as detailed in Table 8 and Table 9.

Table 7: Summary table for die casting equipment settings

Summary Table for Die Cast Equipment Settings		**# of Data Points**		
Table Number	**Name**	**Low**	**Medium**	**High**
Table 8	Die Cast Equipment Settings - Part 1	51	75	99
Table 9	Die Cast Equipment Settings - Part 2	43	54	65
Total		**94**	**129**	**164**

Table 8: Part 1 of die cast equipment settings data

Die Cast Equipment Settings - Part 1			Total Data		51	75	99
Item	**Collection Frequency**	**Collection Time**	**Data Volume**	**Number of Instances**	**# of Data Points**		
					Low	**Medium**	**High**
Identification Data							
Program Identification	Fixed	Each Cycle	1	1	1	1	1
Machine Identification	Fixed	Each Cycle	1	1	1	1	1
Tool Identification	Fixed	Each Cycle	1	1	1	1	1
Part Revision	Fixed	Each Cycle	1	1	1	1	1
Shot Velocity Settings							
Chamber Position	Fixed	Each Cycle	1	4 to 20	4	12	20
Value	Fixed	Each Cycle	1	4 to 20	4	12	20
Control Unit Type	Fixed	Each Cycle	1	4 to 20	4	12	20
Start Testing for End of Shot Velocity	Fixed	Each Cycle	1	1	1	1	1
Target End of Shot Velocity	Fixed	Each Cycle	1	1	1	1	1
Slow Shot Window Start	Fixed	Each Cycle	1	1	1	1	1
Slow Shot Window End	Fixed	Each Cycle	1	1	1	1	1
Slow Shot High/Low Limits	Fixed	Each Cycle	2	1	2	2	2
Fast Shot Window Start	Fixed	Each Cycle	1	1	1	1	1
Fast Shot Window End	Fixed	Each Cycle	1	1	1	1	1
Fast Shot High/Low Limits	Fixed	Each Cycle	2	1	2	2	2
Intensification Pressure Setting	Fixed	Each Cycle	1	1	1	1	1
Target Setting	Fixed	Each Cycle	1	1	1	1	1
High/Low Limits	Fixed	Each Cycle	2	1	2	2	2
Die Height Settings	Fixed	Each Cycle	1	1	1	1	1
Tie Bar Tonnage Settings							
Target Setting	Fixed	Each Cycle	1	4	4	4	4
High/Low Limits	Fixed	Each Cycle	2	4	8	8	8
Other Settings							
Biscuit Size High/Low Limits	Fixed	Each Cycle	2	1	2	2	2
Furnace High/Low Limits	Fixed	Each Cycle	2	1	2	2	2
Intensification Rise Time High/Low Limits	Fixed	Each Cycle	2	1	2	2	2
Start of Fast Shot High/Low Limits	Fixed	Each Cycle	2	1	2	2	2

Table 9: Part 2 of die cast equipment settings data

Die Cast Equipment Settings - Part 2			Total Data		43	54	65
Item	Collection Frequency	Collection Time	Data Volume	Number of Instances	# of Data Points		
					Low	Medium	High
Warmup Process Settings							
Warmup Velocity	Fixed	Each Cycle	1	1	1	1	1
Warmup Downtime	Fixed	Each Cycle	1	3	3	3	3
Warmup Shots Required	Fixed	Each Cycle	1	3	3	3	3
Timer Settings							
Dwell Timer	Fixed	Each Cycle	1	1	1	1	1
Shot Delay Timer	Fixed	Each Cycle	1	1	1	1	1
Tip Lube Timer	Fixed	Each Cycle	1	1	1	1	1
Extractor Cycle Timer	Fixed	Each Cycle	1	1	1	1	1
Spray Cycle Timer	Fixed	Each Cycle	1	1	1	1	1
Ladle Cycle Timer	Fixed	Each Cycle	1	1	1	1	1
Open/Ejection Settings							
Ejection Return Delay Timer	Fixed	Each Cycle	1	1	1	1	1
Ejection Plate Forward Position	Fixed	Each Cycle	1	1	1	1	1
Ejection Plate Return Position	Fixed	Each Cycle	1	1	1	1	1
Ejection Pressure Target	Fixed	Each Cycle	1	1	1	1	1
Fast Die Open Position	Fixed	Each Cycle	1	1	1	1	1
Fast Die Close Position	Fixed	Each Cycle	1	1	1	1	1
Start Core Pull Position	Fixed	Each Cycle	1	1	1	1	1
Start Ejection Position	Fixed	Each Cycle	1	1	1	1	1
Tip Lube Pulse Settings							
Tip Lube Position	Fixed	Each Cycle	1	4 to 8	4	6	8
Tip Lube Direction	Fixed	Each Cycle	1	4 to 8	4	6	8
Slide Pull Settings							
Core In Sequence Number	Fixed	Each Cycle	1	4 to 8	4	6	8
Core Out Sequence Number	Fixed	Each Cycle	1	4 to 8	4	6	8
Cold Water Settings							
Water On/Off	Fixed	Each Cycle	1	2 to 4	2	3	4
Delay Timer	Fixed	Each Cycle	1	2 to 4	2	3	4
Duration Timer	Fixed	Each Cycle	1	2 to 4	2	3	4

INJECTION SYSTEM

With the tool designed and built and the settings to control the machine input into the equipment, the actual process of making the casting occurs. One of the subsystems that generates a significant amount of data is the injection system. This process can be considered the most measured part of a traditional die casting system. Initial data collection systems for injection have existed since the 1960s to help capture the time-series data associated with injection of the liquid aluminum into the die [78]–[80]. Often within the industry, this process receives the most focus. In the 1980s and 1990s, commercial systems became available for die cast companies to purchase and install on their machines [81], [82]. The heightened popularity of injection monitoring drove die casting equipment manufacturers to include these collection systems on machines they sold. As such, literature often focuses on the parameters obtained through this data collection system [17]–[22], [45], [83]–[91] as it is the easiest and most readily available data to collect within the industry.

The oversaturation of use of the injection data should not be mistaken for a lack of importance. The injection process is highly critical to the overall part quality. Therefore, it is important to capture and understand this time-series data that is generated. Most die casting injection processes collect three different time-series data sets as the injection process occurs: velocity, head pressure, and rod pressure. This output data is often stored with two independent variable components to be plotted: time and position. Typically, the position axis is viewed in a shot-profile graph. This position corresponds to the position of the tip within the chamber. Some monitoring systems allow a user to toggle between looking at the graphs from a time-based versus position-based x-axis. If vacuum is used to remove the air from the cavity, a fourth variable of cavity vacuum can be collected and included in these injection profiles. Some monitoring systems also calculate a metal pressure value from the entered tip diameter and hydraulic pressure within the injection cylinder.

The data collected during injection is every 1 to 3 milliseconds. This occurs from the start of the tip moving with the slow shot, through filling the die with fast shot, until the end of the intensification pressure process. This frequency and number of channels creates a large data set of time-series data. An industry focus on the averages within defined windows reduces the power of the data set and will be discussed in more detail later. Figure 25 shows an example shot profile output generated from a Buhler Prince die casting machine. Table 10 contains a summary of the injection process data.

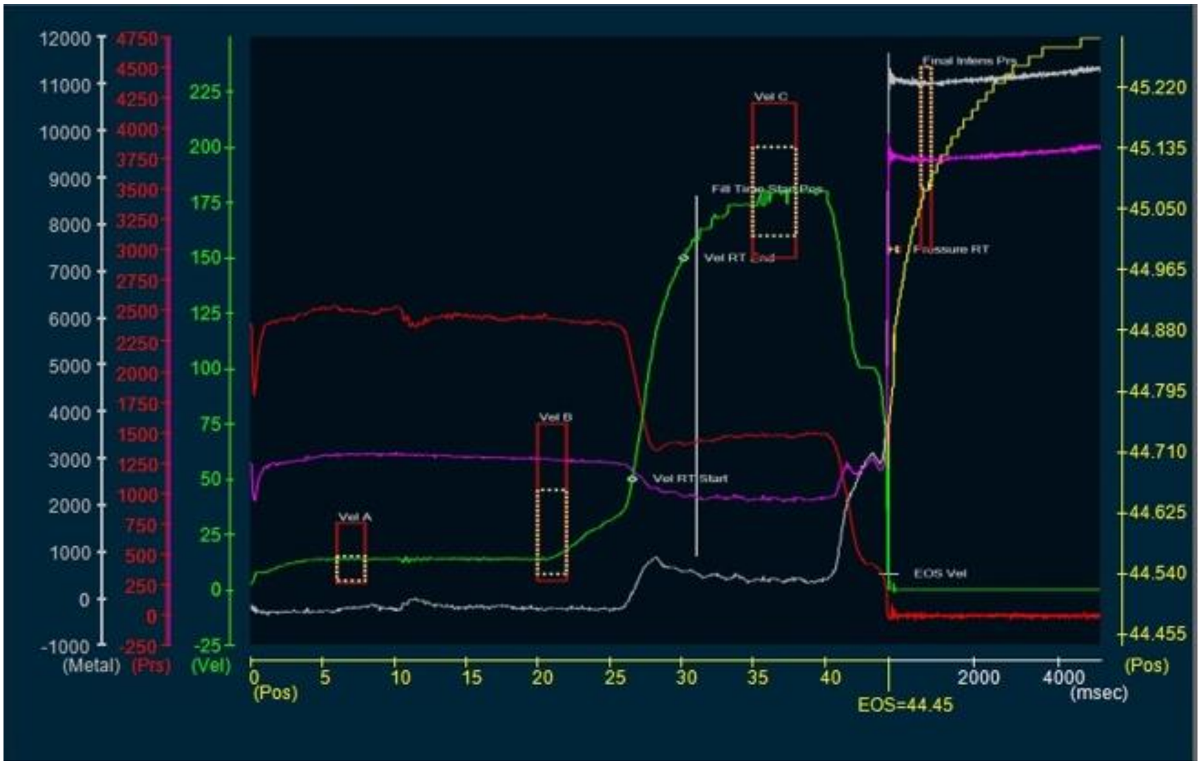


Figure 25: Example shot profile output

Table 10: Injection system data

Die Cast Equipment Settings				**Total Data**	25,641	77,589	156,000
Item	**Collection Frequency**	**Collection Time**	**Data Volume**	**Number of Instances**	**# of Data Points**		
					Low	**Medium**	**High**
Injection Position	333 to 1000 HZ	15 to 30 seconds	Time-Series	1	4,995	14,985	30,000
Injection Velocity	333 to 1000 HZ	15 to 30 seconds	Time-Series	1	4,995	14,985	30,000
Cylinder Rod Pressure	333 to 1000 HZ	15 to 35 seconds	Time-Series	1	4,995	14,985	30,000
Cylinder Head Pressure	333 to 1000 HZ	15 to 35 seconds	Time-Series	1	4,995	14,985	30,000
Vacuum Pressure	333 to 1000 HZ	15 to 35 seconds	Time-Series	1	4,995	14,985	30,000
Squeeze Distance	333 to 1000 HZ	2 to 6 seconds	Time-Series	1	666	2,664	6,000

DIE MOVEMENT AND CLAMPING

The die movement and clamping process produces data that can assist with downtime reduction, part quality, equipment performance, and cycle time consistency. This happens through the opening/closing process of the die and the output measurements of the die cast machine hydraulic interaction with the die.

The clamping tonnage target is selected during the setup process. Die cast machines have strain gages on the tie bars that convert the stretch of the bar from the clamping to a clamp load typically measured in tons. Gages exist on all the tie bars, so four measurements of clamping load per tie bar can be recorded. Significantly offset loads on the tie bar are typically due to poor part placement during the die design process, which may result in problems holding metal within the die or flashing of the tool. These can lead to downtime for tool maintenance or poor part quality. Additionally, upward trends in tie bar loads indicate growth in die steel from increased temperatures. This could indicate thermal management issues in the system or connections. If systemic, it would also indicate poor die cooling designs within the tool.

The hydraulic pressures associated with the die cast machine and die create multiple streams of data that provide insight into the process. Specifically, the slide pull pressure and ejection pressure create time-series data. Pressure sensors can be added to the rod and head side of the cylinders used for slide pulls and the ejection process. As the slide cylinders are triggered to be pulled upon opening or returned before closing, the hydraulic pressure for that movement provides insights into the movement. On the slide in movement, a sudden increase in pressure could indicate buildup of flash in the die or an overflow pocket being stuck. These can be precursors to downtime events for die maintenance. The slide out pressure can provide insight into potential die sticking. Similarly, most ejection systems contain a hydraulic cylinder that moves either a push plate in the machine or connects to the die with an internal ejection plate to push the solidified casting out of the mold. This time-series data provides insight into problems with the casting ejected from the tool, which would result in downtime issues. On the return, an increased pressure could indicate interference with the ejection pin and the tool. This information could be used to plan maintenance, avoid downtime, and ensure the part is ejected without warping.

Data associated with clamping, slide pulls, and ejection provide insight in the interaction of the die cast machine with the die. This information is useful to prevent downtime and die maintenance issues. Table 11 contains details associated with the die movement and clamping process.

Table 11: Die movement and clamping data

Die Movement and Clamping			Total Data		1,004	10,504	36,004
Item	**Collection Frequency**	**Collection Time**	**Data Volume**	**Number of Instances**	**# of Data Points**		
					Low	**Medium**	**High**
Tie Bar Tonnage per Bar	Fixed	Each Cycle	1	4	4	4	4
Slide Pull Pressure - Rod End (full cycle in/out)	10 to 100 HZ	10 to 20 seconds	Time-Series	4 to 8	400	4,500	16,000
Slide Pull Pressure - Cylinder End (full cycle in/out)	10 to 100 HZ	10 to 20 seconds	Time-Series	4 to 8	400	4,500	16,000
Ejection Pressure - Rod End (full cycle in/out)	10 to 100 HZ	10 to 20 seconds	Time-Series	1	100	750	2,000
Ejection Pressure - Cylinder End (full cycle in/out)	10 to 100 HZ	10 to 20 seconds	Time-Series	1	100	750	2,000

EQUIPMENT PERFORMANCE AND ENVIRONMENT

The die cast equipment is the heart of the die casting process. The machine's performance dictates the major systems used for injection, die movement, clamping, and ejection. Failures within the machine's systems result in poor part quality and equipment downtime. The heartbeat of the machine should be monitored to ensure consistent performance for the process. Additionally, environmental factors also influence the machine and the cycle and must be monitored.

Die cast machines often have multiple motors and pumps within the complex hydraulic systems. Monitoring the motors with temperature, vibration, and amp draw sensors provides an understanding of the machine performance and preventative maintenance planning. The same can be done for temperature and vibration for the hydraulic pumps. The hydraulic system on the die cast machine is used throughout the die casting process: injection, die open, slide pulls, ejection, slide in, and closing. As such, time-series data of the entire cycle should be recorded and monitored for the sensors on motors and pumps. This will generate a large volume of data.

Pressure in the hydraulic system and oil temperature is additional data that can be collected within the process. Regeneration rates of these pressure systems may also provide insight into individual components within the machine or where issues may arise. Identification of nitrogen leaks, cylinder failures, and valve performance can be achieved by monitoring the pressure within the hydraulic system. Die cast machines may contain filter systems to help clean the hydraulic oil. Monitoring the pressure drop across the filter can notify maintenance when filters need to be changed. By tracking the time between filter changes, the need for hydraulic oil replacement could be estimated. Oil temperature within the hydraulic system ideally remains consistent. A change in oil temperature could indicate a failure of a cooling system or potentially the need for replacing oil that has reached end of life. Often die cast machine manufacturers will monitor for high temperature limits, but additional tracking of temperature can provide additional insights into the process and equipment.

Along with the machine's performance, environmental factors have an influence on the die casting system. Air temperature and humidity can be monitored. These factors impact cooling within the die casting die through radiation and convection heat loss. Since die cast foundries are typically hot work environments, air make-up units are often installed to improve work environments for employees. However, employees will often still open overhead doors or windows to create additional air flow. Having a localized air temperature within the area of the die is important since temperature could be significantly different machine to machine based on the layout of the plant, movement of air, and employee actions. Additionally, air temperature also impacts many of the plant wide-utility systems. Closed-loop water systems often depend on industrial air-cooled chillers to provide cooling in the recycled water. Air temperatures will impact this final water temperature. Foundries with seasonal temperature changes could find their water systems several degrees cooler in the winter when compared to the summer due to the ambient air temperatures. It is important to monitor this environmental information along with the process data to get a true understanding of the die casting system.

The equipment performance and environmental conditions provide data needed to understand the entire die casting system. The equipment data provides needed insights into the machine's performance and potential needs for preventative and predictive maintenance. Environmental data also dictates temperatures associated with system-wide utilities. A summary of the data from the equipment and environment is seen in Table 12.

Table 12: Equipment performance and environment data

Equipment Performance and Environment			**Total Data**		43,387	785,170	3,067,513
Item	**Collection Frequency**	**Collection Time**	**Data Volume**	**Number of Instances**	**# of Data Points**		
					Low	**Medium**	**High**
Equipment Performance							
Hydraulic Oil Temperature	Fixed	Each Cycle	1	1 to 3	1	2	3
Hydraulic Oil Filter Pressures (in and out)	Fixed	Each Cycle	2	1 to 3	2	4	6
Motor Vibration (3 axial directions)	100 to 1000 HZ	30 to 150 seconds - full cycle	Time-Series	2 pumps x 3 directions	18,000	270,000	900,000
Motor Amp Draw	100 to 1000 HZ	30 to 150 seconds - full cycle	Time-Series	2	6,000	90,000	300,000
Motor Temperature	1 to 5 HZ	30 to 150 seconds - full cycle	Time-Series	2	60	540	1,500
Pump Vibrations (3 axial directions)	100 to 1000 HZ	30 to 150 seconds - full cycle	Time-Series	2 to 4 pumps x 3 directions	18,000	405,000	1,800,000
Pump Temperature	1 to 5 HZ	30 to 150 seconds - full cycle	Time-Series	2 to 4	60	810	3,000
Nitrogen Accumulator Tank Pressure	10 to 100 HZ	30 to 150 seconds - full cycle	Time-Series	1	300	4,500	15,000
Hydraulic Accumulator Tank pressure	10 to 100 HZ	30 to 150 seconds - full cycle	Time-Series	1	300	4,500	15,000
Environment and System Utilities							
Die Cast Machine Water Temp In	1 to 5 HZ	30 to 150 seconds - full cycle	Time-Series	2 to 4	60	810	3,000
Die Cast Machine Water Pressure In	10 to 100 HZ	30 to 150 seconds - full cycle	Time-Series	1	300	4,500	15,000
Die Cast Machine Air Pressure In	10 to 100 HZ	30 to 150 seconds - full cycle	Time-Series	1	300	4,500	15,000
Air Temperature (internal and external to plant)	Fixed	Each Cycle	1	2	2	2	2
Humidity (internal and external to plant)	Fixed	Each Cycle	1	2	2	2	2

METAL AND METAL DELIVERY SYSTEM

Transitioning liquid metal to a net shape part is central to die casting. The metal and metal delivery system therefore provide opportunities to collect and monitor data. The alloy, furnace, and ladle system are the items that must be considered.

The alloy composition and cleanliness of the liquid metal is important to the mechanical properties of the final casting [92]. Spectrometer equipment exists to measure alloy composition in the solid state. Additionally, liquid metal in furnaces experiences mixing due to the filling process. Liquid metal will be pulled from a large holding furnace or smelter and delivered to a smaller holding furnace at a die cast machine. This individual die casting furnace typically ranges from 1,000 to 6,000 pounds, depending the size of the machine and castings being made. Metal is withdrawn from the furnaces one casting weight at a time, thereby lowering the amount each cycle. Eventually, metal must be delivered to refill the die cast furnace. This alloy composition could be slightly different than the previous batch. Because of the testing and holding process, alloy compositions on a per casting basis are near impossible. Until technology advances to provide in-line liquid alloy composition measurements, measurements of alloy compositions will be limited to lot basis and remain variable for individual castings.

The die casting holding furnace does provide opportunities for more complete data framework for die castings compared to the alloy. Foremost is metal temperature. Thermocouples are put into the holding furnace in one or more areas to help monitor the temperature of the liquid metal. Metal temperature will fluctuate through time due to the heating cycle of the furnace. As metal is delivered to the die casting chamber, the temperature can be recorded. If a ceramic filter is used in the furnace to aid with improved metal cleanliness, temperature readings on each side of the filter could provide indication of how plugged the filter is based on the ease of liquid metal flowing through the filter. There are many types of furnaces used in die cast foundries. If electricity is the power source, then data can be collected at set points for elements and the individual power use per element. This data could help predict heating element failures. If gas is used for heating, then gas flow rates can be monitored.

The ladling process to deliver metal from the furnace to the chamber of the die casting machine provides another opportunity to collect and monitor data. Ladling processes can vary significantly. Dosing furnaces exist that pump metal down a trough into the chamber. Robotic ladling systems range from simple 2-axis specialized machines to 6-axis standard robots with a 7^{th} axis ladle system attached. Examples of these different ladling systems are seen in Chapter 1's Figure 3.

The data available varies greatly with different equipment. The settings used within the system and the timing and execution of the process are important data to be collected. Like the die casting machines, ladle systems require technicians to specify the process of filling the ladle, speed of the machine, pour angles, and time in pour, to name some of the standard ones. The goal with these settings and the equipment is to reduce the variation of metal within the ladle bucket. Variation in pour weights impacts the chamber full percentage and thereby the injection parameters. Output data exists within this ladling process as well. Metal level within the furnace can be measured with laser sensors. This feedback can go to the ladling system to aid in movements to the liquid metal surface but also gives an indication of potential timing of the sub-process within the overall die casting process. Additionally, the furnace level information could be routed to a centralized point for all machines across the plant and be used to aid in metal delivery to the die casting furnaces.

The alloy, furnace, and furnace delivery systems provided different types and frequency of data. Alloy information with today's technology is limited to lot sampling. Furnaces provide an opportunity for time-series data. Ladling systems contain settings that should be captured to monitor changes, as well as timing and outputs that provide insight on its overall performance within the die casting process. The data of these systems can be reviewed in Table 13.

Table 13: Metal and metal delivery system data

Metal and Metal Delivery System			**Total Data**		265	1,272	4,029
Item	**Collection Frequency**	**Collection Time**	**Data Volume**	**Number of Instances**	**# of Data Points**		
					Low	**Medium**	**High**
Alloy Composition	Fixed	Per Batch	~ 10 elements	1	Not Included		
Furnace Data							
Metal Temperature in Furnace	Fixed	Each Cycle	1	2	2	2	2
Target Metal Temp Setting	Fixed	Each Cycle	1	1	1	1	1
Metal Level in Furnace	Fixed	Each Cycle	1	1	1	1	1
Actual Element Temperature	Fixed	Each Cycle	1	2 to 4	2	3	4
Furnace Output (on/off)	Fixed	Each Cycle	1	1	1	1	1
Metal Delivery Equipment	*Significant variation will exist in process equipment - below is assuming a robotic ladling system*						
Angle of Fill Setting	Fixed	Each Cycle	1	1	1	1	1
Time in Fill Setting	Fixed	Each Cycle	1	1	1	1	1
Pouring position	Fixed	Each Cycle	1	2 to 6	2	4	6
Pouring Angle Zones	Fixed	Each Cycle	1	2 to 6	2	4	6
Pouring Time Zones	Fixed	Each Cycle	1	2 to 6	2	4	6
Positional Output Value	50 to 200 HZ	5 to 20 seconds	Time-Series	1	250	1,250	4,000

THERMAL DIE MANAGEMENT SYSTEMS

An important part of the die design involves the heat management from the liquid metal into the die steel during casting solidification and then out of the tool to maintain proper die temperature. Within die casting, this thermal management is controlled multiple different ways, including cold water, hot water, high pressure (jet) cooling, and/or hot oil. Die temperature is also controlled by the spray process that is discussed is a separate section.

Cold water, hot water, jet cooling, and hot oil are similar in the goal of thermal management within the die but have different means to that end and are useful in different applications. Typically, hot water, jet cooling, and hot oil units are stand-alone pieces of equipment that are used within the die casting cell. These thermal units contain a means to heat or cool the liquid, pressurize and flow the liquid through cooling channels within the die, control valves to turn flow on and off based on programmed settings, and sensors to verify performance within the unit. Thermal units may contain multiple zones that can be controlled individually. For example, a hot water unit may have two zones, with two unique temperature settings for different time settings. In Chapter 1, photos of a jet cooling and hot water unit are found in Figure 11.

How the liquid media is used creates the different benefits for each of these sub-systems. Cold water cooling is the traditional type of cooling used in die castings. These settings are often controlled within the die cast machine itself to turn on and off valves to allow water to flow into the die, so they are not typically a stand-alone piece of equipment. The settings were discussed in the previous section, but the outputs of flow rates provide additional data that should be collected. Jet cooling pressurizes to hundreds of pounds per square inch (PSI) and is often used in very small cores where traditional cooling is not an option due to size. Jet cooling also provides the ability to extract this heat quickly, which helps with solidification. Hot water and hot oil thermal units are typically identical except for the media used. Hot oil has a longer history within die casting and is capable of heating oil to temperatures of 400 to 650 degrees Fahrenheit. Often hot oil is used to warm the holder block and chamber and to help maintain a constant temperature of the tool steel away from the cavity itself. Hot water units are used in similar ways. Water, however, is only capable of reaching up to 400 degrees Fahrenheit with a pressurized close loop system. Beyond environmental benefits of the water versus oil, the heat transfer capability also differs between the mediums. Water can increase the die temperature much quicker, and these systems can also be used in reverse to pull all the heat out of the die, which provide a benefit to foundries during the setup and change-over of the tools.

Because of the complexity of cooling the die and number of passages, often multiple thermal units are used. It is common for a large tonnage die casting machine to have four to six different hot water units each with two zones, a multi-zoned jet cooling unit, and the use of the cold water system from the die casting machine. A combination of number of units and number of flow zones each creates a rich opportunity for data collection. As with the cold water within the die casting machine settings, the set points selected on this equipment can be collected to monitor process change. Delay timers, flow timers, temperature settings, and pressures are all settings required for the machines to function. Collecting this data provides an understanding of process changes that may occur over time. Additionally, this data provides a baseline for how the equipment performance can be measured. For example, if the goal is a hot water setting of 275 °F, but the temperature sensor within the line is returning a value of 200 °F, a failure is likely occurring in the thermal unit that requires maintenance support. Flow rates monitoring can provide additional insight into the setup of the die. If flow rates are greater than previous runs, a leak within the cooling system is likely happening, which could lead to water within the cavity. If a flow rate is less than normal, possibly a setup issue occurred or passageways inside the die may be blocked. Monitoring flow rates also provides a means to check the equipment, specifically valve performance. If the timer to start flow has not yet occurred, a flow rate of zero is expected. If the flow rate is monitored throughout the cycle, and flow is seen during these downtimes, it can be expected that the valve would need to be fixed.

A large amount of data is created by collecting the settings along with the time-series output data such as temperature, flow rates, and pressures per zone across all the different thermal units. Subtle changes within these flow rates or temperatures could create meaningful changes within the die steel and the steel's ability to remove heat during the solidification process. Data collected within the die possibly highlights these changes. One option is to embed thermocouples within the die steel and then monitor their temperature output throughout the cycle. Ideally, cycle-to-cycle these temperatures should be the same. The trouble with this approach is thermocouple placement. Often only a few spots on the tool can be used to help audit the temperature changes of the die steel due to conflicting use of the space for items like cooling lines, ejector pins, or geometrical part features. These thermal couples are often a distance away from the die steel, so exact temperature readings of the surface of the die steel that the liquid metal contacts need to be inferred. Another approach used is thermal cameras that take surface temperature readings of the die steel. Each image produces hundreds-of-thousands of data points. However, these cameras have their own set of challenges in production. Camera placement to the die steel is often difficult to get the entire die surface, so parts of it are hidden behind other die components. Additionally, the die cast environment with the spray systems can cover the lens in mist, making readings unreliable.

Figure 26 shows one of these thermal imaging cameras inside a case with a shutter, installed in a production die casting foundry. The effects of the die casting environment can be seen within the picture on the equipment. Figure 27 provides example grayscale images captured from this camera system.

Figure 26: Thermal imaging camera in production die cast environment
(photo permission from Mercury Marine)

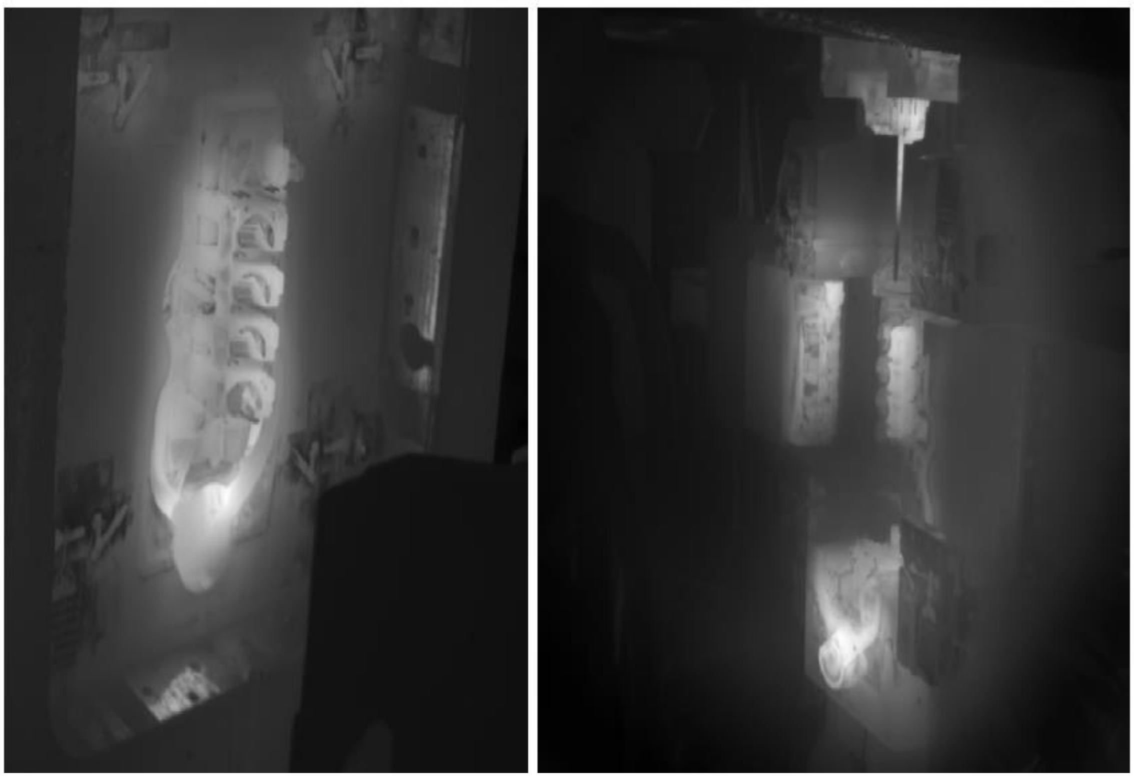


Figure 27: Example thermal image on stationary half (left) and moving half (right)
(photo permission from Mercury Marine)

No solution exists to provide perfect feedback on the die temperature, but these approaches do provide data to help simplify the troubleshooting process when a failure is seen. For example, if a flow rate is low in a zone, the thermal image could pick up that the die is getting hotter in a certain area. By utilizing the cooling line diagrams associated with the die design, the exact line that may not be connected could be identified, eliminating the need for hundreds of water lines to be individually inspected and tested.

The scale of the data generated by the thermal management and thermal monitoring systems varies greatly based on the number of thermal control units used and the technology used by the foundry. This data can be seen in Table 14.

Table 14: Thermal die management data

Thermal Die Management Systems			Total Data		1,806	85,370	513,100
Item	Collection Frequency	Collection Time	Data Volume	Number of Instances	# of Data Points		
					Low	Medium	High
Cold Water							
Water Temperature In	Fixed	Each Cycle	1	1 to 4	1	2	4
Water Temperature Out	5 to 10 HZ	30 to 150 seconds - full cycle	Time-Series	1 to 4	150	1,440	6,000
Flow Rate	10 to 100 HZ	30 to 150 seconds - full cycle	Time-Series	1 to 4	300	9,000	60,000
Hot Water or Hot Oil Units							
Set Points (temp output, time delay, time length, program name/rev)	Fixed	Each Cycle	4	1 to 20 zones	4	40	80
Temperature in to Die	5 to 10 HZ	30 to 150 seconds - full cycle	Time-Series	1 to 20 zones	150	7,200	30,000
Temperature out of Die	5 to 10 HZ	30 to 150 seconds - full cycle	Time-Series	1 to 20 zones	150	7,200	30,000
Flow Rate	10 to 100 HZ	30 to 150 seconds - full cycle	Time-Series	1 to 20 zones	300	45,000	300,000
High Pressure Jet Cooling							
Set Points (temp output, time delay, time length, program name/rev)	Fixed	Each Cycle	4	1 to 4 zones	1	8	16
Temperature in to Die	5 to 10 HZ	30 to 150 seconds - full cycle	Time-Series	1 to 4 zones	150	1,440	6,000
Temperature out of Die	5 to 10 HZ	30 to 150 seconds - full cycle	Time-Series	1 to 4 zones	150	1,440	6,000
Pressure	10 to 100 HZ	30 to 150 seconds - full cycle	Time-Series	1 to 4 zones	300	9,000	60,000
Die Temperature Sensors							
Thermal Couple in Tool	5 to 10 HZ	30 to 150 seconds - full cycle	Time-Series	1 to 10	150	3,600	15,000
Thermal Images of Die Halves	Fixed	Each Cycle	480 x 640 image	2 Pre Spray + 2 Post Spray	Not included in totals Example: 4 images would be ~1.23 million data points		

SPRAY SYSTEM

The thermal management sub-systems integrated with the die are ideally designed to remove all required heat from the die each cycle. The spray application of die lubrication, or die lube, is intended to cover the die halves with the lubrication to ease the ejection of the final casting from the die. Die lube is often mixed in high ratios with water (100 to 1). Industrial practice, however, often uses the die lube spray system to assist with cooling of the die along with the thermal management systems. Given the high water ratio and the control of the application process, it is convenient to spray extra die lube on areas that are extremely hot or that may not have good internal cooling methods to achieve the appropriate temperature on the die. Both the application of lube and the use of cooling is important to the process. Data collected from the spray system can ensure a consistent process and therefore a consistent die temperature needed for part quality.

There are multiple spray systems used within the industry. The two most common systems are a gantry style 2-axis spray system or a spray system mounted on a 6-axis robot as seen in Chapter 1's Figure 12. The gantry style spray system typically has a large manifold that contains dozens, if not hundreds, of nozzles used to spray lube at specific areas of the die. When the die opens, this system travels in the die open position until it is between the die halves and then will move the manifold down between the halves. The nozzles are often tied to a bank with control valves. Multiple banks exist, which allow a technician to program when each bank turns on and off. After the application of the die lube, a "blow off" cycle is usually completed, where high-pressure air is blasted on the die to remove any remaining liquid out of the die cavity. Liquid in the die cavity creates porosity once it vaporizes when the injected metal reaches it. The 6-axis robot follows a similar process but has more degrees of freedom from a programming perspective. As such, the spray systems are often less complicated. Manifolds with dozens of small nozzles still can be used, which is especially useful on dedicated machines running one part number, but may not be preferred given the flexibility gained with the robot. For machines that run multiple part designs, robotic systems have two to eight nozzles in a generic setup. These nozzles are individually controlled and can be aimed directly at critical areas in the die with the robot. Again, these systems also blow air at the end of the cycle, helping remove any liquid from the cavity.

The settings for the different systems vary, but collecting this data allows troubleshooting and understanding of changes made. There are timers associated with a nozzle or bank of nozzles that turn a valve on and off. In many cases, the same nozzle may turn on and off several times throughout the spray cycle. The same is true for the air blow-off at the end of the cycle. Modifications to these timers directly change the amount of lube or air sprayed onto the die. Different levels of lube result in different levels of cooling. In the end, the die temperature with the change will not match the die temperature from prior to the change. Since spray is often used to help cool the die, especially for areas where it may be hard to get internal cooling, it is important to have a consistent process and track the settings used.

Sensors can be added to collect process data associated with this sub-system. Die lube flow rates and pressures are time-series data that can be tracked through the spray cycle. This information can help identify several potential failures. Flow rates or pressures higher than typical may represent a leak in the system or a valve failure, resulting in die lube not being sprayed as intended in the cavity. This could also create puddling of die lube in the die, which could result in porosity in the casting. Low flow rates or pressures could represent nozzles being plugged or valves not triggering. In this case, not enough die lube may be reaching the cavity, thereby affecting cooling and ejection of the casting. Because there are many different nozzles and banks of nozzles toggling on and off throughout the process, flow rates are different when a bank of four nozzles is spraying versus two banks of eight nozzles. Therefore, overall averages of gallons used per cycle may identify system failures but will not specify the exact area of trouble like time-series data can.

The temperature of the die lube is another process parameter that is critical to collet because of its impact on cooling the die. Die lube is often centralized within the plant and is based on water systems used for mixing with the lube. As such, depending on the plant’s utility system, there is a potential for changes in temperature throughout the day or year. Temperatures at night or in winter could create a die lube temperatures that are significantly colder than during the day or in summer. This temperature difference affects the amount of heat removed from the die and should be tracked.

At the end of the die spray process is the blow-off cycle. Collecting time-series data on the air pressure used helps validate a consistent blow-off. Changes from this normal process likely indicate valve failures or leaks. Again, this value changes based on number of nozzles activated at once, so time-series data allows for a better understanding of where these potential failures are within the system. A summary of this spray data can be found in Table 15.

Table 15: Spray system data

Spray Systems			**Total Data**		410	3,622	9,562
Item	**Collection Frequency**	**Collection Time**	**Data Volume**	**Number of Instances**	**# of Data Points**		
					Low	**Medium**	**High**
Significant variation will exist in process equipment - below is assuming a 6-axis robotic spray system							
Settings							
On-Off Timers Settings for each spray zone	Fixed	Each Cycle	4 to 20	2 to 8	8	60	160
Outputs							
Tracking of On-Off for each zone	5 to 10 HZ	10 to 30 seconds	Time-Series	2 to 8	100	560	2,400
Flow Volume of Die Lube into System	10 to 100 HZ	10 to 30 seconds	Time-Series	1	100	1,200	3,000
Air Pressure into System (through blowoff)	10 to 100 HZ	20 to 40 seconds	Time-Series	1	200	1,800	4,000
System Data							
Temperature of Die Lube	Fixed	Each Cycle	1	1	1	1	1
Ratio of Die Lube	Fixed	Each Cycle	1	1	1	1	1

CYCLE TIME ANALYSIS

The controls and feedback systems within a die cast machine are numerous given the needed movements associated with ladling, injection, opening, ejection, spraying, and closing. Multiple switches and sensors exist on the machine to ensure it performs as expected. These devices communicate to the PLC. When combined with the settings loaded into the die cast machine as previously discussed, the system functions execute based on the commands of the PLC and feedback from the sensors.

Tracking the sensor signals provides valuable cycle time data regarding the die casting process. Consistency is critical with die casting, so the same level of heat transfer occurs to keep the same die temperature cycle-to-cycle. Consistency in equipment movements is also an important and can be used to identify wear or pending failure of a component. Additionally, having a detailed cycle time analysis of all movements within the machine also provides quick information on what part of the process changed when an overall cycle time is found to be different. Since the part-to-part cycle time is comprised of so many different systems and operations, investigating a cycle time change to find the cause could take hours if the data is not already available. Because these sensors are critical to the machine, the tags within the PLC exist. Some die casting equipment manufacturers provide this cycle time analysis in the data the machines display. For equipment that does not, a start point of the process needs to be determined, and a time counter within the PLC is then started. A collection system within the PLC must be put into place, so when every switch within the process is hit, the time counter value is stored with the associated switch.

To better understand what some of this data is and how it can be used, a theoretical example is provided. A die casting company starts a timer in the PLC after the ejection complete signal is received, marking the start of the next cycle within the die casting machine. When the switch that records the ejection return is made, that time value is stored. Variation in this time could indicate ejection pin issues with the die. The start and stop of the spray cycle are recorded. Changes in these values influence the overall cycle time and could indicate a change to the spray process. The tip is signaled to return to the full back position after spray. The switch tripped at that full back position is recorded. The time from when the call is initiated to when the switch is hit can be calculated to show the travel time of the tip within the chamber. As a tip wears or binds within the chamber, the time to hit the return switch will increase. The end of spray also starts the closing process of the die and then the ladling process. Switches as the die starts to move are hit, letting the machine know when to close with higher speed versus higher force. Inconsistencies here could mean die linkage needs maintenance or the closing cylinder may be leaking. A lower metal level within the furnace may cause the ladle robot to take more time to fill the bucket properly. As a result, the time the signal comes back saying the robot is ready to pour could cause cycle time variation. This cycle time information continues in more detail than described here and throughout the entire process. Each die cast machine manufacturer has different types and placement of switches. Table 16 provides a summary of Table 17 and Table 18. The latter two tables represent a generic example of data associated with equipment movements that would happen regardless of equipment brand.

Table 16: Summary table for cycle time analysis data

Cycle Time Analysis Summary Table		# of Data Points		
Table Number	**Name**	**Low**	**Medium**	**High**
Table 17	Cycle Time Analysis - Part 1	19	23	27
Table 18	Cycle Time Analysis - Part 2	32	40	48
Total		**51**	**63**	**75**

Table 17: Part 1 of cycle time analysis data

Cycle Time Analysis - Part 1			Total Data		19	23	27
Item	**Collection Frequency**	**Collection Time**	**Data Volume**	**Number of Instances**	**# of Data Points**		
					Low	**Medium**	**High**
Cycle Start Point	Fixed	Each Cycle	1	1	1	1	1
Time to Safety Doors Closing	Fixed	Each Cycle	1	1	1	1	1
Time to Spray Start	Fixed	Each Cycle	1	1	1	1	1
Time to Spray Complete	Fixed	Each Cycle	1	1	1	1	1
Time to Insert Start Signal	Fixed	Each Cycle	1	1	1	1	1
Time to Insert Complete Signal	Fixed	Each Cycle	1	1	1	1	1
Time to Shot Rod Return Limit Switch	Fixed	Each Cycle	1	1	1	1	1
Time to Slide-In Signal	Fixed	Each Cycle	1	4 to 8	4	6	8
Time to Slide-In Limit Switch	Fixed	Each Cycle	1	4 to 8	4	6	8
Time to Die Close Output	Fixed	Each Cycle	1	1	1	1	1
Time to Die Close Mid-Switch	Fixed	Each Cycle	1	1	1	1	1
Time to Die Close Deceleration Limit Switch	Fixed	Each Cycle	1	1	1	1	1
Time to Die Lock Output/Tonnage Confirmed	Fixed	Each Cycle	1	1	1	1	1

Table 18: Part 2 of cycle time analysis data

Cycle Time Analysis - Part 2			Total Data		32	40	48
Item	Collection Frequency	Collection Time	Data Volume	Number of Instances	# of Data Points		
					Low	Medium	High
Time to Pour Signal	Fixed	Each Cycle	1	1	1	1	1
Time to Pour Complete	Fixed	Each Cycle	1	1	1	1	1
Time to Vacuum Start Signal	Fixed	Each Cycle	1	1	1	1	1
Time to Shot Start	Fixed	Each Cycle	1	1	1	1	1
Time to Water Valve On Signal	Fixed	Each Cycle	1	2 to 4	2	3	4
Time to other Thermal Management Unit Signals	Fixed	Each Cycle	1	4 to 8	4	6	8
Time to Dwell Time Start Signal	Fixed	Each Cycle	1	1	1	1	1
Time to Intensifier Fire Signal	Fixed	Each Cycle	1	1	1	1	1
Time to Dwell Time End Signal	Fixed	Each Cycle	1	1	1	1	1
Time to Die Open Output Signal	Fixed	Each Cycle	1	1	1	1	1
Time to Die Open Mid-Way Switch	Fixed	Each Cycle	1	1	1	1	1
Time to Die Open Deceleration Switch	Fixed	Each Cycle	1	1	1	1	1
Time to Die Fully Open Limit Switch	Fixed	Each Cycle	1	1	1	1	1
Time to Water Valve Off Signal	Fixed	Each Cycle	1	2 to 4	2	3	4
Time to Slide-Out Signal	Fixed	Each Cycle	1	4 to 8	4	6	8
Time to Slide-Out Limit Switch	Fixed	Each Cycle	1	4 to 8	4	6	8
Time to Ejection Signal	Fixed	Each Cycle	1	1	1	1	1
Time to Ejection Forward Complete Switch	Fixed	Each Cycle	1	1	1	1	1
Time to Part Extraction Verification	Fixed	Each Cycle	1	1	1	1	1
Time to Ejection Return Signal	Fixed	Each Cycle	1	1	1	1	1
Time to Ejection Return Complete Switch	Fixed	Each Cycle	1	1	1	1	1

EXTRACTION CELL SYSTEM

The last section of die casting data focuses on the process immediately after the casting is created. The extraction cell is important to the die casting process beyond just removing the casting from the die.

Any breakdown in this process directly relates to the overall cycle time of the die casting machine. A failure in any piece of the extraction cell equipment means the die casting process stops. Beyond the financial impact of a lack of production, the thermal balance of the die also changes. This results in wasted warm-up cycles and reduced die life. Because of the impact to the die casting cycle, a focus on the data collection of the extraction cell is a natural progression in the overall cell analysis. Figure 28 is a photo of a production die cast cell at Mercury Marine. This photo of the extraction cell corresponds to the CAD layout of the cell shown previously in Figure 1.

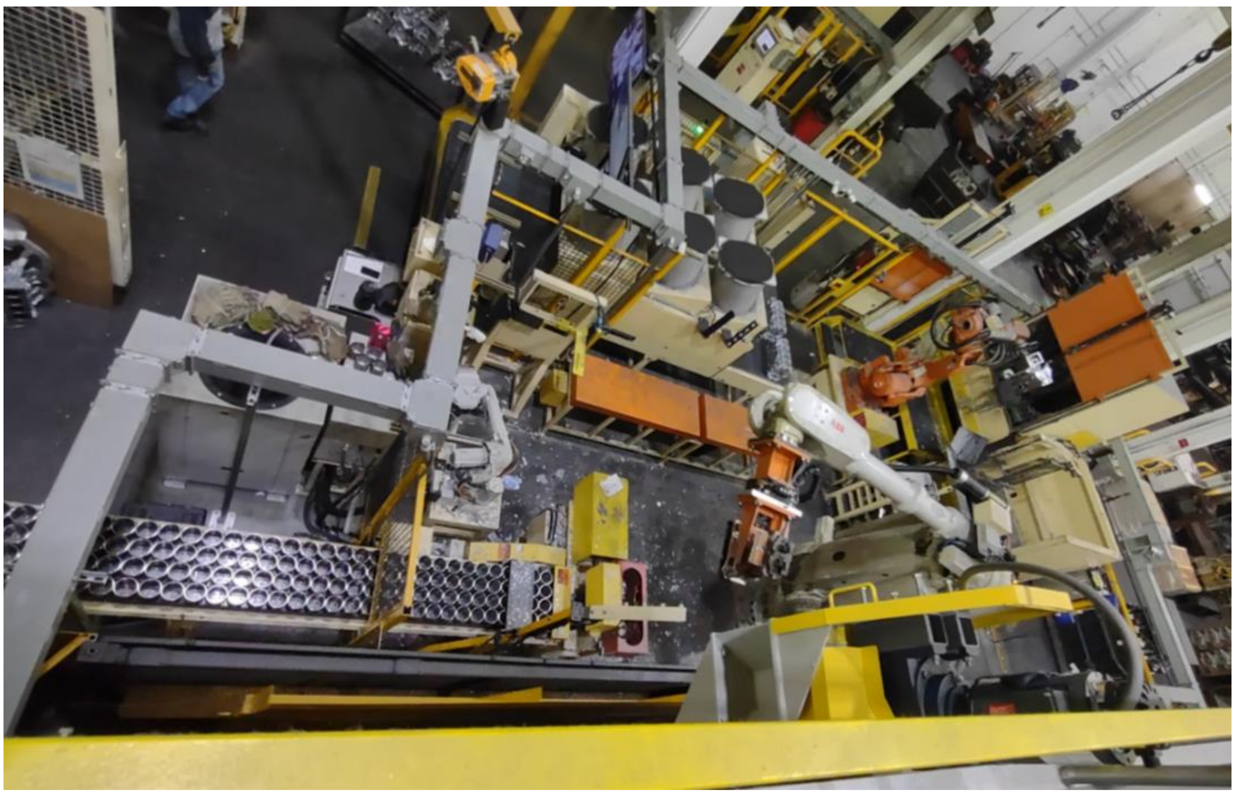


Figure 28: Example extraction cell with multiple robots and post-die cast processing (photo permission from Mercury Marine)

The scale of extraction cells varies greatly based on different foundries and castings produced. This section assumes robotic automation within the extraction cell. The simplest cell would be a robot extracting the casting and placing it on a conveyor to an operator. A highly complex extraction cell could include multiple robots working together to assist with insert placement (ie. liners for an engine block), overflow detection, part serialization, removal of the venting system, sawing of the gating system, and trimming the casting with a trim press. The data generated in this section will be discussed in more generic terms, since collecting data from a saw unit or trim press has some similarities but many differences. The data collected in the extraction cell focuses on part quality and equipment performance.

The extraction cell often completes several quality-related checks on both the casting and the process. Overflow detection is often completed to verify no pockets are left in the die that could cause damage. These devices range from simplistic proximity laser sensors to 3D profiling cameras. When liners are put into engine blocks, often the extraction cell contains sensors to ensure all the liners made it into the casting and did not fall out of the robot or die during the loading process. This validation data should be saved for each casting serial number. Part quality can also be indirectly identified by some of the processes. If the sawing of the gating system sees a significant spike in amp draw, it could indicate an oxide in the casting gating. A trim press that requires significantly larger force to trim the casting could be an indication of a warp or dimensional issue on the casting. If a casting is laser marked or pin stamped with a matrix code, a barcode reader will often validate its readability. This generates data such as contrast ratios, unused error rates, barcode grades, and modulation. This barcode information is important to the casting as its readability in future operations is paramount to the success of connecting results to input parameters, which is needed to apply advanced analytics such as supervised machine learning.

Equipment performance data is also important to prevent downtime within the die casting cell. Saw blades wear out, and the amp draw from the motor or vibration on the spindle will be a precursor to this failure. Motors and pumps exist on saws, presses, and shear units. Time-series data on vibration, temperature, pressures, and amp draw provides insight into potential changes. In some cells, there may be a quench tank. Sensors can tell how full the quench tank is to avoid running out as well as the temperature of the water to see if chiller units are functioning as needed.

The data possibilities in the extraction cell are numerous and have a direct connection to the casting. A thorough review of the process and equipment must be performed to assist with determining critical process data that should be collected. Table 20 and Table 21 provide an overview of data for some of the key pieces of equipment found in common extraction cells and is summarized in Table 19.

Table 19: Summary table for extraction cell system

Extraction Cell System Summary Table		# of Data Points		
Table Number	**Name**	**Low**	**Medium**	**High**
Table 20	Extraction Cell System - Part 1	4,276	53,262	170,320
Table 21	Extraction Cell System - Part 2	3,130	31,084	93,178
Total		**7,406**	**84,346**	**263,498**

Table 20: Part 1 of extraction cell and post die cast process data

Extraction Cell System - Part 1			Total Data		4,276	53,262	170,320
Item	**Collection Frequency**	**Collection Time**	**Data Volume**	**Number of Instances**	**# of Data Points**		
					Low	**Medium**	**High**
Significant variation will exist in process equipment - cells will be comprised of different levels of equipment							
Robots							
Program Number/Rev	Fixed	Each Cycle	1	1	1	1	1
Cycle Time Analysis of all movements	Fixed	Each Cycle	10 to 100	1	10	50	100
Part Detection							
Extraction Sensing	Fixed	Each Cycle	1	1	1	1	1
Overflow Detection Images	Fixed	Each Cycle	480 x 640 image	1 to 4	Not included in totals Example: 2 images would be 614,400 data points		
Overflow Detection Lasers (pass/fail)	Fixed	Each Cycle	1	1 to 10	1	5	10
Hydraulic Shears							
Start/Stop Time Signals	Fixed	Each Cycle	1	1 to 6	1	3	6
Hydraulic Pressure	10 to 100 HZ	5 to 20 seconds (overall)	Time-Series	1	50	625	2,000
Saw							
Settings (RPM, Feed)	Fixed	Each Cycle	2	1	2	2	2
RPM Output	10 to 100 HZ	10 to 40 seconds	Time-Series	1	100	1,250	4,000
Feed Output	10 to 100 HZ	10 to 40 seconds	Time-Series	1	100	1,250	4,000
Amp Draw Output	100 to 1000 HZ	10 to 40 seconds	Time-Series	1	1,000	12,500	40,000
Spindle Vibration (3 dimensions)	100 to 1000 HZ	10 to 40 seconds	Time-Series	1 spindle x 3 directions	3,000	37,500	120,000
Spindle Temperature	1 to 5 HZ	10 to 40 seconds	Time-Series	1	10	75	200

Table 21: Part 2 of extraction cell and post die cast process data

Extraction Cell System - Part 2			Total Data		3,130	31,084	93,178
Item	**Collection Frequency**	**Collection Time**	**Data Volume**	**Number of Instances**	**# of Data Points**		
					Low	**Medium**	**High**
Significant variation will exist in process equipment - cells will be comprised of different levels of equipment							
Part Marking/Serialization							
Part Marking Start/Stop	Fixed	Each Cycle	2	1	2	2	2
Serial Number Marked	Fixed	Each Cycle	1	1	1	1	1
Serial Number Verified (reader)	Fixed	Each Cycle	1	1	1	1	1
Barcode Reader Verification (contrast, unused error correction, axial non-uniformity, grid non-uniformity, modulation, others)	Fixed	Each Cycle	1	5	5	5	5
Barcode Reader Image of Barcode	Fixed	Each Cycle	480 x 640 image	1	Not included in totals Example: image would be 307,200 data points		
Trim Press							
Program Number/Rev	Fixed	Each Cycle	1	1	1	1	1
Cycle Time Analysis of signals and switches	Fixed	Each Cycle	4 to 12	1	4	8	12
Hydraulic Pressure during cycle	10 to 100 HZ	10 to 30 seconds	Time-Series	1	100	1,000	3,000
Temperature Sensor for casting	Fixed	Each Cycle	1	1	1	1	1
Hydraulic Pump Vibration (3 dimensions)	100 to 1000 HZ	10 to 30 seconds	Time-Series	1 pump x 3 directions	3,000	30,000	90,000
Hydraulic Pump Temperature	1 to 5 HZ	10 to 30 seconds	Time-Series	1	10	60	150
Quench Tank							
Water Temperature	Fixed	Each Cycle	1	1	1	1	1
Water Level	Fixed	Each Cycle	1	1	1	1	1
Quality Checks							
Insert Verification (in robot gripper)	Fixed	Each Cycle	1	1	1	1	1
Insert Verification (in casting on extraction)	Fixed	Each Cycle	1	1	1	1	1
Casting Temperature	Fixed	Each Cycle	1	1	1	1	1

SUMMARY AND CONCLUSIONS

The scale of data that can be generated during each die cast cycle is multiple orders of magnitude larger than what industry has assumed to date. The issue within the die industry is the focus is on the injection process [23]. Even though this is time-series data, the industry takes these tens of thousands of data points and creates 20 to 30 descriptive statistics that are tracked [2]. The longest list of published parameters for high pressure die casting was completed based on a European study in 2016. The MUSIC Consortium published 75 parameters that they identified to collect and analyze with machine learning algorithms [86]. This is significantly larger than the three to five inputs that were used to apply machine learning to die casting in recently published papers [17] [18] [21] [22].

When taking a systems engineering approach to die casting, the data framework is not three to five parameters. Nor is it 20 or even 75. The data review completed within this work shows, on the low end, about 80,000 data points and on the high end, more than 4 million data points that can be generated and saved for each die casting cycle. Additionally, this does not include the potential for another 2 million data points from image files that could also be collected. Table 22 shows a summary from all the data framework groups with the low, medium, and high estimates and totals.

A vast majority of this data can be contributed to time-series data. A typical approach may be to collect this data and then summarize it with a descriptive statistic, like an average, for a range of the data. This greatly reduces the amount of data, turning thousands of data points into one. Statistical summaries lose the data context through the calculations. Without the context, the true process will not be known. This will be discussed in more detail with two case studies found in Chapter 5.

Table 22: Data framework group summary table

Die Cast Data Framework Summary				
Data Framework Groups	**Estimated Number of Data Points**			**Images or Others Not Included**
	Low	**Medium**	**High**	
Die Design and Build	66	247	468	
Die Cast Equipment Settings	94	129	164	
Injection System	25,641	77,589	156,000	
Die Movement and Clamping	1,004	10,504	36,004	
Equipment Performance and Environment	43,387	785,170	3,067,513	
Metal and Metal Delivery System	265	1,272	4,029	
Thermal Die Management Systems	1,806	85,370	513,100	1,228,800
Spray System	410	3,622	9,562	
Cycle Time Analysis	51	63	75	
Extraction Cell	7,406	84,346	263,498	921,600
Total	80,130	1,048,312	4,050,413	2,150,400

The optimization and data collection approached used to date have three significant flaws. First, the focus is on quality optimization only and fails to look at any other optimization aspect of the die casting process. With 68% utilization [4], improved machine performance can be a huge financial benefit to the industry. The second flaw is the processes is dynamic. By not tracking any of the settings or inputs of how equipment is programmed, it is unknown if the output changes are driven by humans making modifications to the process or equipment starting to fail. Finally, the scale of parameters used in previous optimization publications is inadequate. Because die casting is such a complex system, it is improbable that representing the system with three to five parameters will allow perfect predictions with machine learning. It should not be a surprise that perfect predictions have not yet happened in the work done to date.

The challenge the die casting industry really faces is being able to collect, store, and analyze this scale of data. When many foundries fail to monitor even 20 variables, hundreds-of-thousands will be viewed as an insurmountable task. Machine learning will help with the analyzation of this volume of data. The thought process on how machine learning is applied must fundamentally change, at least initially. Instead of focusing on quality predictions, machine learning should be utilized to help provide feedback and control to the complex die casting system. The use of machine learning for anomaly detection to help understand when the process changes is incredibly valuable to the industry. Additionally, the volume of equipment performance time-series data available in die casting means the use of machine learning for predictive maintenance could also be a meaningful win for the industry to reduce equipment downtime. Some of the complexity of die casting is removed when one uses machine learning to analyze the feedback data to inform humans involved when something has changed.

Before this work gets into the machine learning case studies within die casting, it is important to further understand challenges with machine learning in die casting. These challenges are highlighted in Chapters 3 and 4. In Chapter 3, the random nature of casting defects is discussed. Chapter 4 will review human misclassification of data and the critical error threshold that creates meaningful challenges with supervised machine learning approaches.

Chapter 3: Stochastic Nature of Porosity Defects[1]

Chapter 3 provides an in-depth review of the stochastic nature of porosity defects in die casting. Random defect formation is a critical topic to comprehend to understand machine learning applications in production die cast environments. An experimental study was performed where key process parameters were held constant through a production run. Castings were then inspected with radiographic equipment. These X-rays showed meaningful variation of porosity formation within the production run. The process used to produce the best-of-the-best and worst-of-the-worst parts was statistically identical.

INTRODUCTION

Porosity is considered a defect, and it is inherent to aluminum high pressure die casting (HPDC) due to the nature of metal solidification [94]. Research completed by the North American Die Casting Association (NADCA) shows that porosity concerns are one of the leading contributors to scrap costs and the biggest problem within die casting [95]. Approximately 30% of the industry has identified addressing porosity as a top concern [96]. There is high motivation to improve porosity scrap given the die casting industry has $8 billion annual sales in North America [3] and suffers from an 8% median scrap rate [4]. Porosity defects can be described by the primary cause: shrink porosity and gas porosity and classified by the size: micro porosity and macro porosity [97], [98]. The cause and size provide one high-level classification system for porosity, although more detailed defect classification systems exist [95], [99], [100].

Porosity is well documented within the HPDC industry [13], [14], [94], [95], [97]. Shrink porosity is caused by volumetric contraction of metals during solidification. Shrink porosity is often

[1] This chapter is an edited version of a paper published in International Journal of Metalcasting, included with permission. D. Blondheim, Jr. and A. Monroe, "Macro Porosity Formation - A Study in High Pressure Die Casting," *International Journal of Metalcasting*, 2021, doi: 10.1007/s40962-021-00602-x [93]

irregular, with ragged shapes. Gas porosity occurs when a gas concentration within the liquid metal is higher than the solubility of the metal. Gas porosity is typically spherical in shape. The mechanisms causing these two types of internal voids can also interact during the solidification of a casting. This interaction is termed gas assisted shrink. This porosity has shape characteristics of both shrink and gas porosity.

The size of porosity is another distinguishing feature of the porosity formed in castings. Shrink porosity and gas porosity can form in macro and micro size voids. Micro porosity is often discussed as voids that form within the mushy interdendritic liquid [101]. Macro porosity is typically discussed within the industry as visible porosity that is compared to quality specifications. Macro porosity may cause rejects within castings with a radiographic (X-ray) inspection or visual inspection after secondary processing, such as machining [12]. A clear definition between micro and macro porosity is lacking and will be discussed.

Micro porosity has been well researched. The stochastic or random nature of micro porosity formation was studied by Atwood [101]. Lee *et al.* [98] reviewed five different models involving complex simulation of micro porosity and microstructure with random nucleation. Several groups of authors have studied micro porosity formation and distribution in castings with micro computed tomography (CT) equipment [102]–[104]. Cao *et al.* [105] and Niu *et al.* [106] studied porosity formation on the micro scale with vacuum-assisted high pressure die casting. In addition to vacuum, Niu *et al.* [106] work also studied other injection parameters to see impact on the mechanical properties with micro porosity. Zhang et al [107] studied micro porosity formation and mechanical properties of both entrapped air and shrinkage. All this work has provided useful insights into how micro porosity forms and the impact it has.

Research on macro porosity is heavily tied to publications on simulations, and it is generally not studied mechanistically like micro porosity. Overall, there is a correlation between porosity results from simulations and where it is found in a casting [108]. Macro porosity can be formed by both shrink and gas porosity. Shrink porosity can be reduced by improving feeding paths in both casting and tool design.

Gas porosity is affected by the pour rate, slow shot acceleration, venting, and die lubrication improvements [95]. Since these features of die and process design are well known, there has not been much research on the formation of macro porosity in production environments. This is an oversight because porosity is still a major cause of scrap [4]. Understanding the formation of macro porosity can create a better understanding of inspection and classification of casting defects.

The goal of this work is three-fold. First, a definition will be provided to help distinguish the difference between micro porosity and macro porosity. This definition will be based on an industry perspective within HPDC. Second, the stochastic or random nature of macro porosity will be reviewed with an industrial case study. Simulation software can predict a zone where porosity is likely to form, but it falls short of showing the random formation of voids produced within that zone. Finally, a review will highlight that from first principles macro porosity should be expected to form randomly in unfed liquid pockets. The end results from this work should challenge the accepted analysis of macro porosity in HPDC.

BACKGROUND ON MACRO POROSITY

The transition dimension between micro porosity and macro porosity has ebbed and flowed between different authors. Huang and Conley [109] said micro porosity was from 500 μm to 1 mm and macro porosity was from 1 mm to 10 mm. Lee *et al.* [98] called voids that have a maximum dimension larger than 5 mm as macro porosity. A different approach used by Liang *et al.* [110] is to call macro porosity anything that can be visibly seen with less than 5X magnification but did not provide a dimension. Zhang *et al.* [111] used any void less than 300 μm as the transition from macro to micro porosity. Consistency in a definition and the reason for the selected value is lacking from literature. The exact transition dimensions between macro and micro porosity may continue to be challenged; however, definitions should be supported by reasoning. Having a practical definition is needed to provide clarity and consistency within industrial research.

Given the significant impact that macro porosity has on HPDC foundries, a well-conceived definition is needed. The variation seen in the literature is because scope of the definition was too small. Three factors should be considered when defining macro porosity: metallurgical formation mechanism, functional part requirements, and ability to inspect the defect. After considering these factors, the limit of detectability in inspection is the best transition between macro and micro porosity for HPDC.

METALLURGICAL FORMATION MECHANISM

From a metallurgical perspective, the size where micro porosity transitions to macro porosity is based on the secondary dendrite arm spacing (SDAS). This relationship is due to the different nature of the feeding flow of the bulk liquid versus the feeding in the mushy region between the dendrite arms. In the bulk liquid, the dendrite arms are not present and therefore cannot affect the maximum size of the porosity. This unconstrained macro porosity is formed when the bulk feeding flow is cut off. Dendrites restrict the feeding flow and constrain the maximum size of the porosity, creating micro porosity [112], [113].

Micro porosity size can vary significantly in different casting processes because SDAS is not a constant value for these processes. This is because SDAS is proportional to the cubed root of the cooling rate [112], [114], [115]. Sand casting and permanent mold castings have a wide range of cooling rates, making a universal definition of interdentric micro porosity in those processes difficult. HPDC cooling rates are different because they are consistently high with an observed SDAS in the range of 10-40 μm [116]. The small value means detecting true micro porosity voids in HPDC is often not industrially practical. This suggests that the definition of macro porosity for HPDC should be driven by the functional requirements or readily available inspection methods.

FUNCTIONAL POROSITY REQUIREMENTS

Functional casting porosity requirements influence a working definition for macro porosity. After castings are machined, they are inspected for macro porosity on machined surfaces. Acceptance criteria for given surfaces depend on the required function of the casting [12]. Thresholds are sometimes determined by past practices for a given original equipment manufacturer (OEM) and recommendations provided by vendors of assembled components such as piston rings, O-rings, and gaskets. In other cases, thresholds are set by failure points identified in finite element analysis (FEA) for the product. These porosity standards are treated as confidential trade secrets by OEMs; therefore, typical thresholds for specific applications are not published.

The North American Die Casting Association (NADCA) has provided an example of four levels of porosity on exposed machined surfaces in *Die Casting Porosity Guidebook* [94] with a Level 1 porosity being specified as a maximum size of 0.4 mm. Experience by the authors has shown typically the tightest specifications are a maximum porosity diameter between 0.4 mm to 0.5 mm (400 µm to 500 µm). The threshold for macro porosity would have to be smaller than this minimum acceptable porosity limit to use functional requirements as a meaningful distinction between acceptable and rejectable levels of porosity during inspection.

<u>INSPECTION ABILITY</u>

Castings are typically inspected in a raw state with X-ray or computed tomography (CT) equipment or by visual inspection after secondary machining. From an X-ray or CT image standpoint, production intent equipment has a pixel size resolution of 100 µm on castings that can fit up to 400 mm x 400 mm image windows [117], [118]. Specialized micro-CT research equipment exists to capture micro porosity resolution of 10 µm to 50 µm on parts typically less than 50 mm cube [102], [119]. This type of equipment has significant limitations in a production castings environment based on size and time needed to perform analysis. Therefore, the capable range of micro-CT equipment should not be considered in defining an industrial application of macro porosity. The focus should be given to capabilities of industrial, production intent X-ray equipment.

Visual inspection of castings is the next challenge. Visual acuity defines what the human eye is capable of detecting. Snellen eye charts, as typically found at optometrists' offices, were developed based on the studies showing human vision can generally resolve a visual target that represents 5 minutes of an arc. This 5 minutes of an arc is typically referred to as 20/20 vision [120]. The equivalent visual angle of a piece of porosity on a casting is determined based on defect size and surface distance as shown in Equation 11. Conversion from degree to minutes of an arc are needed to compare to a Snellen vision chart.

Equation 11 $$Tan\ \theta = \frac{Defect\ size}{Distance\ from\ eye}\ ;\ \theta = \arctan\left(\frac{Defect\ size}{Distance\ from\ eye}\right)$$

Table **23** includes a series of calculated equivalent visual angles for different defect sizes and distances.

Table 23: Equivalent visual angle for different defect sizes and distances

Equivalent Visual Angle (minutes of an Arc)				
Casting Defect Size (mm)	**Distance from defect (mm)**			
	150	**250**	**500**	**600**
0.1	2.29	1.38	0.69	0.57
0.25	5.73	3.44	1.72	1.43
0.5	11.46	6.88	3.44	2.86
0.75	17.19	10.31	5.16	4.30
1	22.92	13.75	6.88	5.73

It has been found that human ability to focus on near objects deteriorates with age. Typically, those in their 20s can focus down to 150 mm from their eyes. This erodes to 250 mm minimum focal distance as they age [121]. Given that 5 minutes of an arc is 20/20 vision [120], selecting 250 µm threshold between micro and macro porosity provides a realistic size that humans can inspect either without magnification in their youth and with minimal magnification as they age.

Based on the SDAS of HPDC, the functional specifications used to inspect castings, and the ability to detect and identify porosity with X-ray and visual inspection, a threshold of 0.250 mm (250 µm) is a defendable choice for the threshold between micro and macro porosity in HPDC. With a size defined and the reasoning supplied, the porosity that causes rejects in HPDC are macro porosity. The focus can now shift to how this macro porosity forms in production castings and what can be learned to improve the HDPC process.

EXPERIMENTAL STUDY

A small housing casting, produced on a 900-ton die cast machine out of A362.0 aluminum alloy, as seen in Table 24, was selected for this experiment. Previous X-ray audits showed varying levels of porosity within this casting even when process parameters remained consistent. The porosity was located near a semi-circular feature on the side opposite the gating. The porosity level found in this location of the casting does not impact overall part quality based on product testing.

Table 24: A362.0 Chemical Composition Limits

Si	Fe	Cu	Mn	Ni	Zn	Ti	Sr
10.5-11.5	0.40	0.20	0.25-0.35	0.10	0.10	0.20	0.05-0.07

Once the die was brought up to temperature with the typical start-up process, 100 castings were produced at identical process settings. These settings matched the production settings and had no adjustments made for the entire sample run. The castings were pin-stamped with a serial number in the robot extraction cell. With this serial number, all process data is traced to the X-ray images. The castings were completed in just under two hours.

Castings were sawed so the area of interest could be easily X-rayed to provide repeatable images and remove background features non-critical to this study. A sample of the sawed casting is seen in Figure 29. One casting was damaged during the sawing process used to prepare it for the X-ray (sample number 76). Its process information was removed from all analysis. This sample showed no significant difference in process values.

Figure 29: Sawed sample casting for X-ray

The X-ray equipment used was a Bosello SRE Max with a 225 max KV power rating. Images were registered using open source Fiji imaging software [122] and the selective plane illumination microscopy (SPIM) registration plug-in [123]. The X-ray images were reviewed following ASTM E505 standard [124] to determine porosity cause. Shrink porosity (ASTM Category C) and gas porosity (ASTM Category A) were seen within the castings. The castings were graded with a 1 (best), 2 (moderate), and 3 (worst) score. Figure 30 provides examples of all three gradings.

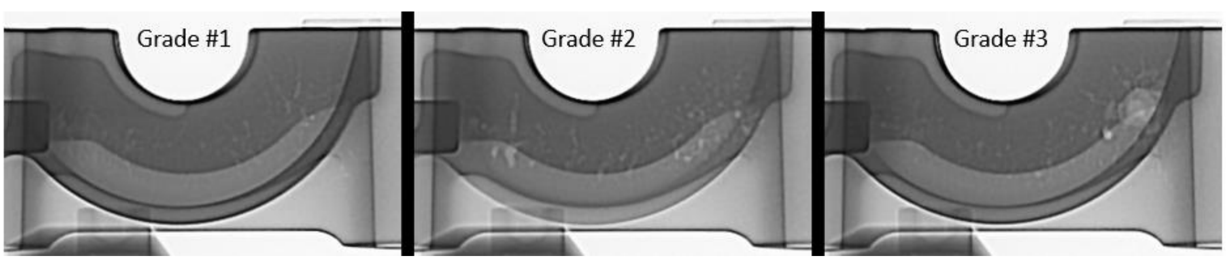


Figure 30: Example X-ray gradings 1 through 3

RESULTS AND ANALYSIS

Critical process parameters were collected for all sample castings (n = 99) during the experiment. The data was within typical production variation for the casting. Mean and 95% confidence intervals for these parameters are seen in Table 25. Based on the grading samples on the images, there were 9 samples each identified as Grade 1 (good) and Grade 3 (worst). These parts that represented the extremes of the macro porosity found within the casting were reviewed in detail.

Table 25: Mean and confidence interval of critical injection parameters

	Mean	95% Confidence Interval for Mean	Units
Cycle Time	68.0	67.8 - 68.2	second
Average Slow Shot Velocity	14.16	14.155 - 14.165	inches/second
Calculated Start of Fast Shot	14.27	14.266 - 14.274	inches
Average Fast Shot Velocity	141.69	141.61 - 141.77	inches/second
Intensification Pressure	5124.9	5119.0 - 5130.8	PSI
Intensification Squeeze Distance	0.174	0.171 - 0.177	inches
Biscuit Size	1.69	1.65 - 1.73	inches

The best castings showed scattered shrink porosity typically 0.2 mm to 0.4 mm thick and up to 1.0 mm long. This porosity was scattered throughout the section in review, with a tendency for it to form furthest way from the gating (part is gated from the left side of the X-ray images). The nine Grade 1 (best) castings can be seen in Figure 31.

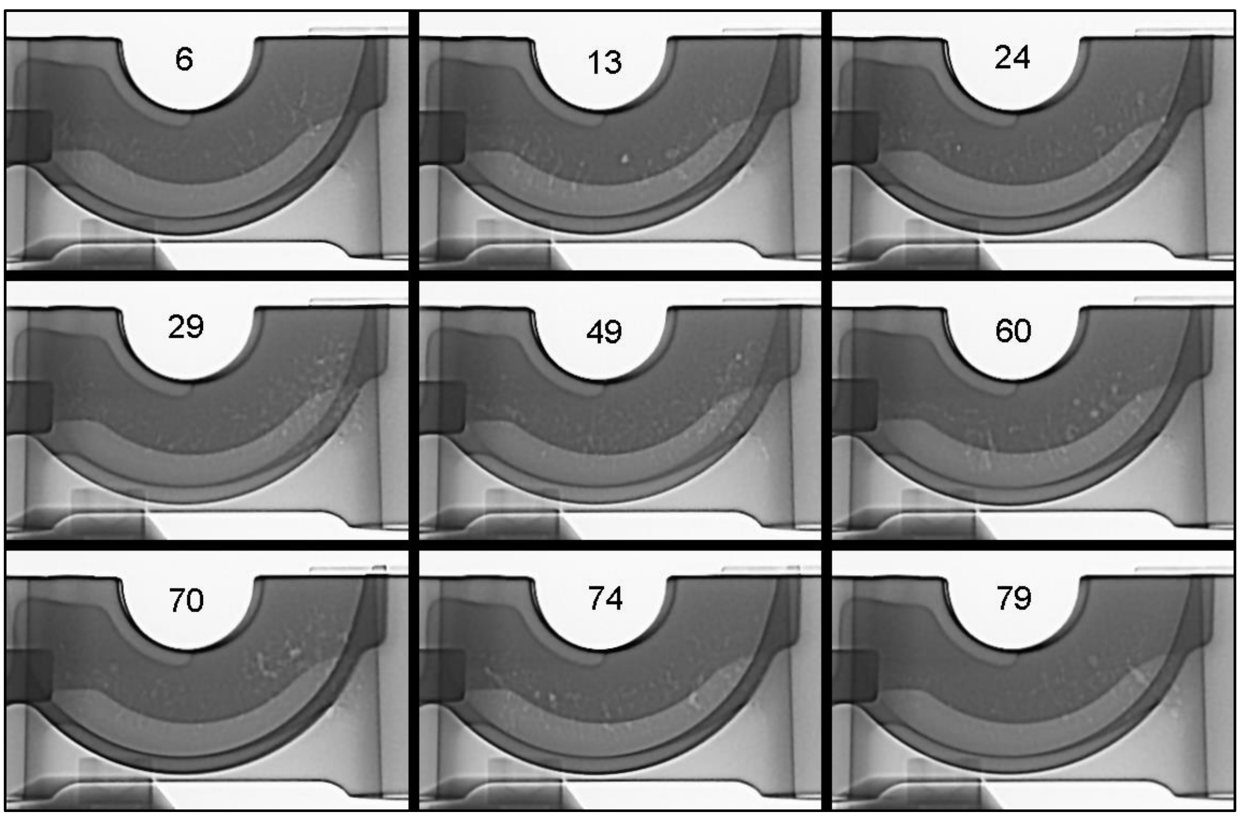


Figure 31: 9 best samples (Grade 1)

The worst castings showed a gas assisted shrink with rounder, but irregular shaped voids, consistent of gas feeding into a shrink porosity. Gas porosity within the worst castings grade were typically 3.0 mm to 4.0 mm in diameter. For shrink porosity, the worst castings had 0.4 mm to 0.6 mm thick and 5.0 mm to 10.0 mm long porosity voids. Overall, the worst of the worst (WOW sample - #43) had an approximate 8.0 mm diameter void in the casting. The nine worst castings (Grade 3) can be seen in Figure 32.

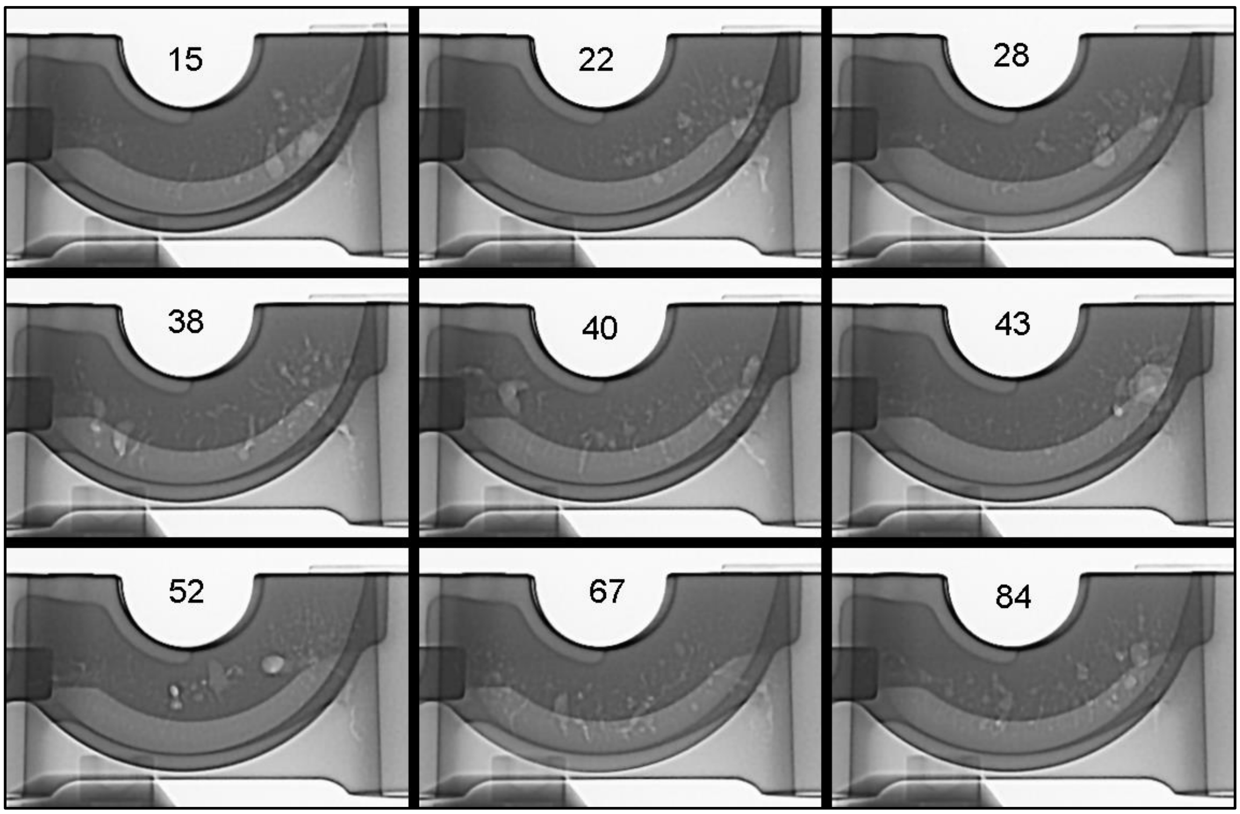


Figure 32: 9 worst samples (Grade 3)

Shapiro-Wilk normality tests [125] showed the process parameters to be non-normal in the 99 samples. As a result, the Wilcoxon Signed Rank test [126] was performed to compare the samples between the best and worst groupings. Table 26 contains all the individual recorded data for the samples. Table 27 shows the p-values calculated with the Wilcoxon test. None of the critical process parameters showed to be significant.

Table 26: Injection parameter data for best and worst samples

#	Grade	Cycle Time (s)	Average Slow Shot Velocity (in/s)	Calc Start Fast Shot (in)	Average Fast Shot Velocity (in/s)	Intens Rise Time (ms)	Intens Pressure (psi)	Intens Squeeze Distance (in)	Biscuit Size (in)
6	1	67.0	14.12	14.30	142.18	69	5096.5	0.157	1.89
13	1	65.8	14.15	14.30	141.34	74	5094	0.167	2.00
24	1	69.3	14.14	14.27	141.65	75	5108.7	0.177	1.79
29	1	67.7	14.18	14.28	141.84	70	5108.7	0.177	1.41
49	1	68.3	14.14	14.28	141.78	71	5123.3	0.157	1.35
60	1	69.0	14.17	14.26	141.87	67	5145.3	0.187	1.67
70	1	68.3	14.15	14.28	141.93	71	5125.8	0.187	1.47
74	1	68.7	14.15	14.27	141.93	72	5164.8	0.177	1.56
79	1	67.6	14.16	14.26	141.58	70	5150.2	0.157	1.59
15	3	67.6	14.13	14.30	142.12	74	5072	0.167	1.91
22	3	67.7	14.15	14.28	141.58	73	5172.2	0.167	1.21
28	3	69.5	14.14	14.29	141.39	70	5116	0.167	1.72
38	3	68.4	14.15	14.28	142.12	73	5101.3	0.167	1.72
40	3	67.8	14.15	14.28	142.12	73	5098.9	0.187	1.71
43	3	66.8	14.17	14.28	142.06	74	5130.6	0.157	1.04
52	3	68.4	14.18	14.27	141.5	71	5223.4	0.128	0.92
67	3	68.5	14.20	14.24	141.14	67	5147.7	0.177	1.76
84	3	68.3	14.21	14.23	142.15	68	5172.2	0.177	1.59

Table 27: Wilcoxon Signed Rank Test p-values

	Cycle Time	Average Slow Shot Velocity	Calc Start Fast Shot	Average Fast Shot Velocity	Intens Rise Time	Intens Pressure	Intens Squeeze Distance	Biscuit Size
p – value	**0.496**	**0.183**	**0.233**	**0.910**	**0.536**	**1.0**	**0.611**	**0.652**

The results of the experiment have shown that macro porosity formation is random when industry accepted control parameters are held constant in a production environment. The size and distribution of the voids varied significantly throughout the casting run, even though no process settings were changed. When comparing the best and worst samples, the process parameters showed no statistical difference. These parameters remain important to the process, but they do not fully explain the randomness associated with the porosity formation. As will be discussed in the next section, the general location of the porosity formation remained predictable, but the actual macro porosity formation was random.

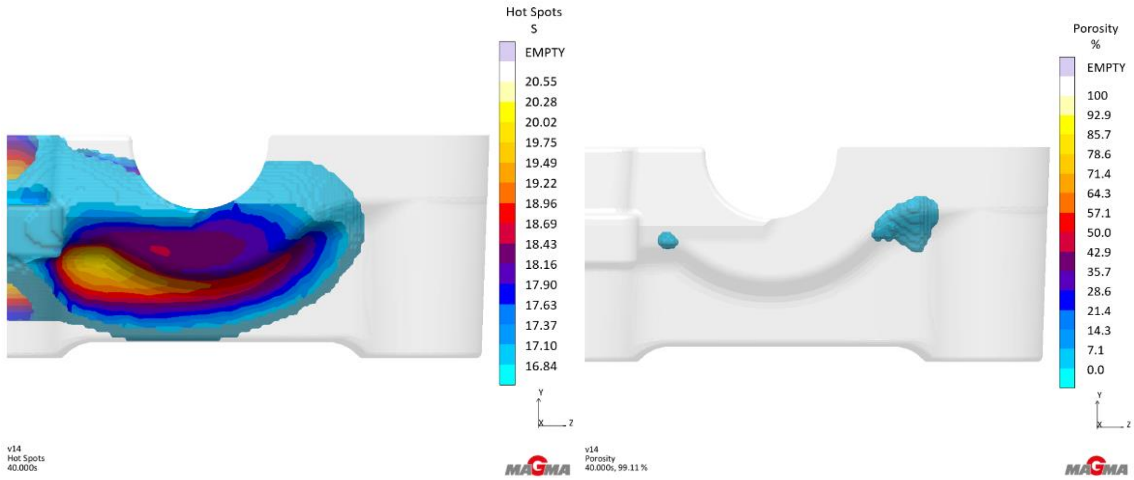
Production controlled processes produce randomly formed porosity. This random macro porosity should challenge many of the accepted analysis and quality implications of porosity in HPDC.

This experiment could did not determine if the Grade 1 castings were more dense (less porous) than the Grade 3 castings. It simply observes that the levels of macro porosity are higher while no process parameters varied significantly. This leaves open the question of whether the Grade 1 castings traded macro porosity for micro porosity that could not be detected by X-ray. Further study is merited and ongoing.

PART GEOMETRY IMPACT ON STOCHASTIC MACRO POROSITY FORMATION

Predicting the morphology of macro porosity is not a useful exercise. Macro porosity is easy to detect with visual inspection on machined surfaces or X-rays. Feeding-based rules have been successfully developed to manage porosity in sand and permanent mold castings. Macro porosity is reduced when an adequately large volume of liquid metal is available to replenish the volumetric contraction of the solidifying metal. The feeding volume must be connected by a liquid path throughout the solidification of the casting's cross sections being fed. From Chorvinov's rule, it can be surmised that thicker sections require more solidification time [113].

Computer simulations of these feeding rules are effective for identifying where porosity can form, but they fail to properly predict the morphology of the macro porosity. MAGMA was used to compare with the experimental results [127]. Six warmup cycles and one production cycle were calculated. Figure 33 shows the predicted porosity zone (a) and the pore volume fraction (b) for the studied casting in the area that was X-rayed. The porosity zone is the predicted hot spot that is the extent of the unfed liquid pocket. Porosity volume fraction is concentrated into two voids that approximate the center of the porosity observed via X-ray. More porosity is predicted on the right-hand side of the image like the experimental results.



(a) (b)

Figure 33: Simulated results of predicted porosity zone

Prediction of the macro pore morphology should be nearly impossible. This is because the pressure drop required to homogeneously nucleate porosity is in the giga-Pascal range. Instead, pores require a heterogenous nucleation site such as an oxide bifilm or pre-existing trapped gas pores [108]. Filtering, degassing, venting/vacuum, and good furnace maintenance practices can reduce the number of heterogenous nucleation sites, but they cannot be eliminated. Their occurrence will also be stochastic by nature. Since the nucleation sites are random, the macro pores also must occur randomly. The simulation software can predict these porosity zones but is incapable of showing the randomness of the size and shape of the macro porosity.

Additionally, HPDC feeds exclusively through the gating system into the part. Pressure is applied to enhance feeding, but gate locations are determined by die design and gate removal considerations. A temperature gradient from the thick sections of the casting to the gate cannot be ensured. Resulting unfed pockets of liquid create the macro porosity common to HPDC. Randomly sized and shaped macro porosity voids should form in these trapped pockets, but also the size and shape of the trapped pockets are random because the solidification during filling and time that pressure is applied to the liquid metal after filling is uncertain. Small extension or reduction of feeding will make significant changes to the total porosity due to the rapid solidification from the HPDC process.

Figure 34 plots the liquid volume in the sample area, as seen in Figure 29 and Figure 33, over the entire solidification time of the casting. Feeding is predicted to be cut off to the sample area at 9 seconds after filling, leaving 18.4 cm^3 of remaining liquid aluminum. 1.9% of the sample area will contain pores with the assumption the aluminum will have approximately 6% shrinkage. Increasing or decreasing the feeding time by 0.25 seconds changes the unfed liquid volume by ±3.5%. These small variations in feeding time can come from a host of uncertainties in the process. For example, cold flakes blocking feeding through the gate, spray variation, initial metal temperature and other variables can reasonably be assumed to affect the feeding time by 0.25 seconds or more.

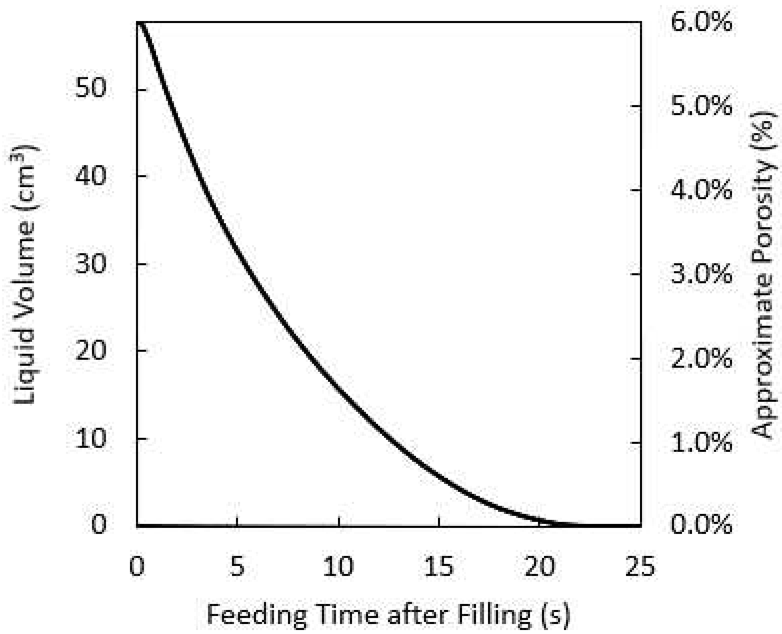


Figure 34: Volume of liquid in the sample area as predicted by MAGMA

RECOMMENDATIONS

The stochastic nature of macro porosity formation over the course of a normal production run and the inability to predict it precisely should challenge accepted practices within the industry. Cross-functional review of risks associated with the random porosity formation must be considered and identified throughout the casting life cycle. Critical areas that should be rethought include the design of castings, inspection of castings, and casting process optimization.

DESIGN OF CASTINGS

Castings free from macro porosity can only exist in a combined effort between casting geometry design and manufacturing process. Manufacturing process will not consistently overcome random macro porosity formation in castings designed with a high probability of porosity. A truly collaborative effort in design for manufacturing (DFM) and finite element analysis (FEA) is required for success. Casting simulation software is accurate at predicting location of macro porosity. However, the random nature of its formation may not guarantee a casting with the worst macro porosity condition is evaluated during functional testing. Casting simulation should be a cornerstone to the DFM activities on new casting designs. The risks associated with porosity results should not simply be written off to something the process could remove, particularly if inadequate sample sizes are reviewed and tested.

The design validation process should also be reviewed based on this random formation of macro porosity. Larger lots of testing samples should be a focus during the initial production validation process phase. Functional testing should look at both worst case macro porosity formation and other random samples with castings produced with normal production settings. This will require large X-ray studies to help review the amount of macro porosity formation that exist and select the castings that should be functionally tested to ensure product performance.

INSPECTION OF CASTINGS

Because macro porosity forms randomly, variation should be expected during a typical production run. This should cause one to question traditional audit and inspection processes in X-ray and at secondary machining. Better audit sizes can be determined by understanding the probability of uncovering a worst-case macro porosity situation.

With an exception of some structural automotive parts, in-line X-ray equipment is not typically found in most industrial die casting foundries. The cost of equipment and time to process when compared to the risks of porosity typically makes this an uneconomical process. Instead, audit X-rays are performed on randomly chosen samples during the production run. The number of samples and a defect rate caused by worst-case macro porosity can be used to understand the probability of selecting a sample lot and seeing no defects. The binomial probability function is used to find this probability. Table 28 shows the probability of finding no defects for different defect rates and sample sizes.

Table 28: Sample Binomial Probabilities

Binomial Probability of Different Defect Rates and Sample Sizes								
Defect Rate	**Number of Defects Found**	**Randomly Sampled Lot Size**						
		3	**5**	**10**	**15**	**30**	**50**	**100**
1%	0	97.0%	95.1%	90.4%	86.0%	74.0%	60.5%	36.6%
2.5%	0	92.7%	88.1%	77.6%	68.4%	46.8%	28.2%	8.0%
5%	0	85.7%	77.4%	59.9%	46.3%	21.5%	7.7%	0.6%
10%	0	72.9%	59.0%	34.9%	20.6%	4.2%	0.5%	0.0%

Understanding this probability should influence how troubleshooting is completed within the foundry. As an example, if a casting truly has a 2.5% scrap rate based on the random macro porosity and the X-ray technician randomly pulls 3 castings for inspection, there is 92.7% chance none of the castings contain that defect. If the defect rate and the process remain constant and the inspection has experience seeing high probability of all good castings, what happens once that inspector finds a defective part? It is slightly better than a coin flip to have a defect occur every two production weeks, based on a three-shift operation. The warning flags in the foundry are sent out and the troubleshooting begins. Another random sample selected could then show all good parts, and the investigation of "what changed in the process" may occur, wasting resources because the process has not changed, and the sampling was just poor. Worse yet, someone may take the failed inspection results and change the process to try to "improve it." Now the actual defect rate shifts from 2.5% to some other unknown value. This can lead to a spiral of process changes over the production life of a part, with limited knowledge of what scrap rates are occurring.

By having a good working knowledge of actual scrap rates for given castings based on machining feedback, one can make an educated decision on how to approach the defective audit sample. The Bayes Success-Run Theorem can be used to help determine a lot size given the historical defect rate and the confidence the manufacturing desires [128].

CASTING PROCESS OPTIMIZATION

The randomness of the macro porosity should cause the industry to review the techniques used for process optimization. An appropriate sample size is needed for any optimization process to ensure the worst-case scenarios are detected in the macro porosity formation. Also, additional process monitoring of the HPDC equipment, beyond the injection system, is needed.

Optimization processes are based on feedback from inspection results. With HPDC, that feedback is based on inspection results in either the raw casting state with X-ray images or in a final processed state with a visual inspection for porosity. As discussed previously, lot sizes need to be reviewed and appropriately selected based on historical scrap rates. Without this, samples may not capture the potential worst-case stochastic formation of the macro porosity. This sample size is also pertinent when setting up an optimization design of experiments (DOE). Without the correct sample size, the variation that exists due to the random formation could be missed.

Injection parameters and metal holding temperature are typically the focus of in optimization publications in HPDC [18], [21], [22], [129]. This approach is logical given the history and commercially available products for the industry to capture this data. Questions should arise on the focus and priority of these process inputs when significant macro porosity variation occurs. It is clear from simulation that injection parameters can influence predicted zones. This study shows the random formation of porosity in these zones. The injection parameters should remain controlled and monitored; however, data collection on additional process monitoring systems should be prioritized [130] to potentially further reduce this predicted porosity zone. These additional systems could include die temperature [9], vacuum [106], and equipment cycle time and performance. Optimization of these additional parameters could reduce the predicted porosity zone. Therefore, the macro porosity that forms in that zone will also be reduced.

CONCLUSIONS

High pressure die casting suffers from a porosity problem. By better understanding how macro porosity forms, improvements can be made within the industry. Like micro porosity, macro porosity is randomly formed in HPDC. Casting design, inspection, and optimization processes are all affected by macro porosity. This random macro porosity formation has been shown by an industrial case study and fundamental theory.

Simulation software uses casting geometry and tooling design to predict porosity zones but cannot predict the actual random size and distribution of those voids in the zone. Understanding this stochastic formation should challenge previously accepted practices and improve the comprehension and classification of macro porosity defects in die casting. Specifically, this provides the industry with three areas of further study.

The first area of study is within the industry's process control approach. Injection parameters are the main focus within the industry and academic research for HPDC process control. The conducted experiment shows castings produced with statistically similar injection parameters, cycle time, and biscuit size can produce significantly different levels of porosity as seen in the Grade 1 and Grade 3 examples. This work shows the traditional parameters are not fully capable of reducing the randomness that can exist in the HPDC process. Additional research is needed to understand if other process parameters such as variability in metal cleanliness, furnace temperature, die temperature before die close, spraying, die thermal management, or others could reduce the predicted porosity zone and therefore region where stochastic porosity forms.

The next area to research is the density in the predicted porosity zones between castings. Research has shown that density is not a good predictor of mechanical properties [131], [132]. The difference levels of void space visible in the Grade 1 versus Grade 3 X-rays leads to questions regarding the density of the predicted porosity zones. Are the densities of these grade differences the same with different distributions of size of the macro and micro porosity? This is a useful question to have answered as its impact on quality inspection results (acceptable versus scrap casting) and perceived mechanical properties could be misleading to the industry.

From this work, we know there is randomness in the size and shape of macro porosity. This randomness influences classification of defects and process optimization decisions. Misclassification of macro porosity can lead to poor predictions of quality when supervised machine learning algorithms are used. In this case, two significantly different outputs on X-ray images are produced from nearly identical inputs. It becomes impossible for machine learning to find a pattern in what fundamentally becomes noise in the “results” created by random macro porosity formation. Furthermore, sample sizes for optimization studies must be carefully planned based on these random macro porosity formations. Small sample sizes will have a higher probability that the true worst-case macro porosity formation is not seen, thereby providing misleading optimization guidance. These areas are worthy of additional research.

In conclusion, the stochastic nature of macro porosity formation within the prediction porosity zone should challenge the industry to further research HPDC process in production environments. By researching these topics further, the industry will be better positioned to help improve overall HPDC casting design and manufacturing of parts.

Chapter 4: Classification Issues and Critical Error Threshold[2]

Chapter 4 builds on the randomness demonstrated in Chapter 3, but this time from a human perspective. Production manufacturing environments rely heavily on human inspection for result classification, and humans are poor visual inspectors. Four elements of defect misclassification are defined in this chapter. These misclassifications of results create dataspace overlap within datasets, which makes applications of supervised machine learning difficult. Further, Chapter 4 introduces the concept of the critical error threshold. The accuracy of a supervised machine learning algorithm must be greater than this critical error threshold for the algorithm to provide business value.

INTRODUCTION

Machine learning (ML), a specific subset of technology in the Artificial Intelligence (AI) field, has seen an explosion in commercial use in the past 30 years. From utilizing computer vision to recognize hand-written numbers for the post office in the 1990s [134] to current applications to trace and treat the COVID-19 pandemic [135], the ability of machine learning to unlock patterns and provide insight into data has provided great benefits to society.

Machine learning has also found increased use in manufacturing and operational systems. Industry 4.0 and Smart Manufacturing have driven digitalization of manufacturing operations. Microsoft and Amazon have made large investments in this technology as they partner with manufacturers to provide data collection and analytics tools [136] [137]. Estimates show a 10.1% compounded annual growth rate of investment into smart manufacturing technologies with a 2024 global spend rate predicted at $400 billion USD [138].

---

[2] This chapter is an edited version of a paper published in International Journal of Metalcasting, included with permission. D. Blondheim, Jr, "Improving Manufacturing Applications of Machine Learning by Understanding Defect Classification and the Critical Error Threshold," *International Journal of Metalcasting*, 2021, doi: 10.1007/s40962-021-00637-0 [133]

This growth of ML within manufacturing comes with trepidation by industry leaders. A 2021 survey by KMPG [139] shows 55% of industrial manufacturing business leaders believed AI adoption within manufacturing is moving faster than it should. Within this same survey, the industrial manufacturing segment had the largest number of respondents who believed that AI initiatives delivered somewhat or significantly less value within the organization, with more than 21% falling in these two categories. There is a considerable challenge to implement ML in industrial settings. Experience has shown many software providers oversell the benefits of ML/AI and downplay the effort needed to implement ML to an organization's executives. The effort to successfully implement ML should not be underestimated. Many ML projects in manufacturing fail to provide value. Unfortunately, the KMPG survey highlights the gap that exists in leadership ranks of organizations. Seventy-eight percent of executives found value within AI initiatives while only 50% of managers reported finding the same value in these projects. This work will show how misclassifications of product within production environments can assist with these ML failures.

Beyond the leadership survey, research has also highlighted the challenges with applying ML in manufacturing. Wuest *et al.* [5] describe challenges that manufactures face with ML, including the acquisition and pre-processing of data; high dimensionality of manufacturing data; highly unbalanced data sets; selection of ML algorithms; and interpretation of the results. Baier *et al.* [6] interviewed multiple industries, including manufacturing, to identify the most significant challenges in implementing ML. They identified challenges existing in three key areas:

1. **Pre-deployment stage:** companies must gather the right quality data in sufficient amounts.
2. **Deployment stage:** challenges are associated with gathering large amounts of data and ensuring hardware and software systems can handle the volume of data.
3. **Non-technical items:** acceptance of ML models for people who have no background in data science and people questioning the practical value ML can provide in manufacturing.

The pre-deployment stage is critical to ensure the data used for ML training is correct. Misclassification of training data can lead to errors, specifically within supervised ML algorithms. Supervised and unsupervised ML are the most used algorithm types in ML applications. In supervised ML, algorithms are trained from labeled data sets. Multiple input values are provided to a supervised ML algorithm to find a pattern to help accurately predict a future result. The algorithm is trained on both the inputs and results. It is then tested against a data set that the algorithm was not trained on to see how well the model can predict results [140], [141]. As will be discussed, misclassification of products occurs in production manufacturing environments. This noise introduced into the supervised ML algorithm leads to poor modeling, prediction failures, and ultimately ML not being implemented. In unsupervised ML, no results are known in the data sets. Only the input parameters are provided for the algorithm to detect patterns. Unsupervised ML focuses on clustering analysis and anomaly detection [140], [141]. The focus of this work is on supervised ML as it is commonly used in manufacturing applications of ML. Supervised ML would be used in classification predictions, such as predicting part quality. However, due to some of the shortcomings associated with misclassification of data used in manufacturing, applications of unsupervised ML will be discussed in the recommendations.

Challenges with applying ML in the metal casting industry have also been studied. Sun *et al*. [7] discusses unbalanced data sets, sporadically labeled data, and how metal casting data is not seen as the "big data" associated with ML given that some casting processes can take years to generate tens-of-thousands of rows of data. Traceability and classification of castings are challenges in a production environment [8]. Classification of casting quality is of high interest when attempting to optimize the process within the foundry. The technical process to serialize parts is not as difficult as the actual process of collecting and tracing these parts through multiple manufacturing operations, facilities, and companies within the supply chain [9]. Often a human is at the end of this process inspecting the casting and classifying the results against print specifications. Studies have shown that humans are inconsistent visual inspectors, missing 20% to 50% of defects in various manufacturing processes [120], [142]–[146].

Misclassification of results causes data space overlap in machine learning. This overlap can fundamentally diminish the results of supervised ML. Overlap occurs when multiple classifications of training data results occur in the same dimensional data space [147]. A simple 2-dimensional example of data overlap is found in Figure 35. The left side of the figure shows why it is difficult for ML applications to find a pattern in noisy, overlapped data. ML can build models that accurately predict the situation with the limited overlap as seen on the right side of Figure 35. Misclassification of results, either done by human error or by process variation, is a significant contributor to overlap within manufacturing data sets.

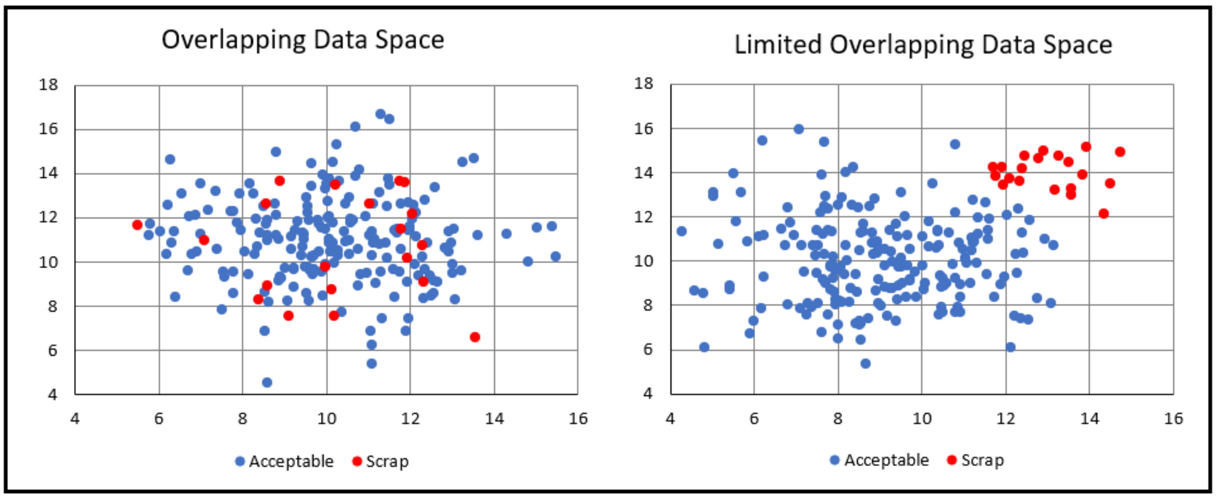


Figure 35: Example of overlapping data space

This work will show how defect classification systems used in product manufacturing influence ML models. Misclassifications of defects can make applications of supervised ML exceed a critical error threshold (CET) rendering the model financially useless. Proposed are four elements of defect classifications in production environments that cause these issues for ML:

1. Binary Acceptance Specifications
2. Stochastic Formation of Defects
3. Secondary Process Variation
4. Visual Defect Inspection

These four elements will be reviewed from the high pressure die casting (HPDC) perspective. Due to the modularity of these elements and the generality of the CET, these concepts have applications in other manufacturing processes such as sand casting, permanent mold casting, machining, painting, and assembly. Finally, the issues with the misclassification will be discussed in terms of data space overlap, which creates a bias shift while training supervised ML.

The goal is to help organizations understand and avoid the pitfalls commonly misclassified manufacturing data can have on ML. With this knowledge, organizations should drive to improve data classification process especially when there is an intention to implement ML. At a minimum, this work can provide insights into why some applications of ML may have failed in the past due to poor, but commonly used, practices of data collection in manufacturing.

BACKGROUND IN HIGH PRESSURE DIE CASTING

High pressure die casting (HPDC) is composed of multiple systems that control hydraulic, mechanical, and thermal processes to produce near net shape castings with short cycle times in complex metal molds [1]. HPDC typically focuses on large production volumes due to the capital investment in equipment, tooling costs, and short production cycle times.

The design of the casting and tooling, combined with the setting and control of the manufacturing system, dictates casting quality and equipment performance. The North American Die Casting Association (NADCA) estimates annual sales of $8 billion USD for aluminum die casting in 2019. This represents more than 80% of the American Foundry Society's (AFS) forecast in all aluminum castings of $9.67 billion USD [3]. Based on a 2014 NADCA study, current methods for control and optimization in HPDC produce a median scrap rate of 8% of parts produced and equipment utilization of 68% within the industry [4]. Improving uptime and reducing scrap costs can create meaningful value.

The process associated with HPDC makes castings prone to defects. As aluminum turns from liquid to solid during the casting process, its volume shrinks by approximately 6% [94], [97]. If unfed, the casting will have void space called porosity. Porosity is typically found in the thickest portions in the casting. HPDC's goal is to reduce this unfed area by applying high amounts of pressure on the liquid metal to continue feeding the casting during solidification [1], [13]. Casting geometry design, gating and die design, and equipment size all factor into how well this void space can be fed [14], [94]. In production, all castings will have some level of void space. It is the goal of the design and manufacturing engineers to place this in non-critical areas of the casting. Unfortunately, as seen in the 8% median scrap rate, porosity-related defects continue to challenge the quality and reliability of the HPDC process.

Porosity is just one of multiple potential defects in HPDC. Defects can be described both in the location as to where they form and the metallurgical cause. Multiple publications are available that describe different classification systems for HPDC defects in these terms [95], [99], [100]. These defect classification systems are important to provide consistent levels of defect feedback to the foundry. It is difficult to successfully complete quality control and improvement projects without a common defect language [100]. In addition to porosity, the other most common defect types in HPDC include laminations, inclusions, leaking, cracks, blisters, soldering, erosion, and deformation. Although all these defects are important, the focus of this work is on the porosity defects. Porosity is a leading cause of scrap within the die cast process [95]. Approximately 30% of the foundries within the industry have identified the need to address porosity as top concern [96]. The HPDC community will see great value by applying ML to improve porosity and other quality defects.

Porosity that impacts the quality in production castings is created from entrapped gas, called gas porosity, or volumetric shrink, called shrink porosity. These two types of porosity can combine to form gas-assisted shrink. These defects' causes and physical descriptions are well published [13], [14], [94], [95], [97]. Production examples of porosity defects can be seen in Figure 36.

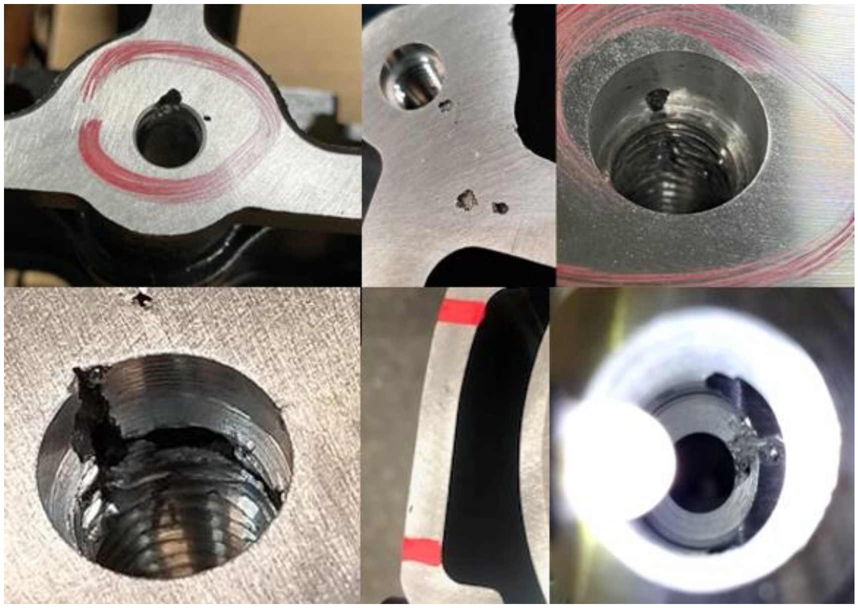


Figure 36: Examples of HPDC porosity

Porosity defects are commonly internal to the casting. As a result, the defect is usually uncovered after additional processing. This means manufacturing costs such as trimming, shot blasting, painting, machining, testing, and inspection are added to the casting cost prior to the defect being found. In addition to the costs, there is a time delay associated between casting production and when the feedback is received. In the ideal situation, this feedback time is extremely short due to single piece flow. In real-world manufacturing facilities, it is common for there to be several weeks, if not months, between casting production and defect feedback. Reasons for these delays include batch processing typical in manufacturing; complex, international supply chains; transportation time and safety stock levels; multiple vendors performing different manufacturing operations; and insufficient inspection and analysis approaches for product. To minimize this risk, foundries utilize casting simulations; experiments to help determine process settings and limits; and radiographic, or X-ray, audits to identify changes to the HPDC process which may result in increased porosity.

Casting simulations have proven to be successful at identifying high-risk areas of porosity within HPDC [108]. By simulating the filling and solidification of the given casting geometry and die design, software can predict areas of entrapped gas and unfed shrink regions. These predicted areas represent a probability of where porosity is most likely to form in the casting. Ideally, these areas are eliminated by geometry changes in the casting, changes to the die's gating or venting system, thermal management of the tool, or process changes on the injection of the metal [12], [15], [23], [70], [148]. This cannot always be accomplished due to functional requirements of the final part assembly and process limitations. Since die cast tooling is so complex, once a die is created it is expensive, difficult, and risky to make changes to tooling while trying to support production volumes. Additionally, simulations are successful at identifying concern areas, but they are not 100% accurate due to the assumptions and generalizations made when setting the initial conditions of the simulation. As a result, the HPDC industry often works in less than ideal situations when trying to produce castings free of porosity.

Foundries will often X-ray inspect production castings to minimize the risk of creating castings with porosity. In the current industrial environment, usually only structural automotive [20] or aerospace [149] castings utilize 100% X-ray inspection. These types of castings can justify the added costs of equipment and processing time to inspect every produced casting in critical locations. In most other casting applications, the production cycle time, cost, or functional need creates an economic situation where castings will only be sampled for X-ray inspection. This audit will help identify changes in typical porosity levels, which may indicate special cause variation in the manufacturing process. Examples of porosity in X-ray images are seen in Figure 37.

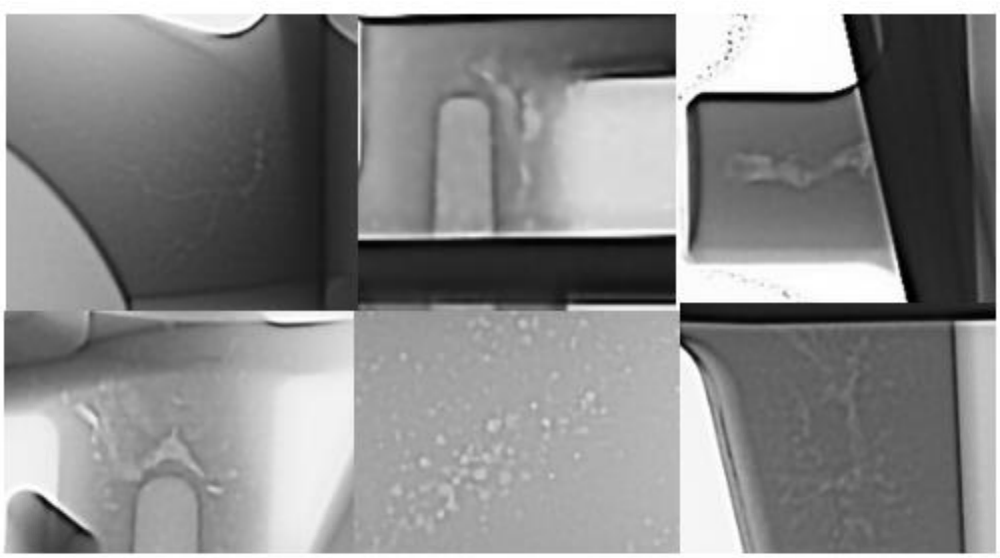


Figure 37: Sample X-Ray images of porosity

Even by reducing the possibility of porosity with simulations and control of the process with X-ray, porosity defects still pass through the supply chain and are identified after additional value is added. Without 100% X-ray, porosity scrap is found after machining when the void is exposed with milling or drilling operations. After the machining operation is complete, human operators visually inspect the machined castings to determine if they pass a porosity specification. The decision made is binary. Castings will either be classified as *acceptable* or *scrap*. When a *scrap* part fails to meet porosity requirements, a repair method such as welding or epoxy may be allowed in certain applications. For this work, a focus will remain on the primary classifications of *acceptable* or *scrap* since the goal is to create ML predictions that would help eliminate poor quality regardless of its ability to be repaired. This binary classification of castings contributes to the issues with applying machine learning and will be discussed in detail in the next section.

FOUR ELEMENTS OF DEFECT CLASSIFICATION

In HPDC, obtaining accurate classifications is a challenge that needs to be fully understood. Misclassified training data has consequential impact on the training bias of supervised ML. There are four key elements of classification: *Binary Acceptance Specifications, Stochastic Formation of Defects, Secondary Process Variation,* and *Visual Defect Inspection.*

<u>BINARY ACCEPTANCE SPECIFICATIONS</u>

In Juran's Quality Handbook, a defect is described as "anything that does not meet or exceed the requirements of the customer, the business, or the process" and states the "importance to have a realistic threshold for what is called a defect" [142]. In manufacturing environments, this threshold is given as a quality specification for the product. The specification is typically set during the design and testing phases to ensure the product achieves the functionality intended.

A formal specification is included in manufacturing prints or as a stand-alone document to ensure conforming product is provided to the customer. This specification becomes an aid during the inspection and will include a threshold of acceptable level of defects. Parts that exceed this threshold are classified as *scrap*, while parts below are classified as *acceptable*. A binary classification is made.

A common approach used for porosity in a production environment is defining a maximum permissible porosity size and number of pores per region [14] . Figure 38 shows an example of a theoretical specification for a casting based on a max pore size of 2.0 mm and a max allowable number of pores to be 4 in a 25 mm² area.

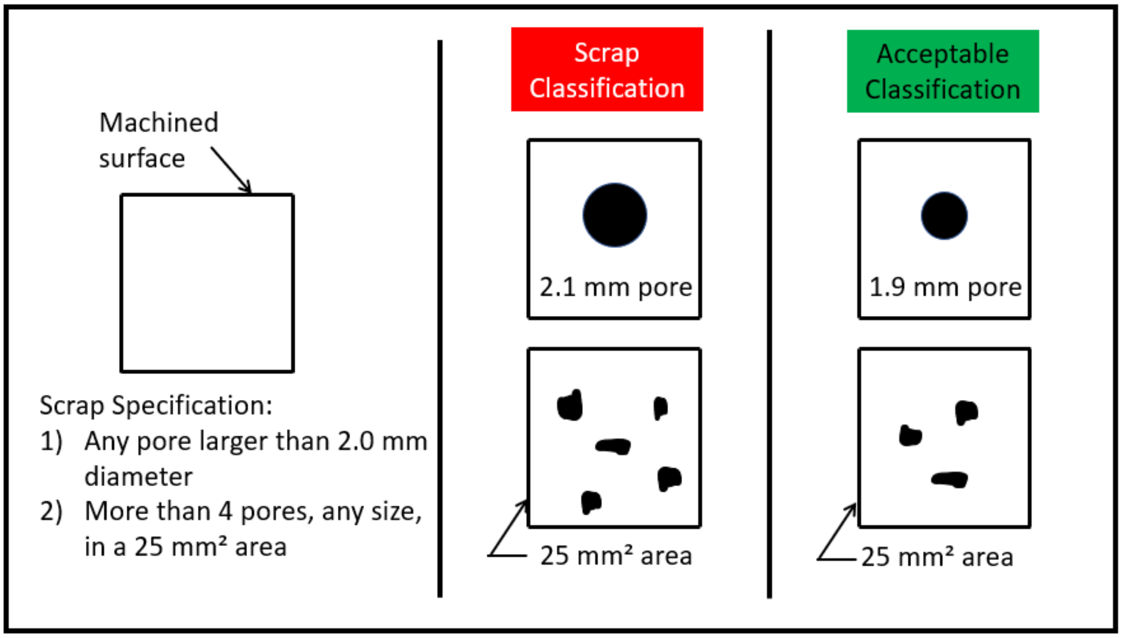


Figure 38: Classification issue due to binary acceptance specifications example

In most applications, some level of porosity is acceptable. It is common for one casting to have varying porosity zones on different part features. Each zone may have its own unique threshold based on functional needs. Sealing surfaces between machined castings or critical threaded holes might have a tighter tolerance than non-functional machined surfaces or clearance bores. Typically these zones are identified on a manufacturing print [14]. No universal standards exist for acceptable porosity levels within castings. Threshold are set based on supplier requirements on assembled products (like sealants or o-rings), past design practice, and functional testing. These thresholds are often deemed proprietary information to original equipment manufacturers (OEMs) and are protected through non-disclosure agreements.

The binary classification process associated with specification requirements creates a problem for ML applications. Defects form along a continuous measure of size. That important detail is lost when a binary *acceptable/scrap* result recorded. As seen in Figure 38, a pore with a 2.1 mm size is labeled as scrap, but a pore at 1.9 mm is acceptable. The loss of fidelity on the defect measurement with a binary classification creates problems for supervised ML. Results in the data space may overlap due to this lack of distinction with a binary classification.

STOCHASTIC FORMATION OF DEFECTS

The injection and solidification of castings follows known physical rules that are modeled in casting simulation software. Rules for fluid flow, heat transfer, feeding, cooling, and many physical calculations are factored into the simulations. As a result, simulations have proven to be good at predicting locations or zones as to where porosity defects will occur during the casting process. Figure 39 shows an example of this porosity predicted zone produced by MAGMA simulation software [127].

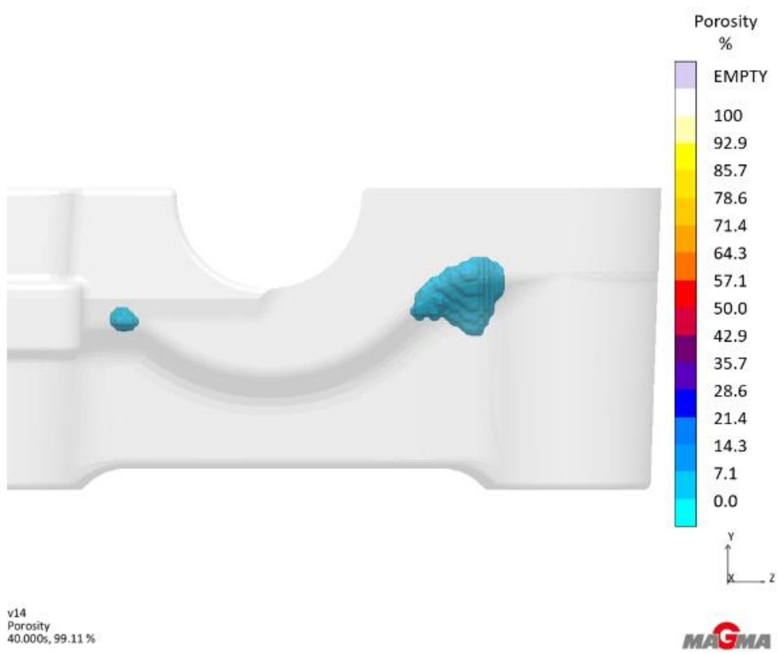


Figure 39: Example of simulated predicted porosity zone

In production, this predicted porosity zone does not create the same porosity from casting to casting. There is a stochastic, or random, nature to porosity formation within a casting. Theory says this stochastic formation occurs due to the random formation of dendrites as the metal starts to solidify which cause shrink porosity [101], [150] and the heterogenous nucleation sites for pores that can cause gas porosity [97], [101], [150]. Oxides and inclusions are examples of these heterogenous nucleation sites for porosity that are randomly distributed through the liquid metal [97].

This stochastic theory was shown in a recent industrial experiment [93]. In this experiment, one hundred castings were produced with no process changes. These castings were serialized and inspected using a Bosello SRE Max with a 225 max KV power rating X-ray unit. Results showed significantly different porosity formation between castings. Figure 40 shows two castings from this experiment that were sequentially produced. The formation of the porosity was in the simulated predicted zone, as seen in Figure 39. However, the porosity was random and different between sequential castings even with no process changes. That paper showed no statistical difference in the critical process parameters between the best nine castings and the worst nine castings.

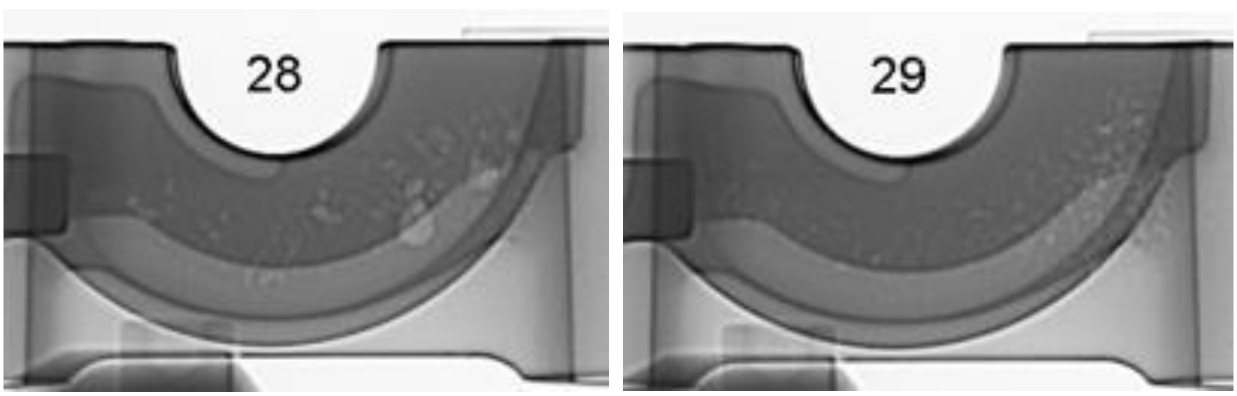


Figure 40: Sequential castings showing stochastic porosity formation

If the predicted porosity zone is away from any machined surface, the randomness associated with the porosity formation will have no impact on the classification of the final part. The porosity will not be uncovered or seen during visual inspection. If a hole or machined surface is cut into a zone predicted to hold porosity, the machining could potentially expose the porosity, pending random formation of the porosity. Figure 41 provides a visual example of how different stochastic porosity formations can alter classifications of the castings.

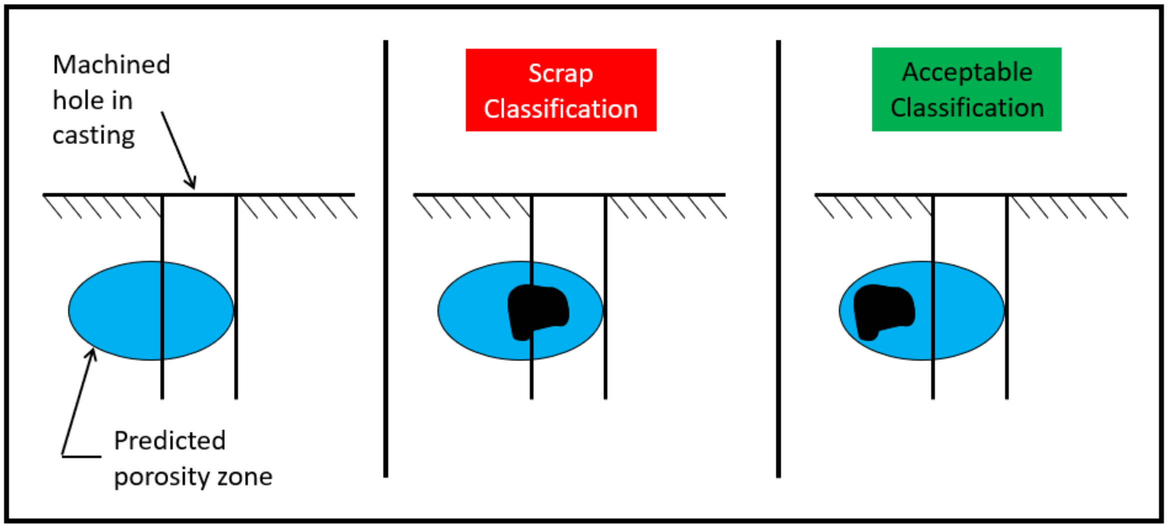


Figure 41: Classification issue due to stochastic defect formation example

The randomness of the porosity formation directly connects to the ML problem of overlap. The data space of these points will be the same if two castings have identical input parameters. However, the random formation of the porosity causes one casting to be scrap and the other acceptable. The overlap will cause the ML algorithm to struggle and possibly fail at providing meaningful insight. The collected data amounts to noise that the ML cannot pattern.

SECONDARY PROCESS VARIATION

Selecting machining tolerances for a casting is a critical part of the product design. Machining tolerances are selected based on process capability, manufacturing costs, quality, life-cycle impacts, and functional requirements [151]–[153] . There are various methods for optimizing the tolerance selection that have been studied and published [151]–[156].

The natural variation that occurs in machining processes and the tolerances associated with the feature create allowable differences part to part. Geometric dimensioning and tolerancing (GD&T) are used on manufacturing prints to control the measurement of features. The ASME Y14.5 standard [157] is often referenced for GD&T requirements on prints [158]. The variation associated with the manufacturing process by the tolerancing affects the classification of casting defects. Figure 42 shows an example of the effect tolerancing and machining variation can have on defect classifications for machined surfaces and a drilled hole.

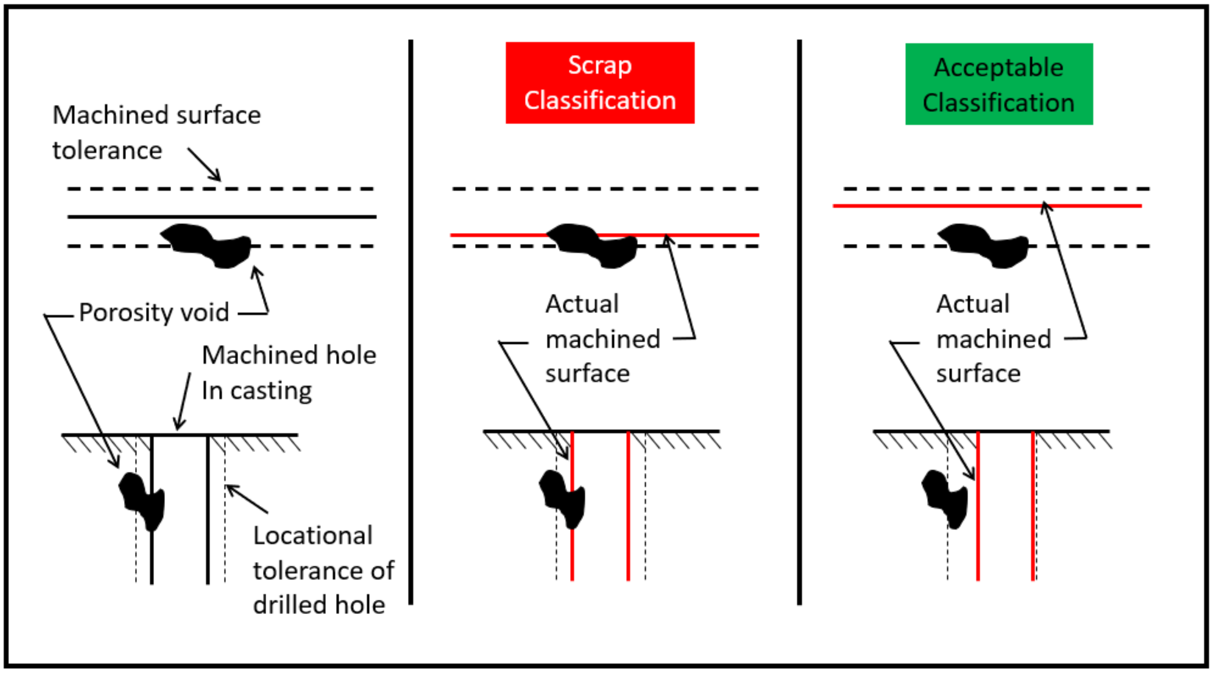


Figure 42: Classification issue due to secondary process variation example

Like the random formation of defects, the variability within the processing of the part has the potential to create data space overlap. This overlap could potentially be avoided if every machined dimension was also collected and included into the algorithm. However, this additional inspection would be extremely costly to do in a production environment. Additionally, it would still not guarantee results of the ML prediction since the true condition of the part below the machined surface is still unknown. The ground truth needed to train ML algorithms for accurate predictions is not collected by traditional means within manufacturing.

<u>VISUAL DEFECT INSPECTION</u>

The last element that affects the classification of castings is the visual inspection process. Although the technology exists for computer vision inspection in certain applications such as two-dimensional surfaces [120], [159], [160], cost and product mix within most manufacturing plants prevent this from being widely applied [120]. Humans often complete the inspection and classification of defect results.

Much research has been published on a human's ability to complete visual inspections on machined product. A person is capable of identifying 50% to 80% of defective machined products with 100% visual inspection [120], [142]–[146]. Considerable training on visual inspection is needed to achieve the high side of this range. This includes a comprehensive knowledge of defects, planned eye scanning paths on parts, and appropriate environmental conditions such as lighting [120], [161], [162]. Many manufacturing companies do not undertake this considerable effort, even though there are large costs associated with poor quality being passed as acceptable [120], [142]. As a result, classification rates will be on the inferior end of the range.

The ramifications of poor classification practices because of visual inspection cannot be overlooked for supervised machine learning. Most manufacturers stay in business because they can develop a process that is capable of a high yield. As a result, the data sets generated are highly unbalanced. Acceptable results greatly outnumber scrap results. This issue is compounded when potentially half the defective product is labeled as acceptable instead of its true scrap classification.

The adage "garbage in, garbage out" can easily be apply to supervised ML algorithms based on visual inspections. Poor visual inspection leads to incorrect classification of results. These incorrect results create data space with overlap. Overlap will cause ML algorithms to struggle to find a pattern in the noise collected within manufacturing data sets.

COMBINATION AND SUMMARY

The four elements described can individually contribute to misclassification of defects. Unfortunately, these elements do not act independently of each other. Instead, they combine and change through time to create more classification confusion for supervised machine learning.

In particular, the *Stochastic Formation of Defects* and the *Secondary Process Variation* combine to create misclassifications. In some combinations where the porosity forms and how it is machined, the results are classified as acceptable. In other cases, the casting may be classified as scrap. Figure 43 was created to visualize these combinations. The first column in the figure shows different levels of random porosity in the predicted porosity zone. A machined hole is drilled into the casting. In the first row, the parts are classified as acceptable based on a theoretical specification. In the second row, the parts are classified as scrap. In the second column, the random porosity formations are exchanged between the top and bottom row. Changes to the location of the hole based on machining variation are shown that make the previously acceptable parts scrap and scrap parts acceptable. The true classification is never known without a 100% X-ray inspection of each casting.

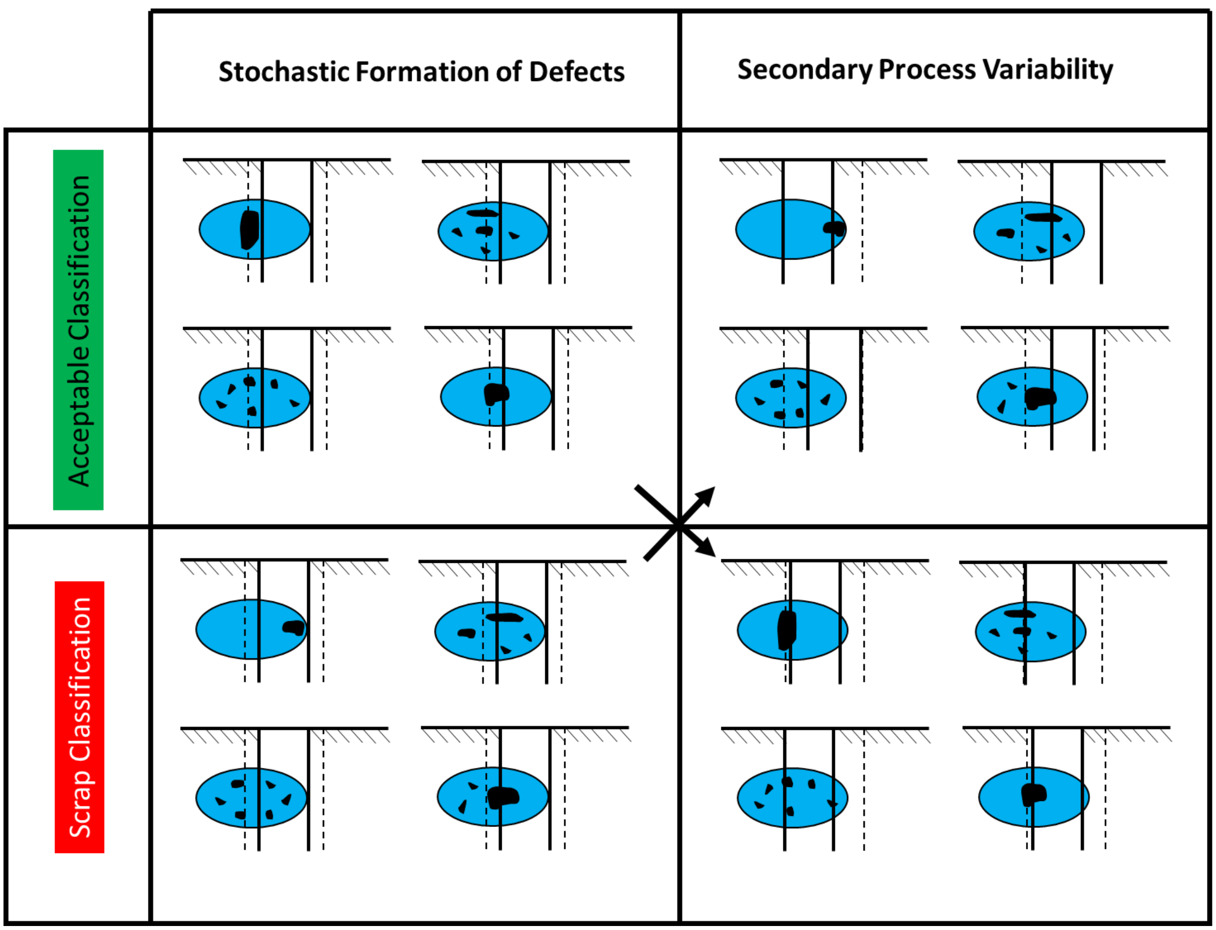


Figure 43: Combination of elements example

These two elements can also be combined with the *Visual Defect Inspection* to further complicate results. Visual inspection is dependent upon the person who is performing the task. Operator-to-operator performance of visual inspection can vary considerably [120], [142], [162]. Proper classification becomes a probability based on the chances of porosity forming, if the machining opens it up, and whether the inspection catches the defect.

All four of these elements change through time. As previously mentioned, operators who perform the inspection task will change shift-to-shift but also will likely change with turnover. New inspectors face a learning curve of defect identification while simultaneously fighting off the repetitive nature of the work. The probability of detecting defective castings changes through time. Unfortunately, the validity of the classification results is unknown as one looks at historical data.

Specification thresholds can also change based on new suppliers or additional testing. Perhaps a different vendor can allow a slightly larger porosity specification because its o-ring is improved. Or a new part failure has shown the product is used in ways it was not designed. Now a maximum porosity size previously accepted could be rejected. To build large data sets for ML, this data would need to be consistent through time. This knowledge is lost with the binary classification of scrap. The previous data becomes useless for supervised ML.

Finally, manufacturing processes vary through time. Both the casting process and machining process change based on equipment maintenance, tool wear, tool replacement, die changes, or process setting improvements. To provide a detailed example, consider tool changes in machining. Part-to-part variation in machining tolerances are often very small, given the repeatability of modern machining equipment. However, once a tool breaks and is replaced by a new tool, there is a functional change in the manufacturing system. Provided the new machined dimension falls within the designed tolerances, manufacturing will proceed without a second thought even if the new dimension is a step change from the previous tool. This changes the dynamic of the porosity exposed and the classification of the part through time.

These four elements of *Binary Acceptance Specifications, Stochastic Formation of Defects, Secondary Process Variation,* and *Visual Defect Inspection* all influence the final classification of a part. As discussed, many of these elements can create an overlap within the data space, which ML will struggle to produce meaningful predictions. This struggle can be compounded by the highly unbalanced data sets that often exist in manufacturing.

A cursory review of current operational practices would suggest manufacturers are collecting the needed data to apply ML in production manufacturing settings. However, without understanding how the ML algorithm interprets the results, the user will struggle to gain the promised value from ML technology. The industry may have very clearly defined specifications, but if it only collects acceptable/scrap and not the actual size or clustering of the void, valuable information for ML is lost. A surface is machined and then visually inspected. However, the actual truth is unknown since an operator cannot see below the surface to understand if the void still exists in the casting. Also, dimensions and locations of features are not captured on 100% of the product, so that data does not exist for ML to utilize. Manufacturers may feel good they inspect 100% of castings to ensure only good parts get to their customer but fail to realize the fallacy of human inspection rates in repetitive visual tasks.

In the end, it is not the complexity of ML technology and implementation that fails the manufacturer. Instead, the data the manufacturer has put considerable effort in to specify, create, and inspect parts to prevents the ML from being successful. The next section will provide insight into how the bias-variance tradeoff within ML is influenced by these misclassifications and the importance of a critical error threshold with highly unbalanced data that exists in manufacturing. This will help a user understand the financial impact and ML accuracy levels required for successful implementation of ML and why they often fall short in providing value in manufacturing.

IMPACT ON SUPERVISED MACHINE LEARNING

Supervised machine learning algorithms are trained on input data and the known results to find patterns to accurately make predictions [163], [164]. There is limited ability to overcome poor training data that does not represent the prediction trying to be made. Accurate classifications are critical to successful ML applications [165], [166]. The impact of misclassification of results will be reviewed within this section.

BIAS-VARIANCE TRADEOFF

The Bias-Variance tradeoff graph is often used to visualize the error associated with supervised machine learning models. The generalization error of a ML model is comprised of three components: inherent error, bias, and variance, as described in Equation 12 [164]. The generalization error is the error produced by the model when applied to independent test data.

Equation 12 $$\boldsymbol{Generalization\ Error = Inherent\ Error + Bias^2 + Variance}$$

The generalization error is used to select the tuning parameters during the training of a supervised ML algorithm and even which choice of learning algorithm used for a specific data set [164]. The goal is to avoid underfitting and overfitting to the training data. This is done by minimizing the squared bias component while ensuring the model can be generalized to future data as represented by the variance. The inherent error is error that exists based on the data used and will be discussed in detail later. This bias-variance tradeoff can be seen in Figure 44.

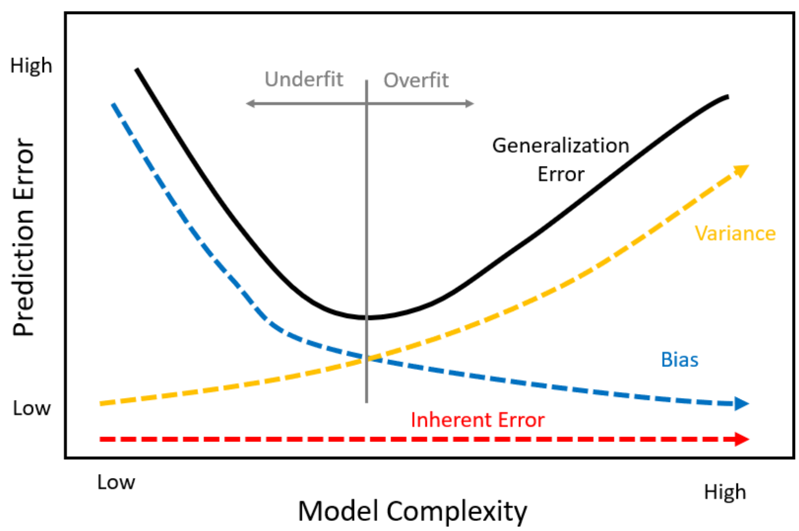


Figure 44: Bias-Variance tradeoff graph

The accuracy and the generalization error of the ML model are related as seen in Equation 13 where Accuracy is the fraction of samples correctly classified, with a value of 1 representing perfect classification. When there is a high level of accuracy in the model, then there is a low generalized error rate.

Equation 13 $$\boldsymbol{Generalized\ Error\ Rate = (1 - Accuracy)}$$

CONFUSION MATRICES

A confusion matrix is used in ML classification problems to understand how well the trained model performs. The confusion matrix shows the accuracy of the predictions made by the ML model in comparison to the true conditions of the test data. If the ML model performs well, it will correctly identify True Positives (TP) and True Negatives (TN), while only having a small number of False Positives (FP) and False Negatives (FN). The organization of the traditional confusion matrix is seen in Figure 45.

Traditional Confusion Matrix (counts)		True Condition		Predicted Totals
		Positive	Negative	
Predicted Condition	Positive	True Positive (TP)	False Positive (FP)	TP + FP
	Negative	False Negative (FN)	True Negative (TN)	FN + TN
True Totals		TP + FN	FP + TN	Total = TP + FP + FN + TN

Figure 45: Traditional confusion matrix based on counts

A confusion matrix can be normalized when the individual counts, such as TP, are divided by the overall total count. This normalization process allows the user to understand the percent predictions the model makes in each of the categories. The details of the normalized confusion matrix are seen in Figure 46. This normalized percentage for the TP, FP, FN, and TN is carried forward in the balance of the equations presented in this work. The values used in equations are now labeled with a percent sign (%) to indicate they are a fraction and no longer a count. This approach was taken because the discussion focuses on error rates and percentages and not a count as traditionally used in a confusion matrix.

Normalized Confusion Matrix (percent)		True Condition	
		Positive	Negative
Predicted Condition	Positive	TP % = TP / Total	FP % = FP / Total
	Negative	FN % = FN / Total	TN % = TN / Total
where TP % + FP % + FN % + TN % = 1.0			

Figure 46: Normalized confusion matrix based on percentage

Figure 47 was created to help illustrate a generic example of how the counts of a model as shown in Figure 45 are calculated as percentages in a normalized confusion matrix as seen in Figure 46.

Example Calculations

Traditional Confusion Matrix		True Condition: Positive	True Condition: Negative	Predicted Totals
Predicted Condition	Positive	870	15	**885**
Predicted Condition	Negative	55	60	**115**
True Totals		**925**	**75**	**1000**

Normalized Confusion Matrix		True Condition: Positive	True Condition: Negative
Predicted Condition	Positive	87.0%	1.5%
Predicted Condition	Negative	5.5%	6.0%
where 87% + 5.5% + 1.5% + 6% = 1.0			

Figure 47: Example calculations on traditional and normalized confusion matrices

The concept of measuring the accuracy of a ML model can be hotly debated. Multiple types of accuracy measures exist to summarize the model in different ways. Traditional accuracy, balanced accuracy, F1 score, or Matthews correlation coefficient are some of the most commonly used accuracy metrics [167], [168]. A shortcoming of the traditional accuracy measure is evident in highly unbalanced data. These unbalanced data sets often exist in manufacturing. A high traditional accuracy value may give a user a false sense about how well the model performs. For example, a recorded accuracy of 95% may lead one to believe the model is "good". In fact, this could be a result of the model predicting 100% acceptable product and the product having a 5% scrap rate. In this case, the prediction provides no value, even though its 95% accurate. Other accuracy calculations provide a different perspective on the overall model accuracy for unbalanced data. A theoretical example in Figure 48 highlights the value differences in the accuracy metrics.

Accuracy Calculation Example		True Condition		Predicted Totals
		Positive	Negative	
Predicted Condition	Positive	948	48	996
	Negative	2	2	4
True Totals		950	50	1000

$$Traditional\ Accuracy = 95\%$$
$$Balanced\ Accuracy \cong 72.6\%$$
$$F1\ Score \cong 97.4\%$$
$$Matthews\ Correlation \cong 13.1\%$$

Figure 48: Example of different accuracy calculations

The best calculation of model accuracy is not for this paper to help decide. Instead, consistency must occur when comparing a decision made before and after the implementation of ML. The traditional accuracy approach is used in this work for several reasons. First, this metric was selected because it is commonly used in industry even with its shortcoming for unbalanced data. Additionally, it is expected the ML user reviews the entire confusion matrix and not just an accuracy statistic calculated to truly understand the prediction of the ML model. Finally, the traditional accuracy creates a situation where the math is simplified and provides insight into creating the critical error threshold equation.

Prior to ML implementation, the decision automatically made in manufacturing is all product is acceptable, so there are no predicted negative conditions. The product with actual negative conditions is rejected after additional processing. These are the false positives created by assuming all product is good. After ML implementation, the costs associated with False Positives (FP%), False Negatives (FN%), and True Negatives (TN%) all must be considered. The traditional accuracy calculation for normalized confusion matrices is used and can be seen in Equation 14.

Equation 14

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{TP_{\%}+TN_{\%}}}{\boldsymbol{TP_{\%}+TN_{\%}+FP_{\%}+FN_{\%}}} = \frac{\boldsymbol{TP_{\%}+TN_{\%}}}{\boldsymbol{1}} = \boldsymbol{TP_{\%}+TN_{\%}}$$

$\boldsymbol{where}$:
$\boldsymbol{TP_{\%} = Percentage\ of\ True\ Positives\ as\ decimal}$
$\boldsymbol{FN_{\%} = Percentage\ of\ False\ Negatives\ as\ decimal}$
$\boldsymbol{TN_{\%} = Percentage\ of\ True\ Negatives\ as\ decimal}$
$\boldsymbol{FP_{\%} = Percentage\ of\ False\ Positives\ as\ decimal}$

With Equation 14, the generalized error rate in Equation 13 can be simplified to focus on the incorrect predictions of the ML model as shown in Equation 15.

Equation 15

$$\boldsymbol{Generalized\ Error\ Rate = (1 - Accuracy) = 1 - (TP_{\%} + TN_{\%}) = (FP_{\%} + FN_{\%})}$$

CRITICAL ERROR THRESHOLD (CET)

A concept not found in literature to date, but of great importance to applications of ML in manufacturing, is understanding a *critical error threshold* (CET) exists for the generalized error in the bias-variance tradeoff graph. This CET defines applications of trained ML that provide value to the organization to implement. Figure 49 shows the CET updated bias-variance tradeoff graph. Financial value is achieved by the manufacturing organization if a model's generalized error falls under the CET.

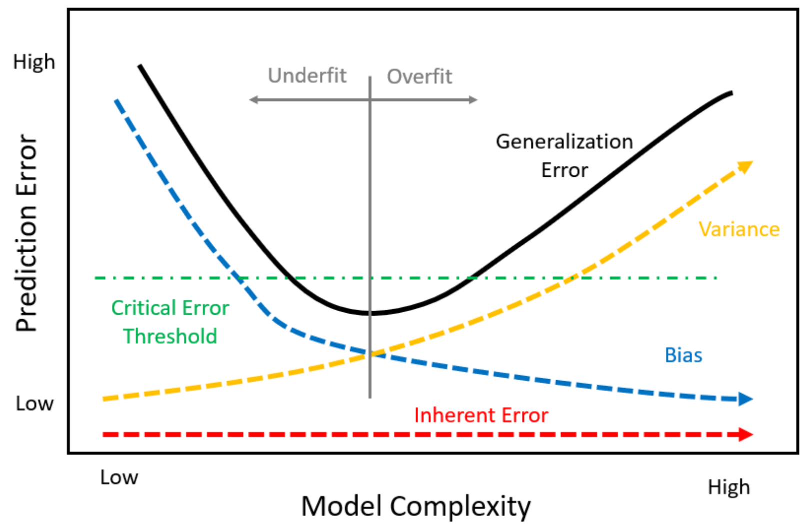


Figure 49: Bias-Variance tradeoff graph with Critical Error Threshold

CET is a financial review of the accuracy associated with a machine learning model. The prediction accuracy of the ML model, the costs associated with the scrap, and the cost to implement must be justified when compared to current scrap costs as seen in Equation 16.

Equation 16 $$\boldsymbol{Total\ Scrap\ Cost\ Before\ ML > Total\ Scrap\ Cost\ After\ ML + Cost\ of\ ML}$$

With details regarding specific current scrap performance, scrap costs, value-add costs, volumes, cost for ML implementation, and the percentages from the normalized confusion matrix, Equation 16 can be expanded to provide additional analysis in defining the CET as seen in Equation 17. It is important to note that the value-add costs could represent more than scrap costs added to the part before the defect is uncovered. This variable can also represent additional inspection costs, schedule adjustments, additional setups, warrantee claims, and other items that drive manufacturing costs due to defective products in the supply chain.

Equation 17

$$\boldsymbol{CET} \rightarrow \boldsymbol{Accuracy\ of\ the\ ML\ model\ in\ terms\ of\ FN_{\%}, TN_{\%}, FP_{\%}\ is\ such\ that}:$$
$$(\boldsymbol{S_{\%}})(\boldsymbol{C_{\$}} + \boldsymbol{V_{\$}})(\boldsymbol{EAU}) > (\boldsymbol{FN_{\%}} + \boldsymbol{TN_{\%}})(\boldsymbol{C_{\$}})(\boldsymbol{EAU}) + (\boldsymbol{FP_{\%}})(\boldsymbol{C_{\$}} + \boldsymbol{V_{\$}})(\boldsymbol{EAU}) + \boldsymbol{ML_{\$}}$$
$$\boldsymbol{where}:$$
$$\boldsymbol{CET = Critical\ Error\ Threshold}$$
$$\boldsymbol{S_{\%} = Current\ Scrap}\ \%\ \boldsymbol{of\ Casting\ as\ decimal}$$
$$\boldsymbol{C_{\$} = Casting\ Scrap\ Cost\ per\ part}$$
$$\boldsymbol{V_{\$} = Processing\ Value\ Add\ per\ part}$$
$$\boldsymbol{EAU = Estimated\ Annual\ Usage\ or\ Volume}$$
$$\boldsymbol{FN_{\%} = Percentage\ of\ False\ Negatives\ as\ decimal}$$
$$\boldsymbol{TN_{\%} = Percentage\ of\ True\ Negatives\ as\ decimal}$$
$$\boldsymbol{FP_{\%} = Percentage\ of\ False\ Positives\ as\ decimal}$$
$$\boldsymbol{ML_{\$} = Annualized\ Machine\ Learning\ Implementation\ Cost}$$

If there is no optimization within the manufacturing process with the implementation of the ML model, then the scrap rate before ML must equal the scrap rate after model implementation. As per the normalized confusion matrix, this scrap rate is the summation of the True Negative condition, regardless of prediction as shown in Equation 18.

Equation 18
$$\boldsymbol{S_{\%} = TN_{\%} + FP_{\%}}$$

The generalized error rate and CET relationship exist per Equation 19. There is motivation to implement ML in manufacturing when the generalized error rate is below the CET.

Equation 19
$$\boldsymbol{Generalized\ Error\ Rate} < \boldsymbol{CET}$$

If no optimization has happened with the implementation of the ML model as defined in Equation 18, the CET value can be simplified as seen in Equation 20. This is accomplished by solving for the ratio between the right and left hand of Equation 17 and substituting the generalized error rate from Equation 15 and the scrap percentage as per Equation 18.

Equation 20
$$\boldsymbol{CET = TN_{\%}}\left(\frac{\boldsymbol{V_{\$}}}{\boldsymbol{C_{\$}}}\right) + \boldsymbol{FP_{\%}} - \frac{\overline{\boldsymbol{ML_{\$}}}}{\boldsymbol{C_{\$}}}$$
$$\boldsymbol{assuming\ no\ scrap\ optimziation, where}:$$
$$\overline{\boldsymbol{ML_{\$}}} = \boldsymbol{Annualized\ ML\ cost\ per\ EAU}\ (\boldsymbol{volume\ normailzed})$$

The values for FN and FP will vary by tuning and training of the ML algorithm. The casting cost and value-add costs for defects are dependent on the part and processing. Slight improvements in predictions could yield financially favorable results for highly expensive value-add costs. If the value-add costs are low, the accuracy of the model is poor, or the cost to implement the ML is expensive, then value gained by implementation will not be justified. The generalized error must fall under the calculated CET for the organization to achieve a financial benefit. There is no motivation to adopt and implement ML within manufacturing if the accuracy of the model prediction does not yield financial benefits. Therefore, it is important to understand how the inherent error and squared bias affect the generalized error and implication of the CET.

INHERENT AND BIAS ERROR

In Equation 12, the inherent error exists because the data used for the prediction does not completely represent the ability to predict the output [164]. There likely exists additional input variable data that would improve the prediction on the model. To reduce the inherent error, these additional variables would need to be identified and included as predictors in the model. Within manufacturing, the impact of the inherent error may be large. The data collected may struggle to accurately train the model because critical variables are not included. This would increase the inherent error, thereby increasing the generalized error of the prediction. Additionally, by not including the critical parameter, the data space that exists will likely overlap since the true cause for the results are not known. If the generalized error is increased above the CET, as seen in Figure 50, the value of the ML model is not worth the investment.

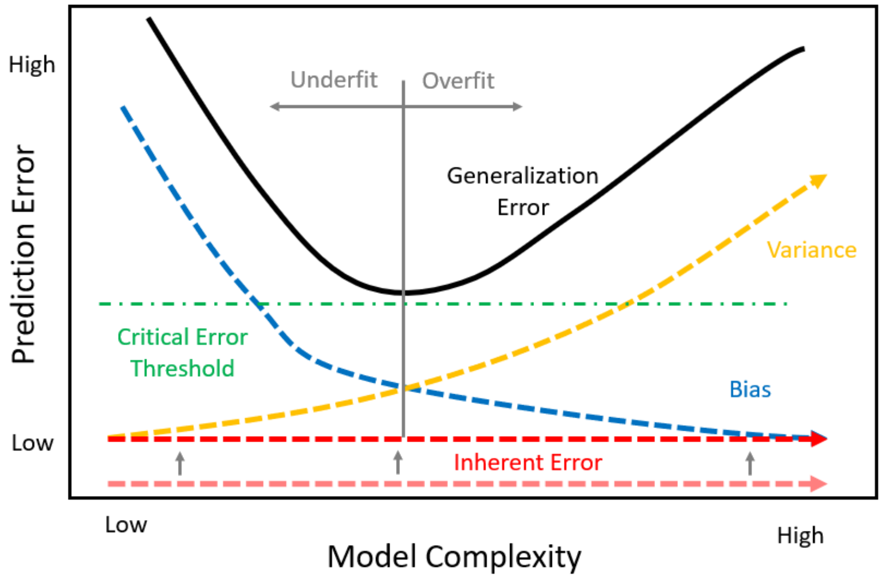


Figure 50: Increased inherent error

The squared bias term from Equation 12 is described as the amount by which the mean model estimate based on the trained data differs from the true mean [164]. Typically, a more complex model created from the training data means a lower squared bias term. The model will be overfit to the training data. When applied to other data beyond the training set, it will perform poorly. This performance is captured in the variance term.

The squared bias term is directly related to the four elements of classification issues identified within this paper. By misclassifying the training data, the squared bias component shifts upward, increasing the generalized error of the model. The bias term in the graph could be thought of as approaching an asymptotic limit of the misclassification rate associated with the training data. The larger the rate of misclassification within the training data, the higher the shift upward in the bias term and the resulting generalized error of the algorithm. If this shift up is substantial enough, the generalized error may exceed the CET as seen in Figure 51.

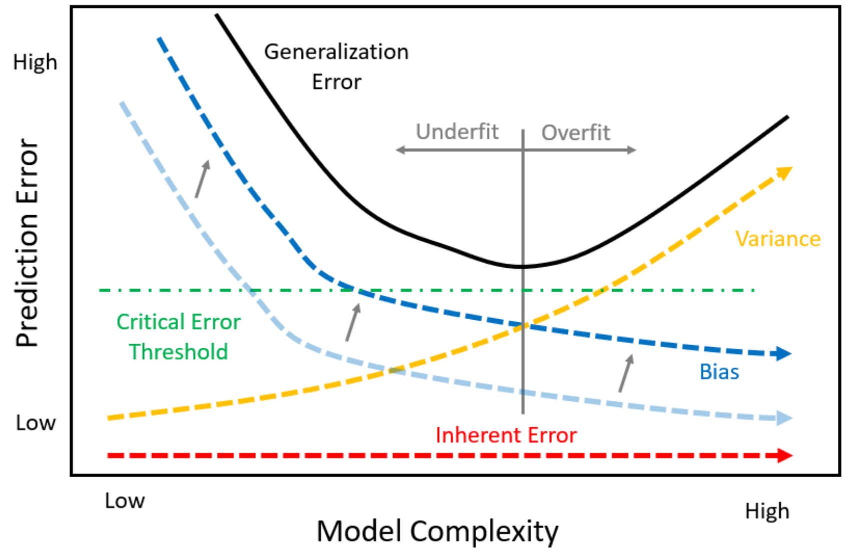


Figure 51: Increased bias

Although machine learning is a powerful tool to help find patterns, the data it is trained on must represent the critical parameters required for a good prediction to eliminate data space overlap (inherent error) and classified correctly to provide proper training (bias). The challenge for supervised ML in manufacturing is addressing both items. Data collected in manufacturing is limited and may not capture the total data space needed for an accurate prediction. The probability all the results are properly classified, especially from visual inspection, is poor. The CET calculation could provide a benefit if the ML is tuned to focus on false positives or false negatives, provided there is an economic advantage with the scrap or value-add costs.

This is further compounded by the highly unbalanced nature of manufacturing data. Consider an example in a ML algorithm created to improve the click-rate for web advertising. Success in this application could be achieved by increasing from a 2% click rate to a 5% click rate. This can also be considered as utilizing ML to move from a 98% error rate to a 95% error rate. Getting under the CET may be easy in this advertising case. Compare this to a manufacturing example where the success of product being produced is theoretically 95% with a 5% scrap rate. The goal of ML is to increase this to a near perfect prediction. Because quality yield rates within manufacturing are typically good, the ability to implement ML requires perfect data collection and process knowledge. It should not be a surprise to see this technology fail to produce meaningful results in many manufacturing settings. The generalized error rate is above the CET needed due to this imbalance of the data, missing data, and misclassified results.

RECOMMENDATIONS AND FUTURE STUDY

The work presented here should challenge the typical approaches for classification and use of supervised ML within the foundry and manufacturing industries. A systems approach is needed to gather all data parameters within the process to minimize the generalization error. If the CET is crossed, the supervised machine learning model adds no financial value to the operation. There are three recommended areas of study or improvements needed.

SYSTEMS APPROACH TO PROCESS DATA (INHERENT ERROR)

Although briefly discussed in this paper, there is a critical importance in reducing the inherent error for the model. Those interested in applying supervised ML in manufacturing must ensure the proper input variables are included in the model. The inherent error within manufacturing is likely a large portion of the generalized error. Limited data collection systems exist for the entire manufacturing system. There is a long-standing history with shot injection parameters within HPDC. However, the industry lacks commercially available or widely adopted data collection systems for thermal balance of the tooling, cycle time analysis of the entire process, and overall machine equipment performance. Even though the importance of thermal equilibrium in the tool is known [94], the industry is left to make assumptions and focus on what is easily measured. Quality performance of an 8% median scrap rate, as described at the beginning of this paper, highlights the results of current practices.

Correct data and process understanding is needed to reduce the inherent error. Adding variables that help pattern the prediction and eliminating variables that do not contribute will reduce the generalized error of the model. Understanding the entire die casting system is the first step to know all the process data available. A systems approach is needed to review and document all possible parameters associated with the process. There are many sub-systems within die casting creating data to be collected and studied. Some of these sub-systems include the thermals of the die, the variation in the lube spray system, and the equipment performance. This is an area of research that needs continued focus to ensure the correct variables and data are available for ML models. The data collected within the industry to date has not solved the quality problems die casting faces. Additional research is needed in this area to document all sources of variability introduced into the HPDC process.

FOUR ELEMENTS OF CLASSIFICATION ISSUES (BIAS)

As described, there are four key elements that impact classification of casting: Binary Acceptance Specifications, Stochastic Formation of Defects, Secondary Process Variation, and Visual Defect Inspection. These misclassified results create data space overlap thereby shifting the bias of the model and potentially making the model useless. This challenge is complicated further with the highly unbalanced data sets associated with manufacturing and the CET associated with the process.

To ensure bias is reduced, the classification of the result variables in the data collected must be accurate. This will be a challenge and must be addressed through equipment and training. An accurate model requires eliminating the binary classification of scrap. Acceptable and scrap classifications will not be sufficient. Even multiple classifications for different types and location of scrap have limited benefit given the continuous nature of the random defect formation and secondary processing.

Technologies like in-line X-ray equipment must be considered in applications where ML is applied. The bias associated with ML models would be improved by knowing the amount and location of the defect 100% of the time. This would also eliminate the delay of quality results from machining to the foundry. Additionally, it could reduce the human visual inspection of X-ray images if technologies like automatic defect recognition are used.

Without investment in X-ray equipment, the largest component of classification issues given the historical performance researched, is the visual inspection of final product. If only 50% to 80% of the defects are properly identified, a portion of results are being misclassified as acceptable. The ML algorithm cannot overcome this overlap. Additionally, ML is faced with challenges regarding highly unbalanced data sets from the manufacturing process. As a result, considerable effort must be given with defect inspection. Texts regarding how to improve visual inspection for castings are published and provide guidance to train and improve operators [120].

Because humans are fallible, the inspection task should also be investigated for automation. Computer vision systems utilizing ML present opportunities to automate a large portion of this visual inspection. Even with technology limited to 2-dimensional surfaces, implementing this can reduce the required workload on human inspectors. This can allow inspectors to focus just on hard to capture areas, such as holes where visions systems may not be as successful. In addition, the resultant images gathered from these vision systems would be combined with the process data to provide additional ML opportunities.

UNSUPERVISED ML AND FEATURE IMPORTANCE

The challenges with supervised machine learning are considerable due to the limitations of data collection of the process (inherent error) and misclassification of training data (bias). Unsupervised ML is an approach that can provide value in manufacturing. Unsupervised ML is learned on data inputs without any knowledge of the results. It typically focuses on clustering and anomaly detection algorithms. Data processing and process control are two areas where unsupervised ML can benefit manufacturing today.

Manufacturing processes are highly complex systems. If collected, the volume and velocity of data that can be produced by equipment is extremely large. Manufacturing can produce more columns of data with the complex system than individual rows of events. In HPDC, it has been estimated that there are hundreds of thousands of possible columns due to the time-series nature of some variables used such as speed, pressure, and temperature [2]. Often reducing these time-series down to an overall statistic loses critical information. The use of unsupervised ML to complete anomaly detection is highly advantageous. Unsupervised ML can be used to help identify when process data [45], images [9], and time-series data [169] are outside normal ranges.

A human operator working in the manufacturing cell could not handle the level of data the cell produces. A computer with ML could manage the process control analysis and alert the operator when there are anomalies. The anomaly thresholds are set by the ML algorithm based on past performance of the process much like the control limits are set on a statistical process control (SPC) chart. Traditionally, shot monitoring system limits are human driven by past operator experience or limited developmental experimentation and can be considered a specification limit for that parameter. Unsupervised ML can take a high-dimensional SPC approach to data analytics, unlike traditional shot monitoring systems. An additional benefit of unsupervised ML anomaly detection is it can apply to time-series data sets to detect changes in the entire profile of the time-series set. Traditional shot monitoring systems only monitor statistics such as average fast shot speed calculated from those profiles. This creates the potential to miss subtle changes in the profile, such as additional braking done at the end of the shot to prevent flashing, which could have consequential implications on the casting quality.

The other machine learning application that can be utilized is feature importance. Related to supervised ML, feature importance is the process of utilizing machine learning to identify which variables have the most important influence on the prediction. Feature importance can provide an advantage even by identifying the few process parameters to investigate with a design of experiment (DOE). This DOE could optimize the process and reduce the scrap even if the ML model is incapable of creating a highly accurate model. Additional study is needed on the impact of misclassification of results on feature importance, but it appears a promising ML tool to assist manufacturing applications.

CONCLUSIONS

Supervised machine learning is a powerful tool that has been successfully applied in many industries. There have been substantial advances in ML domains such as image classification and natural language processing. Comparing these to applications of supervised ML in manufacturing is like comparing apples and oranges. Both use ML, yet they are distinctly different and must be treated accordingly.

The challenges of supervised machine learning in manufacturing are significant due to classification issues and limitations in data collection. The four elements proposed (*Binary Acceptance Specifications, Stochastic Formation of Defects, Secondary Process Variation,* and *Visual Defect Inspection*) influence the final classification of a part. Misclassification creates data space overlap. This overlap alters the bias in the training of supervised machine learning, possibly rendering the model financially useless in a production environment. Understanding the critical error threshold provides economic guidance on when ML can be successfully applied.

Until manufacturing can establish system-wide gathering of process variables and eliminate classification issues, the success of supervised ML will be limited to highly controlled research or academic experiments. Much noise exists in the system today. This does not mean ML is to be abandoned, but instead different approaches are needed for manufacturing to see the benefit.

Beyond traditional uses of supervised ML, feature importance and unsupervised ML provide entry points for manufacturers looking to enter and start using machine learning. The potential time savings and guidance on critical input parameters feature importance can provide needs to be better understood and utilized within manufacturing. This could be a noteworthy savings in experimentation and the optimization process for die casters. Additionally, utilizing unsupervised ML for process control and anomaly detection allows for the use of machine learning in manufacturing while creating the foundation needed for future supervised ML. This foundation is created when a company improves its classification of parts produced (reducing the bias and overlap) and optimizes the data that could improve the prediction model (reducing the inherent error). In the end, these changes will position manufacturing to benefit from accurate predictions of supervised machine learning while obtaining an improved understanding of the process.

Chapter 5: Machine Learning Case Studies

Chapter 5 is comprised of two main sections. The first section introduces machine learning concepts for the reader and then highlights the challenges associated with applications in complex, production manufacturing systems. Although industry tends to focus on quality prediction, the challenges identified in this work and the opportunities that exist in the data should guide the production user to focus on unsupervised machine learning for data processing for anomaly detection and process control. The second section of this chapter is a set of four case studies of machine learning applied in a die casting foundry. These case studies are published works presented at the North American Die Casting Association (NADCA). The case studies review topics like process optimization, thermal image analysis, and anomaly detection of time-series data. Significant value can be created by appropriate use of machine learning within the die casting industry.

INTRODUCTION

Artificial intelligence (AI) is one of the key pillars of Industry 4.0 [92], [170]–[172]. The term artificial intelligence was coined in the 1950s defining a machine system that could learn and match the intelligence of a human [173], [174]. This generic definition still holds true today for uses of AI. Artificial intelligence consists of multiple components: language processing, knowledge, reasoning, problem-solving, and planning [174]. Today, the reasoning and problem-solving components of AI are typically referred to as machine learning. Machine learning is a subset of AI. Machine learning is often defined as the act of a computer learning specific patterns in data without being explicitly programmed [140], [141]. Applications of machine learning within die casting is the focus of this work. Machine learning is not comprised of just one method; instead there are a large range of algorithms that comprise machine learning. The algorithm that has gained much popularity recently is Deep Learning (DL), which is a neural network algorithm comprised of many neuron layers in efforts to mimic the human brain [175]. Figure 52 shows the hierarchical relationship between AI, machine learning, and deep learning.

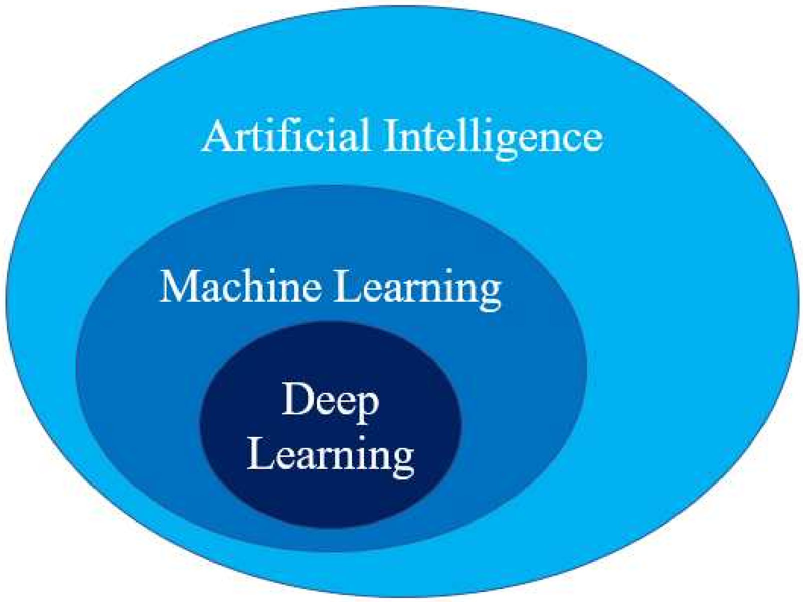


Figure 52: Artificial intelligence, machine learning, and deep learning hierarchy

Machine learning is comprised of three significant categories of algorithms: supervised learning, unsupervised learning, and reinforcement learning [140], [176]. These algorithm categories are used for different purposes. Supervised learning uses labeled data to learn the pattern in the input variables to predict future unlabeled datasets. Unsupervised learning uses unlabeled data and instead looks to cluster or group data based on input variables. The third category of algorithms is reinforcement learning. Reinforcement learning are algorithms based on an agent interacting with its environment to try to maximize a reward. Details of the three different categories of learning algorithms can be found in Table 29.

Table 29: Details for Supervised, Unsupervised, and Reinforcement Learning

	Supervised Learning	**Unsupervised Learning**	**Reinforcement Learning**
Data Type	Labeled data	Unlabeled data	Learn from environment
Types of Problems	Regression and classification	Clustering and anomaly detection	Exploitation or exploration
Goal	Predict outcomes	Discover groupings	Learn a sequence of decisions to maximize reward
Algorithm Examples	Decision Trees, Support Vector Machines, Neural Networks	K-Means Clustering, Autoencoder, Apriori	Q-Learning, SARSA, Deep Q Network

The focus of the case studies provided is unsupervised learning, although supervised learning was initially attempted in the first case study and will be discussed. There are applications of reinforcement learning published within manufacturing, such as robotic programming [177] or scheduling [178]. However, reinforcement learning has more specialized applications and fell outside of the scope of this research. Publications discussed in Chapter 1 are all supervised learning applications [17]–[22]. Supervised learning dominates the publication space of machine learning applications in die casting. Beyond the author's work presented in 2017 through 2021 in the upcoming case studies [9], [45], [169], [179], only one other publication was located within the die casting industry. Žapčević and Butala utilized unsupervised learning to data mine previous production parameters to build dynamic knowledge rules [180]. This work, however, presented questionable data that could be representative of a setup or sensor issue. The Tablet Thickness was exactly 0.0 for the first approximately 15,000 cycles of 56,000 cycles used in analysis. After approximately 15,000 cycles, the balance of the variable showed normal process variation. This leads one to believe the 0.0 data was incorrectly collected and likely should have been not included in the analysis. Regardless, the clustering helped identify an issue within the process, thereby highlighting an argument that will be made that a focus on unsupervised learning can provide value to the die casting industry.

There are many different components and definitions that need to be understood when discussing supervised and unsupervised learning. Entire textbooks have been written on these machine learning topics. The balance of this section provides a high-level overview of components and nomenclature needed when discussing machine learning.

- *Algorithm*: An algorithm is a set of instructions and logic used in machine learning programming to learn from the data presented to typically provide predictions on future data. Different algorithms have unique approaches on how to pattern data to help make predictions. Numerous algorithms exist such as linear regression, logistic regression, naïve bayes, neural nets, decision trees, random forest, support vector machines, k-means clustering, etc.

- *Input Data*: Input data is represented by the variables used to describe the process that the machine learning algorithm will try to learn from. Input data is often described as the "x-axis" variables that lead to some dependent "y-axis" output or *Results*. Another name used for input data is "features."
- *Results*: Results are the output from a process with the input data. Results are the "y-axis" output based on the "x-axis" inputs. Results are sometimes called "labeled data" since in the training data, the inputs contain a labeled output. Results can be a classification or a predicted value.
- *Data Pre-Processing*: Raw data gathered for machine learning applications often needs some level of data pre-processing before it can be used within an algorithm. Some of this pre-processing is basic such as formatting or removal of rows of data with missing values. Other pre-processing operations are more advanced, such as normalizing or scaling input values, which is required for some algorithms.
- *Training Data*: Training data is the data used to train the machine learning algorithm. For supervised learning, this data set includes both the input data and results.
- *Validation Data*: Validation data is data separated out of the training data that the algorithm does not use during the initial training. The term often used is this validation is that this data is held back from training. After training is completed, the validation data is used to help tune the hyperparameters of certain machine learning algorithms. There are multiple approaches for developing the validation data including the hold-out method, k-fold cross validation, and bootstrapping.
- *Test Data*: Test data is data that has not been seen by the algorithm during any training or validation process. Test data is the final test to see how well the machine learning algorithm performs and if it was trained effectively. It is typically called an unbiased

evaluation of the algorithm since the model has not seen any of the data prior to the testing.

- *Confusion Matrix*: A confusion matrix is used to help evaluate the performance of a classification machine learning model. The confusion matrix compares the actual target values of the test data compared to the algorithm predicted values. Ideally, the predictions will match the actuals to provide a high level of accuracy within the model. Terms true positive, false positive, false negative, and true negative are used to describe a two-class confusion matrix as seen in Figure 53. Additional details on confusion matrices can be found in Chapter 4.

<table>
<tr><th colspan="2" rowspan="2">Confusion Matrix</th><th colspan="2">Actual Condition</th></tr>
<tr><th>Positive</th><th>Negative</th></tr>
<tr><th rowspan="2">Predicted Condition</th><th>Positive</th><td>True Positive</td><td>False Positive</td></tr>
<tr><th>Negative</th><td>False Negative</td><td>True Negative</td></tr>
</table>

Figure 53: Example 2-class confusion matrix

- *Neural Net*: Neural net is a popular type of machine learning algorithm based on neural connections like the human brain. A neural net is comprised of connected neuron layers that learn to pattern the input data to make a prediction. Deep learning is a neural net with many different layers. An autoencoder is an unsupervised version of a neural net.
- *Over-Fitting*: Over-fitting is a term used in machine learning when the algorithm creates a prediction that performs extremely well on test data but does not with the test data. In these cases, the algorithm over-learned the parameters within the test data and no longer has the generalization needed to accurately predict data on which it was not trained. An example of an overtrained graph can be seen in Case Study I in Figure 66.
- *Anomaly Detection*: Anomaly detection is the process of finding outliers within a data set. Anomaly detection can be accomplished through supervised learning if the result labels are known or via unsupervised learning if anomalies are unknown within the data set.

Anomaly detection can provide insight and control into the system that is generating the data.

With a basic understanding of critical machine learning terms and concepts, a review of machine learning in complex and challenging manufacturing systems is needed. After understanding these challenges, approaches for utilizing machine learning in production die casting applications will be reviewed before the case studies are presented.

CHALLENGES OF MACHINE LEARNING IN DIE CASTING

Implementing machine learning for any type of application can be challenging. Success is never guaranteed. As the complexity of a system increase, the challenges of success follow suit. With a systems engineering mindset, by breaking down the difficulties and risks one expects to see within the system, actions can be taken to mitigate the potential challenges. It is important to understand the technical and non-technical challenges associated with applying machine learning in die casting. Based on publications within manufacturing and experience gained through the last five years, there are six key challenges faced when trying to implement machine learning in die casting. The six challenges are seen in Figure 54 and described in detail thereafter.

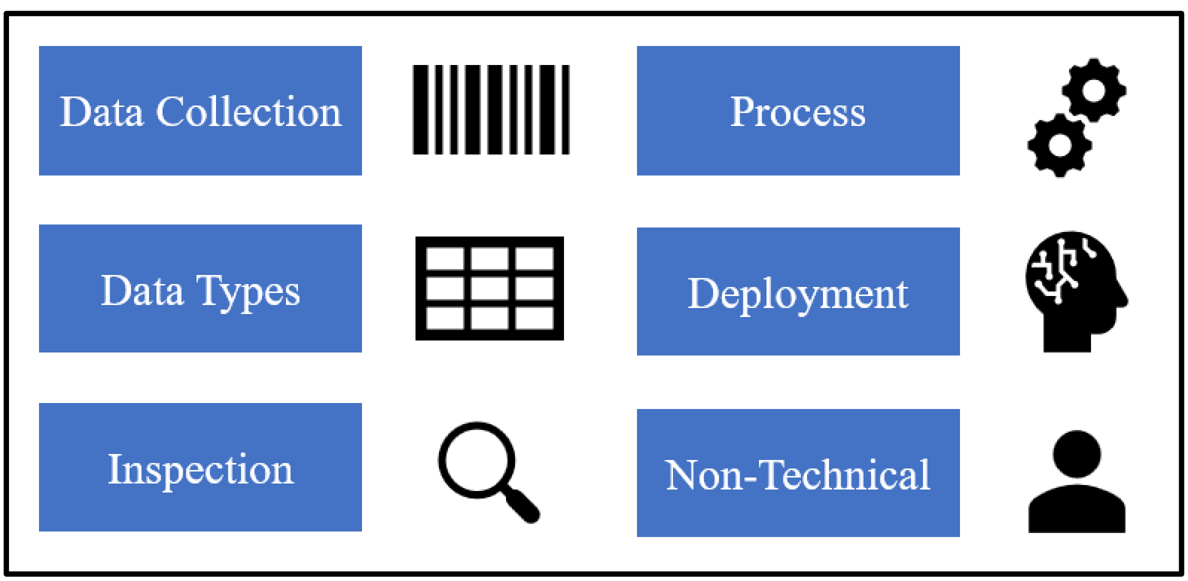


Figure 54: Six challenges of machine learning in die casting

DATA COLLECTION

Collecting data from manufacturing is the basis of Industry 4.0 and is a prerequisite for any machine learning application. Die casting is a data-rich environment as shown by the data framework described in Chapter 2. Collecting the volume and velocity of data generated within die casting comes with challenges. These challenges present in three different phases. The first is the physical act of collecting the data from the equipment. The second phase is the traceability of the product throughout the life cycle. The final phase is ensuring data is stored and preprocessed, so it can be linked and accessible.

Communication protocols that have become popular with Industry 4.0 have simplified the process of connecting equipment within manufacturing plants. OPC UA (Open Platform Communications Unified Architecture) and MQTT (Message Queuing Telemetry Transport) are two popular types of industrial communication protocols, although others exist [181], [182]. Determining which protocol works for both current and future equipment is one of the first decisions a foundry must make. This is a balancing act, as the age and type of equipment may make these connections difficult [27]. Additionally, having resources with the knowledge to implement this technology along with machine learning is often a constraint within manufacturing [7]. Additionally, selecting and then connecting the sensors, hardware, and network equipment provides more decisions and effort that must be undertaken prior to collecting manufacturing data. Cyber security adds another layer of challenges within this entire process. Implementation of cyber security within Industry 4.0 is critical to reduce the threat of an intrusion, loss of data, or ransomware attack [172], [183]–[185]. Finally, the last challenge with collecting the data is ensuring it is the right data that will provide value within a machine learning prediction. This can be the most challenging, as often companies focus on the data that is easy to collect versus what may be controlling the process [5], [27], [186], [187].

Assuming the initial data can be collected and stored as the product is created, the next challenge is traceability and collection of data within the lifecycle of the manufactured part. Serialization and traceability become key components for tracking results and labels throughout the processing. A unique identifier on the manufactured part is the initial step required to create a training dataset with results.

Additionally, time and effort must be taken to collect and store the inspection and processing feedback of the part throughout the supply chain. These efforts are considered large enough that a separate section entitled “Inspection” was created. The process of creating a serialized number on a die casting can be a difficult task. The location must be agreed upon by the designers, so it will not interfere with any part function. The foundry team must agree, so the casting can be put into the proper orientation within the extraction cell to be marked. Also the machinists must agree as they will need to read the serial number to collect quality and inspection data on the casting as it is processed. Oftentimes, castings are shot blasted and painted, so having a serial number survive and be read is a necessity [8]. Figure 55 shows an example of a 2D laser barcode etched on a die casting. Serialization and traceability provide additional technical hurdles that need to be overcame to have data available for machine learning.

Figure 55: Example 2D laser barcode on a die casting
(photo permission from Mercury Marine)

The final challenge of collecting data comes with storage and preprocessing. Data storage presents a host of additional decisions a company needs to make. One of the problems that exist is there is a plethora of solutions that exist for data management within a company [186]. Determinations must be made on types of data storage (SQL, JSON, BLOB, or Oracle) and location of the data (cloud based, server based, or local) [188], [189]. Data accessibility from these locations must be selected, and often a hierarchy is built to accommodate hot, warm, or cold storage of data [190]. Additional challenges come from data linkage. Value is created when data is connected. For machine learning to work, the data from the casting variables stored in the foundry's database must be able to connect to the results from the postprocessing of the casting. If this is one company, this may not pose a challenge. However, with a typical supply chain, there may be two or more companies involved with the casting, shot blasting, painting, and machining of a casting. If this data cannot be connected and supplied, then supervised machine learning cannot be implemented as no results will exist. Additionally, supply chains like this can take significant time to process castings. Month-long delays from casting to machining result feedback are commonplace. Finally, as data is being stored, and for sure before being used in machine learning algorithms, data often needs to be preprocessed [171], [188]. Data preprocessing includes elimination of missing data, determining mislabeled data, dealing with sporadically labeled data, detecting outliers, and normalizing data [5], [7].

Mayr *et al* suggests that data preparation takes up to 70% of time in a machine learning project [27]. The time commitment alone on the collection, storage, and processing of the data provide many challenges for implementing machine learning within manufacturing. When traceability and serialization are added throughout a multi-company supply chain, it becomes even more difficult.

DATA TYPES

As stated previously, collecting the right type of data is important to successfully implement machine learning. Die casting creates several different types of data in both the information generated in the process and the size of terms of length and width of the data set. Additionally, die casting produces highly unbalanced data sets, which will be discussed.

As shown in Chapter 2, die casting data with more than hundred thousand data points per cycle produces an extremely wide dataset. Die casting data suffers from Bellman's *curse of dimensionality*, which makes applications of machine learning more difficult [92], [191]. Especially for supervised learning, if the training data set is extremely wide but not long, there will not likely be enough rows of data to determine which features are the important ones to use to pattern from. This can lead to poor performing algorithms when tested in real life. The Hughes Phenomenon is named after Gordon Hughes. In his publication, Hughes stated “with a fixed design pattern sample, recognition accuracy can first increase as the number of measurements made on a pattern increases, but decay with measurement complexity higher than some optimum value” [192]. This phenomenon is visually depicted in Figure 56.

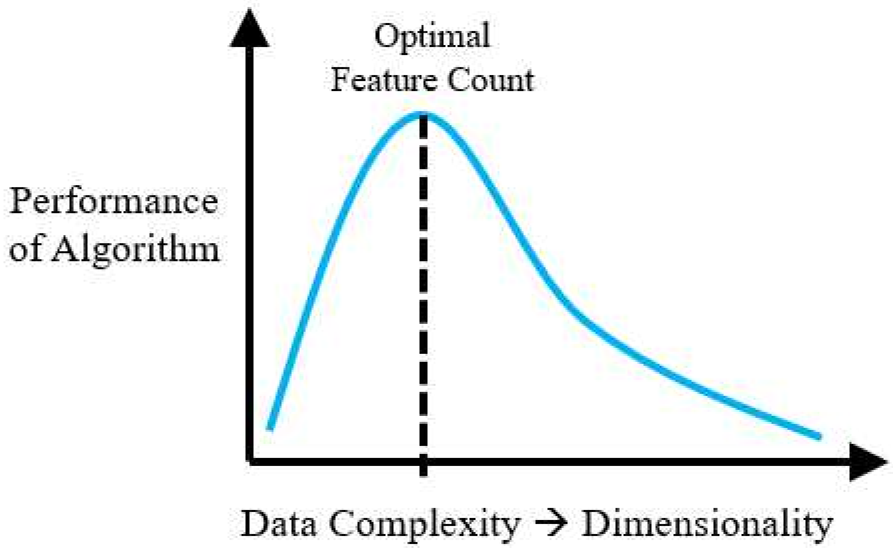


Figure 56: Hughes Phenomenon visually depicted

Additionally, with a fixed number of rows in the training dataset, additional features included reduce the fraction of samples in any given input space due to the increasing number of dimensions [193]. If there are not enough rows of data for each given example, the algorithm will not be able to determine which features are the important ones to use to pattern.

The curse of dimensionality would go away if enough additional samples are collected that the dataset ends up longer than the length of all the features included. However, this too poses a significant challenge in die casting. Although die casting is often tied to high-volume production, the datasets that are created are not the “big data” that is typical of machine learning applications. Sun *et al* describe research performed on datasets in the metal casting industry only producing 300 to 7,000 rows of data per year for a given part number [7]. Even with multiple years of data, these parts will never produce big data. In some automotive applications, the potential for hundreds of thousands of rows of data exist [85], but these could be spread out across multiple machines and tool designs with meaningful differences between them. Machine learning applications within the die casting environment will need to find success in “small data” without the luxury of millions of rows of data.

Part of the challenge with the huge width of data available is determining which data is important to solve the problem at hand. Tools exist for data reduction and feature engineering to help reduce the curse of dimensionality, but those items also bring with them additional resources, machine learning knowledge, and possible lack of understanding of how to optimize the results given the data reductions on features included. Making an accurate prediction with machine learning is the goal and beneficial, but the end state for manufacturing is using that prediction to help optimize the process to never have the negative state occur again. Understanding the type of data is important within die casting. The data presented in Chapter 2 was classified into five categories: design parameter data, input settings data, output – discrete data, output – time series data, and cycle time data. It is important to state that even though the data framework identifies hundreds of thousands of potential data points, it is not expected that a user incorporates all this data into one model. In fact, to reduce the challenges associated with machine learning in die casting, understanding the problem will point to the data that is needed. If one is trying to predict quality, then the output data types would likely be the focus. However, the data provides many other opportunities for machine learning, specifically in the predictive maintenance arena and with anomaly detection for process control. This will be discussed in more detail later in this chapter. The key point though, is one of the challenges of applying machine learning in manufacturing is understanding what data is available and how that data can be used to solve the problem at hand. Subject matter experts are more important in this conversation than data scientists programming the algorithms.

Finally, the last challenge associated with die casting data is the unbalanced nature of the results, specifically for quality predictions. For manufacturers, it is a financial positive they can produce such a high yield within their process. However, from a machine learning perspective, this greatly skews the number of samples of defective parts that an algorithm can use to train [5], [7], [133], [194]. When this is combined with smaller data sets, and larger number of data features, the unbalanced nature can be overwhelming. Some have used SMOTE and other oversampling techniques to train algorithms in this unbalanced area [7]. Having highly unbalanced data sets is challenging, but what compounds this problem is the inspection process used to label the results used in the algorithm.

INSPECTION

Another important challenge when implementing machine learning in manufacturing is the inspection of product through the supply chain. Inspection determines the classification of the given part. This classification becomes the label for the dataset, provided the serial number of an individual casting is included with the inspection results. If results are not recorded, the opportunity for supervised machine learning does not exist.

Many different potential failure points exist in the inspection process. Results data will not exist if serialization information is not on the casting. The serial number marking must survive any post processing operation of the casting. Then, this serial number must be read and the appropriate inspection result recorded. Manufacturing facilities often track cases of defective product, but machine learning also requires the tracking of all product. This is needed so classification algorithms can also know the data patterns that exist for the good castings as well as the defective ones. This approach is not typical of the industry. Additionally, the entire serialization of castings is almost nonexistent within the industry. From a 2014 NADCA study, 55% of foundries surveyed store the injection parameter data from the process. Only a small number of these foundries identify the casting with a serial number [4]. One of the biggest challenges the industry has is serializing the data to link results to parameters. The final potential failure point is the connection of the process data collected at the casting creation with this inspection data. Often this data exists in multiple databases, and in many cases, this data could belong to two separate companies.

Merging the process data with the results data must happen for successful supervised machine learning. Beyond technical challenges of traceability with serial numbers and linking different databases, time creates additional difficulties or delays in preparing data for machine learning. Inventory management through the supply chain becomes an important part of building complete data sets. Since die casting is typically a batch process [1], [13], foundries could run weeks to months of inventory ahead of post-processing operations. This means it takes weeks to months to start getting initial data available to merge between the process parameters and the results of the inspection. This could mean months to years to start building big data typically required for machine learning.

Finally, as discussed in detail in Chapter 4 [133], the inspection process for most quality requirements is a human performing visual inspection. This visual inspection has long been studied within manufacturing and has demonstrated a person is capable of identifying 50% to 80% of defective machined products with 100% visual inspection [120], [142]–[146]. This misclassification of product has meaningful impact on the ability for machine learning algorithms to be successful. Dataspace overlap becomes a concern when trying to obtain accurate machine learning predictions [133].

Obtaining highly accurate results data is a challenge when implementing machine learning. Historically, the industry has not serialized castings, let alone tried to track the results of acceptable and defective parts throughout the entire supply chain. The industry also relies heavily on human inspectors, which are unreliable. Culturally, there is significant hill for the industry to climb to overcome the challenge of inspection for machine learning.

PROCESS

As reviewed in detail in Chapter 2, die casting is a highly complex system. This complexity presents multiple challenges when implementing machine learning including the number of sensors needed to collect the data, the environment where the sensors are implemented, human error and intervention throughout the process, dynamic system changing due to improvements and changes, and the ability to test and optimize while trying to produce production castings.

Collection of hundreds of thousands of data points requires hundreds of sensors implemented throughout the die casting system. This means adding or reading sensors from multiple equipment manufacturers, sensor brands, and age of equipment. This is a considerable controls engineering challenge when implementing for the first time. Additionally, the die casting environment is harsh. Liquid metal, die lube mist, hydraulic oil, and large temperature variation await most sensors placed within the die casting cell. This results in an additional level of maintenance for sensors, even after implementation is completed. It also results in failed sensors that may provide incorrect data. Additionally, not all sensors are static within the system. Sensors added for flow rates or temperature readings within the die add another level of complexity to the process. Not only do the setup personnel need to correctly install the die casting tooling correctly, but each sensor that moves into and out of the cell must now be properly identified and connected, so the data remains consistent setup to setup.

The die casting process is heavily influenced throughout the system by human intervention and error. The complexity of die casting produces thousands of potential failure points for humans. Some of these failures are catastrophic, while others may go unnoticed within the process. Fortunately, the industry has checks and balances within most foundries to minimize catastrophic failures that destroy tooling or equipment. Unfortunately, the unnoticed failures add noise and variation into the process data that make it difficult for machine learning algorithms to be successful. The idea with the die casting data framework is to identify areas where data can be collected to help ensure strict process control and identify potential human mistakes within the system, but this concept takes much effort. In addition to the human errors, die casting is ripe with a continuous improvement methodology. Casting defects and part engineering changes drive improvements to the tooling and process even while parts are in production. Changes to physical aspects of the process, like the die, could make previously collected data irrelevant to future machine learning applications.

The complexity of die casting also requires the conversation of initial conditions for each cycle. It has been shown that heat transfer and fluid flow are sensitive to these initial conditions, thereby making the thermal balance of the die casting die subject to these nonlinear equations. The industry uses warm-up cycles to get the die to a steady state temperature, so it is well understood these initial conditions change through time. Additionally, Miller has argued that die casting reaches only a quasi-steady state [44], meaning the temperature remains dynamic for most of the production run. The thermal temperature of the die surfaces dictates the location where porosity defects form in the casting. Tracking and collecting all this data is difficult. Advanced technologies like thermal imaging cameras taking images for every cycle can accomplish this task but come at large financial costs with significant time commitments to implement. Additionally, the type of machine learning algorithm becomes more complex as well. In these cases, a recurrent neural network that considers sequence and time-series data is required. This creates the need for additional data for training and testing of the algorithm before it can be implemented.

The final challenge to discuss about the process is the range of settings used on the die casting machine during production. Die casting is a capital intensive process. Millions of dollars are invested into one large tonnage die casting die. Given the constraints of program launches, the development of these tools in a production system may not be as detailed as needed to help with optimization. Additionally, if a process is found that appears to produce high-yield production, the motivation to search larger setting ranges becomes a challenge. If time is not allowed during a testing phase with the process development, the data generated and stored is limited from a domain standpoint. Instead of a larger dataspace of ranges, the data stored are limited to the process setting that was successful initially. Other ranges of input parameters go untested. To be clear, the ranges suggested differ from the approaches described in the literature reviewed back in Chapter 1. Some publications have created optimization models and tested input parameter ranges that would not be considered from an operations standpoint. Tsoukalas was one of these publications that uses a second stage plunger speed of 1.2, 2.5, or 3.8 m/s [21] which is an unrealistic range based on production experience. The concept proposed here is varying a speed determined to produce high yield castings, such as 2.5 m/s from 2.1 to 2.9 m/s to provide realistic production values for to the machine learning algorithm to pattern.

DEPLOYMENT

Even if 70% of the effort for a machine learning project is completed with the data preparation [27], it still means there is a significant portion of effort left in the deployment of machine learning. With this effort comes challenges within the deployment process. These challenges come from selecting the right algorithm, knowing when to update and modify the algorithm, and making IT decisions including hardware/software architectures and cyber security for the entire system.

Selecting the correct machine learning algorithm is a challenge for any machine learning implementation, regardless of the type of industry [5], [6]. The data collected dictates which types of machine learning algorithms can be considered for deployment [5]. If data is unlabeled, without any results, only unsupervised algorithms can be considered. If data has categorial values, a tree model may be considered. Selecting the correct type of algorithm can be a difficult task for those not extremely familiar with machine learning. Unfortunately, many manufacturers and die cast foundries fall into this category. Selecting the right algorithm may require some level of trial and error or use of a third-party vendor, thereby making it more difficult within manufacturing.

The complexity of the algorithm itself poses another challenge. Algorithms like neural nets are described as a black box, since the process used to make a prediction is not clearly understood or easily interpreted by humans. This is different than a decision tree, in which the algorithm clearly can be interpretated to understand the logic behind the classification [6], [28], [171], [195]–[198]. As described by Paleyes *et al*, often model selection ends up being simplified to assist with the understanding and acceptance of the algorithm, as well as to reduce the hardware requirements to perform the calculations [196].

The challenges do not end when the algorithm is selected, trained, or implemented. Maintenance during deployment also must be considered. Changes within the process and the data being collected will require modifications of the implemented solution. The question of when to update the algorithm must be considered by the team. This is not always an easy decision to make. Additionally, maintenance is not only on the trained algorithm. The hardware and software associated with the solution must also be considered. From network and computer hardware to software delivery systems, there are many interconnected parts that need to work together for a long period of time to implement machine learning within die casting. Given the harsh environment of die casting, quick and easy replacement of hardware must also be factored into the decision-making process.

As mentioned in the previous paragraph, a significant portion of deployment comes with the IT hardware and software used to deliver a machine learning project. Managing the IT process can be as challenging as managing the data collection through the OT (Operational Technology) process [199], [200]. Hardware and software architecture becomes another series of decisions when implementing machine learning in die casting. Location of hardware must be considered. Is the timing critical that an edge device is needed and integrated into the data collection system to run streaming machine learning on the data? If time is not that critical, then an architecture with data pulled from the cloud or an on-premise storage may be adequate. Each decision becomes an investment with technology implemented in production. A generic example of a hardware and software architecture for a machine learning solution implemented on a factory floor can be seen in Figure 57. The challenge is making the right IT decision. Implementing a solution and then seeing downtime, usability issues, or outright failure with the hardware and software would likely cast doubts on the entire machine learning approach.

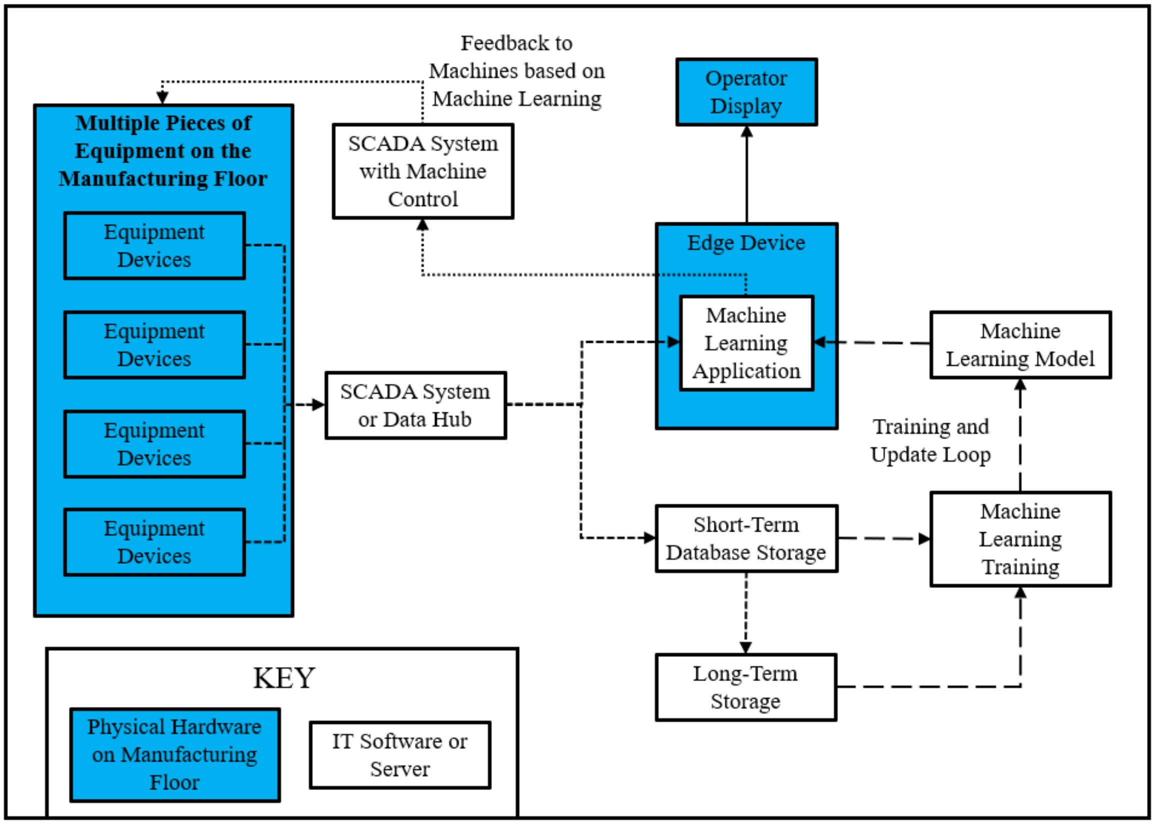


Figure 57: Generic IT architecture for machine learning implementation

The final challenge to review in deployment comes with the ever-increasing cyber security threat on connected equipment [31], [33], [199]–[203]. The increase in intrusions, ransomware, data interception, denial of service (DoS) attacks, and malware has made cyber security on the OT and IT devices a critical piece of the overall machine learning implementation [200]–[203]. This becomes another complex system layered in with the complexity already discussed involving the die cast process and data collection from Industry 4.0 efforts. Specialized skillsets in cyber security are needed to implement and maintain this technology. Finding and maintaining these resources becomes another challenge for the foundry.

NON-TECHNICAL

Technical challenges with data collection and the implications of machine learning in die casting are numerous. The non-technical portion of adaption and use of this technology is also noteworthy. Foundries face four key non-technical challenges when implementing machine learning. These challenges include a data-driven decision-making culture, interpretation and transparency of the models, experience where machine learning solutions developed in research have minimal or no real-world value, and the skills gap that exists between manufacturing and data science fields [6].

Using data to make decisions within manufacturing is a challenge that intuitively seems like it should not exist. Unfortunately, the manufacturing industry has a history of a no common data delivery platform, lack of data in some applications, uncertainty of what data to utilize, and trust issues with data [204], [205] . With these shortcomings, it should be no surprise that data-driven decision-making culture is not deeply embedded within manufacturing. This is compounded by lack of transparency and interpretation of machine learning models in research publications having minimal value in real-world applications.

As discussed in the Deployment section, the black box nature of some algorithms makes it hard for the humans involved to accept the results. Nunes and Jannach argued the lack of transparency and interpretation of complex machine learning algorithms prevent humans from trusting them [206]. This is echoed by Rudin and Wagstaff, stating a predictive model, regardless of how accurate it is, cannot be useful unless a human understands it [207]. Many of today's most popular machine learning algorithms, like neural nets or deep learning, suffer from this negative black box effect. It can be argued that lack of interpretability in some algorithms is additionally problematic for the die casting industry. In a real-world application, accurate predictions are not the end game. Machine learning is really being used to optimization of the process by predicting and eliminating quality issues and downtime events. The interpretability of the machine learning models is paramount for this success. Without the transparency, this becomes another hurdle for die casting to overcome with machine learning applications.

As initially reviewed in Chapter 1, and one of the motivations of this dissertation, literature and research that has been published to date on machine learning applications to die casting have minimal real-world value to the industry. The scope and approaches used were narrow and did not follow industry norms. This problem does not just apply only to die casting but also the lack of real-world machine learning value has also been experienced in the medical field [208] and computer networking [209].

Boutaba *et al* built on this by stating one of the problems with machine learning research is the use of synthetic data that does not reflect the complexity of real-world settings [209]. The lack of demonstrated value achieved by machine learning within the die casting industry only continues to challenge further cultural acceptance of it. As such, the focus of machine learning applications needs to shift to gain successes within the industry.

Finally, the last non-technical challenge with machine learning applications within die casting is the skillset gap that exists within the industry [85], [92]. Machine learning within die casting has seen limited publications. The die cast industry has also not been actively pursuing data scientists to implement this technology. This may be a fortuitous situation since there is likely a shortage of machine learning researchers available [207]. A skills gap of data science and machine learning knowledge exists within foundries. This is a large, non-trivial, non-technical challenge for the industry to overcome. Small steps are being taken within the industry. The 2021 North American Die Casting Association Congress and Exposition will have a technical session dedicated to machine learning [210]. Additionally the Advanced Casting Research Center (ACRC) and The Minerals, Metals & Materials Society (TMS) have been advancing training and conferences on topics of machine learning and AI withing the metals industry [7], [211], [212]. However, this challenge still exists, and much work remains to continue increasing the die cast industry's use of machine learning. The skills gap will remain wide for a while, which will be challenging for those trying to use the technology.

SUMMARY

Upon completing this section, the reader may feel a bit overwhelmed or, believe that the task of implementing machine learning within die casting is hopeless. That is not the intent or argument being made with this work. The die casting industry is dealing with a complex system. To be successful within a complex system, all the risks need to be identified and then mitigated. These challenges all present risks that must be thought through to be successful at machine learning applications in foundries. The research completed for this work points at a non-traditional machine learning approach to assist with this application. As will be discussed next, the die cast industry must focus on the correct applications of machine learning to add value.

ADDING VALUE WITH MACHINE LEARNING

The challenges of applying machine learning in a complex system are numerous. However, these challenges are not insurmountable. Value is created by understanding the correct applications of machine learning within die casting. To find the correct applications, it is worthwhile to consider predictions in complex systems in general. Even though Friedrich Hayek, in his 1974 Nobel Prize Lecture [213], was talking about economies, his idea can be applied to other complex systems:

> "Organized complexity here means that the character of the structures showing it depends not only on the properties of the individual elements of which they are composed, and the relative frequency with which they occur, but also on the manner in which the individual elements are connected with each other. In the explanation of the working of such structures we can for this reason not replace the information about the individual elements by statistical information, but require full information about each element if from our theory we are to derive specific predictions about individual events. Without such specific information about the individual elements we shall be confined to what on another occasion I have called mere pattern predictions – predictions of some of the general attributes of the structures that will form themselves, but not containing specific statements about the individual elements of which the structures will be made up" – Friedrich Hayek 1974

Hayek argues that without perfect information about each individual element of the system, it is impossible to make precise predictions. Instead, one is left to only predict general patterns, and it becomes an illusion to chase precision [213] :

> "Allow me to define more specifically the inherent limitations of our numerical knowledge which are so often overlooked. I want to do this to avoid giving the impression that I generally reject the mathematical method in economics. I regard it in fact as the great advantage of the mathematical technique that it allows us to describe, by means of algebraic equations, the general character of a pattern even where we are ignorant of the numerical values which will determine its particular manifestation. We could scarcely have achieved that comprehensive picture of the mutual interdependencies of the different events in a market without this algebraic technique. It has led to the illusion, however, that we can use this technique for the determination and prediction of the numerical values of those magnitudes; and this has led to a vain search for quantitative or numerical constants. This happened in spite of the fact that the modern founders of mathematical economics had no such illusions. It is true that their systems of equations describing the pattern of a market equilibrium are so framed that if we were able to fill in all the blanks of the abstract formulae, i.e. if we knew all the parameters of these equations, we could calculate the prices and quantities of all commodities and services sold. But, as Vilfredo Pareto, one of the founders of this theory, clearly stated, its purpose cannot be "to arrive at a numerical calculation of prices",

> because, as he said, it would be “absurd” to assume that we could ascertain all the data. Indeed, the chief point was already seen by those remarkable anticipators of modern economics, the Spanish schoolmen of the sixteenth century, who emphasized that what they called pretium mathematicum, the mathematical price, depended on so many particular circumstances that it could never be known to man but was known only to God.” - Friedrich Hayek 1974

The connection of Hayek’s statements to the complex die casting system has become clear through this work and experience within industry. The call for more data to improve the quality predictions of supervised machine learning is likely the absurd assumption that Hayek describes. To gain full knowledge of every element in the system in die casting, one would have to know the exact atomic placement of all the alloy elements as they are poured into the chamber and injected into the cavity. Knowledge about the miles of cooling lines and pipes would have to be captured, so any plugged line or hardwater build-up that changes the heat transfer within the die is captured. One would have to figure out how to capture the data associated with air evacuation of the die during the injection. Human influence on the process, including the setup of the thousands of individual tooling components and decisions on classification of scrap, would need to be fully detailed for these predictions. These are all highly complex and chaotic processes that influence part quality in die casting. Much of this knowledge is physically impossible to collect while still producing quality castings. Others are financially unjustifiable to gather. The goal with collecting more data is not to gain a perfect and precise quality prediction.

The concept provided in Chapter 4 for the Critical Error Threshold creates an economic threshold when a manufacturer should apply machine learning. Manufacturing systems must perform at a high yield rate to be profitable, so the datasets produced are highly unbalanced. The goal of manufacturers is to increase this unbalance by decreasing the scrap. To do this, machine learning must be applied differently. If the accuracy of the prediction does not reduce the cost, then it is not financially viable to implement the technology. This hurdle becomes harder to clear as quality improves and the critical error threshold reduces. As seen in Figure 58, the accuracy of the prediction model must get increasingly higher to create a generalized error rate that is below the Critical Error Threshold (CET). The high yield of manufacturing creates accuracy thresholds that demand near 100% prediction precision. As shown in the work throughout Chapters 2, 3, and 4, die casting is a highly complex system with randomness and human intervention. Perfect predictions in die casting are not likely. Therefore, the calls for more data to capture the single additional variable that will create this perfect prediction is likely a fool's errand.

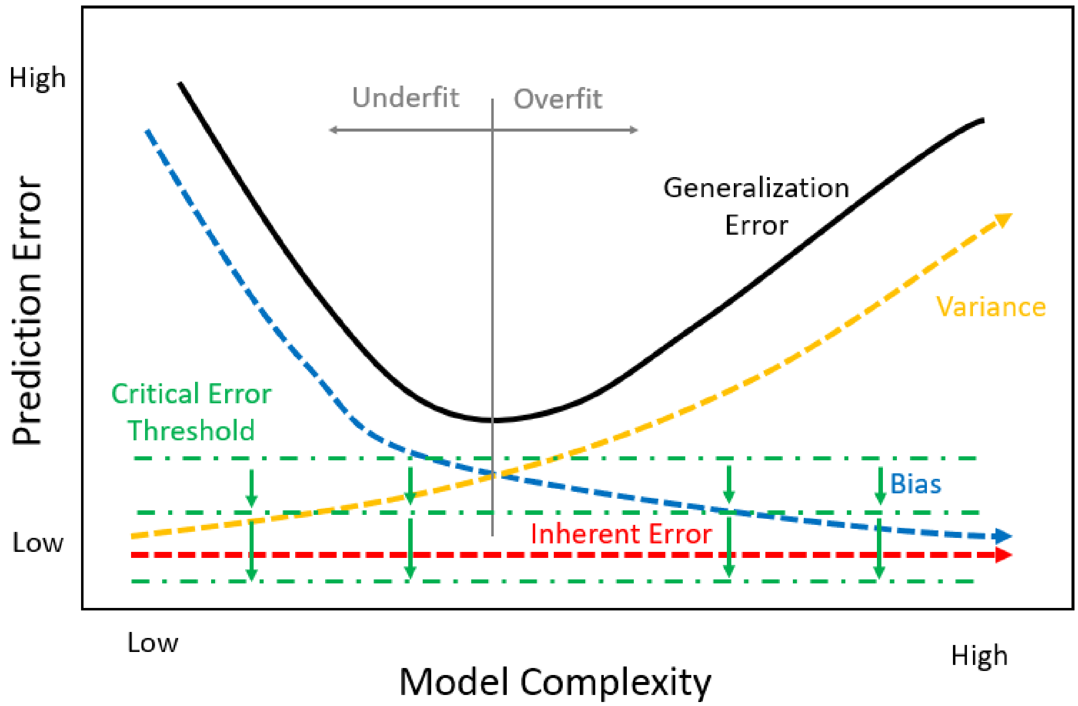


Figure 58: Reducing scrap rates drives a need for increased accuracy and a lower CET target

The goal with collecting more data from die casting should not be focused on applications of supervised machine learning to predict quality. Using machine learning to predict part quality will not likely create the value or motivation needed for mass adoption of this technology within die casting. Instead, more pragmatic applications of machine learning are needed to create real, tangible value for foundries. This requires machine learning approaches that are not new in theory but novel in their application within die casting system. The industry needs the data as outlined in the framework from Chapter 2 to provide the feedback to the die casting system to reduce variation and noise that exists today. Machine learning is needed for handling the volume of data to validate process control.

Substantial knowledge and financial value can be created within the die casting industry by using unsupervised machine learning. The volume and velocity of data generated within this system requires machine learning. Humans will never be able to analyze hundreds of thousands, if not millions, of data points every few minutes. This problem has always plagued the industry. Time-series injection data of velocity and pressure that has existed for decades. The industry simplifies this down to a handful of averages. To make matters worse, the lack of traceability and serialization makes collecting meaningful result data nearly impossible. Based on current practices, the die casting industry is not set up for supervised machine learning success. Accurate and meaningful quality predictions based on data currently available for most foundries is highly improbable, nearing the point of impossible. Again, if it was just one more sensor to add or variable to collect and control, the industry would have found and exploited it already.

The problem the die casting industry has is it's trying to control a highly complex process, with a handful of feedback variables. The rest of the system is assumed to be controlled and therefore ignored. This is the fallacy within die casting that needs to change. The industry must understand the entire system and be able to collect feedback data on the full process to comprehend the actual level of process control that exists. This work puts forth an effort to solve this problem by defining the data framework for feedback on the system and providing examples on how unsupervised machine learning can provide the analysis needed on that data. The first phase of die castings foundry and machine learning combination is utilizing unsupervised machine learning to review the feedback data, detect anomalies, and provide process control. With the complex system fully understood and in control, efforts at that point can be taken to predict quality with supervised machine learning. Without control, noise will dominate the data, and predictions in real-world settings will not provide value. The right application of machine learning is where the value is created for die casting.

Before the unsupervised machine learning case studies are explored, we return to Hayek's lecture one last time [213]:

> "If man is not to do more harm than good in his efforts to improve the social order, he will have to learn that in this, as in all other fields where essential complexity of an organized kind prevails, he cannot acquire the full knowledge which would make mastery of the events possible. He will therefore have to use what knowledge he can achieve, not to shape the results as the craftsman

> shapes his handiwork, but rather to cultivate a growth by providing the appropriate environment, in the manner in which the gardener does this for his plants"

The goal of this work is to cultivate the use of machine learning in die casting. Unsupervised machine learning is appropriate for this environment. By understanding the process and reducing the noise in the system, the die cast industry has the best chance to leverage this technology.

CASE STUDY OVERVIEW

The balance of this chapter is comprised of four case studies of machine learning applications applied at Mercury Marine aluminum die casting plant in Fond du Lac, Wisconsin. These case studies are publications from the past four years of this dissertation work through the North American Die Casting Association (NADCA). Each publication was included in this dissertation with the content as published through NADCA, with updates for figures and table numbers, references, and formatting to align with the rest of this dissertation.

The first case study was the initial publication in 2017 based on the earliest applications of machine learning at Mercury Marine to help predict part quality and optimize the process. This study really became the foundation for much of the rest of this dissertation. As will be seen, the applications of supervised learning did not yield useful results. Instead, unsupervised clustering was used to group input parameters. After the fact, results were applied to the clusters and found that certain clusters performed better than others. As a result, the process was optimized and scrap was reduced, although an accurate supervised quality prediction was never achieved. These findings provided much motivation in the study of the challenges of supervised machine learning in complex systems.

The second case study was published in 2018. This work focuses on the use of unsupervised machine learning applications for anomaly detection in thermal images. As described in Chapter 2, thermal images of die surfaces can be taken that contain millions of points of data. Being able to develop a process to review and detect changes within the images is important to help detect process changes within die casting. This case study reviews an approach to analyze and provide anomaly detection feedback on die temperatures that is novel and a first of its kind within the industry.

The third case study was completed in 2019. This work focused on time-series data related to the injection system, which has been the main focal point of process data within the die casting industry. The aim was to find other means to evaluate the entire time-series of data points versus descriptive statistics like average fast shot speed used in the industry. By developing a modified cosine similarity function, anomalies could be detected within the data. An Azure web-based app was developed and implemented to assist with the communication of these anomalies directly to foundry operations.

Finally, the last case study was presented in 2021. This study built on the use of time-series data with the use of autoencoders for anomaly detection. The goal in this case study was similar to the third case study: to demonstrate the evaluation of the entire time-series data stream is critical within die casting. The difference was the use of a neural-network based algorithm that has become popular in the machine learning field and the application on time-series data beyond the injection data. In this case study, the volume of die lube used during the spray cycle was evaluated to find anomalies between cycles. Since this technology is so new to the die casting industry, a major portion of the paper walks through an example using handwritten digits before applying it to the die casting data. This was done to educate the industry by providing a visual representation of what is happening with an autoencoder and how the anomaly detection process works.

The outcome of all four of these case studies has been the same: value was added to the die casting foundry without the use of supervised machine learning to predict quality. These works also show the wide breadth of process data used and available within die casting. Additionally, many of these algorithms have applications outside the examples given. Much of the volume of data from the framework in Chapter 2 is time-series based. The use of autoencoders can detect changes in the motor or pumps that can fix equipment downtime issues as well as it can detect flow changes in the spray cycle. The key lesson learned from these case studies is the right applications of machine learning within die casting can help the industry better understand the process and changes that happen within it. The feedback of the complex system can be monitored by machine learning and can inform humans when something has changed that needs attention. Given the complexity of the process and the volume of data it can generate, this is a huge step within the industry.

CASE STUDY I: INITIAL DEVELOPMENT OF MACHINE LEARNING ALGORITHMS TO PREDICT CASTING DEFECTS IN HIGH-PRESSURE DIE CASTING [3] [45]

Data giants like IBM, Google, Amazon, and Facebook have been using big data and machine learning algorithms for years and, in some cases, decades to help drive extraordinary results and insight for their companies and customers [214]. The high-pressure die cast industry lives in a data-rich world. A review of the die casting process reveals hundreds to thousands of variables that may affect the process or equipment, and therefore, the quality of a casting. Some of these variables are easily measured, while others are technically difficult.

This chapter will review an initial approach used at Mercury Marine's Casting Business Unit (a division of Brunswick Corporation) to experiment with big data from the high-pressure die casting process and then test its application in machine learning algorithms to improve the understanding of the casting process. Data sources for the analysis include thermal images of die steel after spray and shot end process parameters collected during the production cycle. The results of these initial algorithms will be reviewed to show the effectiveness of utilizing the data to help predict casting defects and what future areas of development, data collection, and improvement are needed.

INTRODUCTION

E.A. Herman stated in the preface of *Die Casting Process Control* that the "revolution of process control" had ended and that the degree of control that could be applied to die casting is orders of magnitude better than it was just a few years earlier [23]. Fast-forward to today and all manufacturing,

[3] This section is an edited version of a conference paper, included with permission from the North American Die Casting Association. D. Blondheim, Jr. "Initial Development of Machine Learning Algorithms to Predict Casting Defects in High-Pressure Die Casting," 2017 NADCA Congress and Tabletop, Atlanta, GA, Sep. 2017

not just die casting, faces a revolution. This is a cyber-physical system revolution, connecting manufacturing equipment to the digital world. There is no shortage of buzzwords that encircle this momentous change we face: Industry 4.0, Industrial Internet of Things (IIoT), Smart Factory, Digital Revolution, and the 4^{th} Industrial Revolution.

Buzzwords aside, the benefit of the technology now available is massive amounts of data for any manufacturing process and equipment. This is on a scale significantly larger than ever before, especially in die casting. We no longer live in the world described in *Die Casting Process Control* where "continuous measurement records are impractical" and "a reasonable number of measurements is the goal, and that is usually between 50 and 150" [23]. The technology exists today to collect tens of thousands of measurements across each of the thousands of parameters that may affect the die casting process and equipment. Die cast companies are no longer challenged to collect data, this technology exists. Instead, the issue is trying to process massive amounts of streaming data and, more importantly, learning how to improve the casting quality and equipment reliability based on this data.

The die casting industry has entered naturally entered the world of big data by pursuing short cycle times with a highly automated process. Die casting is also a complex process, with many subsystems that must work together to produce a high-quality casting. Data streams are available throughout the entire casting process as well as the subsequent inspection process. Examples of subsystems that can provide critical data include the:

- Furnace and metal delivery
- Die lube/spray
- Shot end of the die cast machine
- Clamp end of the die cast machine
- Thermal regulation units (cold water, jet cooling, hot water, and hot oil)
- Internal die sensors
- Thermal imaging of dies and castings
- Vacuum and venting process
- Input parameters used for machine control
- Cycle time data of every signal generated by the die casting machine
- Cycle time data of the extraction process
- Die and tooling design metrics
- X-ray or CT images
- Dimensional inspection data

- Ambient environment data (temperature, humidity, etc.)

Our ability to generate manufacturing data in die casting outpaces many other industries with significantly longer cycle times or less complex processes. A recent multi-million dollar European Union research project outlines more than 75 unique process parameters that were collected and analyzed to make improvements to high-pressure die casting [86]. These 75 parameters are just the tip of the iceberg, given the ability to incorporate thermal images and time series data in advanced analytics.

Traditionally, data is seen of as columns of unique parameters, with rows for individual parts. Often these unique parameters are calculated statistics that simplify a larger data set. For example, the average slow shot speed during a cycle is usually a singular number stored by a data collection system to summarize a time-series measurement. Many data collection systems store time-series information, like shot velocity and pressure throughout the shot cycle. A means to quickly and automatically analyze the entire string of time-series data has not been commercialized or embedded into the industry. Collection of hundreds to thousands of data points is realistic given the systems collect time-series data millisecond increments. Also consider the data difference between collecting die temperatures with a few thermocouples throughout the die versus thermal images of each die half offering 640 x 640 resolution of the entire surface (or almost 1 million pixels of data when the moving and stationary half images are combined). In short order, the die cast industry will be measuring an exponentially increasing amount of data versus, what should now be considered, a small number of traditional die cast parameters in a columned table.

This data is extremely important to the die casting industry. There is significant room for improvement with benchmark median values of equipment utilization at 68% and scrap rates at 8% [4]. Traditional statistical methods, such as statistical process control (SPC), have moved the industry to where it is today. Not clearly identifying or understanding the interactions of parameters in traditional SPC charts limits how effective they are for multivariate data such as die casting parameters. Big data combined with advanced methods of data analysis, anomaly detection, image recognition, and other machine learning algorithms will drive the next big step for improvement. Consider what could be defined as a world-class manufacturer producing castings with a 60-second cycle time and a 2% internal scrap rate. This process, which would require a high degree of control to establish and hit the world-class standard, is still producing scrap at a rate of 1.2 parts per hour. Optimizing and controlling the die casting process is a large challenge due to the complexity, variation, and interactions of multiple subsystems, along with the overall stochastic nature of metal casting. Helping to solve this problem is the motivation behind this experiment.

EXPERIMENT

One of the initial challenges of a large-scale data collection project is automating the means of collecting and storing data in an operation that runs 24 hours a day, 7 days a week. Commercialized systems and equipment manufacturers often have solutions on data collection for the shot-end process, which are readily accepted and widely used within the die cast industry [4]. Storing that data, on the other hand, is currently only done by 55% of the industry [4]. Experience has shown gathering data beyond the shot monitoring systems requires much internal development. Thermal images, subsystem cycle time data, thermal management systems, and even input parameters used to set up the die cast process are all typically independent systems. Pulling and then finding a means to store all that data in a database can be difficult. Although this chapter focuses on processing once the data is collected and not the collection process, the collection is recognized as a significant challenge a foundry will face when trying to collect non-traditional die cast process data.

The next challenge faced is part traceability. To solve the problem of "1.2 scrap piece an hour," data is needed down to the specific part for both castings deemed as good and scrap. Data published in 2014 stated that only a small number of the 55% of foundries that store shot monitoring data identify the castings produced with serial numbers for traceability [4]. Part identification and downstream data collection on casting quality is another significant hurdle a foundry must overcome to get meaningful data. Castings are often deemed scrap after machining and leak testing. Die cast companies that cast and machine in house gain a significant advantage in achieving quality data collection. In the case of this experiment, it has required years' worth of work to automate part traceability to the thermal image and shot data collection systems. Part traceability is another significant hurdle to complete the data analysis.

Knowing the challenges of data collection, as this experiment was launched, a casting and a die cast machine needed to be selected. The desire was to collect as much data as possible on a casting. A high-volume casting was needed, so the plant could run the product through multiple shifts, days, and weeks to experience all the normal process disruptions. This was not to be a controlled experiment, allowing any variation in parameters as established during the part approval. In the end, an external customer's casting was selected because of its volume, the plant's ability to collect the quality data with on-site personnel, and the home machine that had a thermal imaging systems installed. Not all the desired data-collection systems were in place at the time of the experiment, but the systems used established a baseline of what learning could be gained from a large data set and machine learning algorithms.

Casting, Process, and Equipment Details

The casting selected was a crankcase half made of A380 Aluminum. The finished casting weighs approximately 13 pounds and has been produced by the foundry for several years. The scrap rate from the foundry and machining due to casting defects, but not including warm-up scrap, typically ranges from 3% to 6%. The casting includes an isolated heavy section that is machined and unable to be fully cored due to part design. Porosity in this area is the leading contributor to scrap. An example of this porosity can be seen in Figure 59. The next largest contributor of scrap is gate break-in defects that are repairable after machining depending on size and location relative to the split-line edge as seen in Figure 60

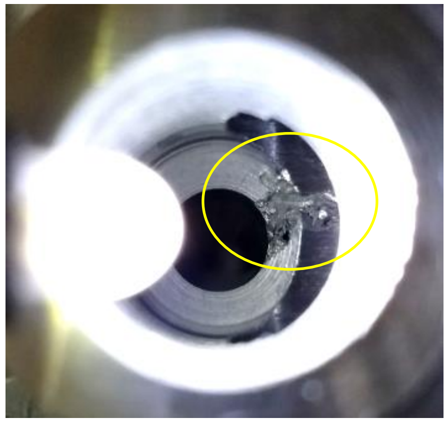


Figure 59: Porosity defect uncovered after machining

Figure 60: Gate break-in defect

The machine that ran the experiment is a 1200 ton IdraPrince manufactured in 2001. The machine has a 7.5-inch shot cylinder diameter and a 1:1 charged accumulator for intensification with a max 4500 PSI. The hydraulics of the shot end are controlled by an Olmstead Control Valve with a Textron Servo. The electronic controls are Prince Profiler and Prolink. The plunger tip size used to produce the part has a diameter of 4.25 inches.

Two thermal imaging cameras with an image collection system are installed on the die cast machine. This system was developed as a collaboration between Mercury Castings and a third-party integrator. The camera system takes thermal images of each die half after the spray cycle is completed, immediately before the die is closed for the next shot to be fired. Thermal image files are stored on the network and saved with the serial number information in the file name. Figure 61 shows a picture of the camera system mounted on the die cast machine.

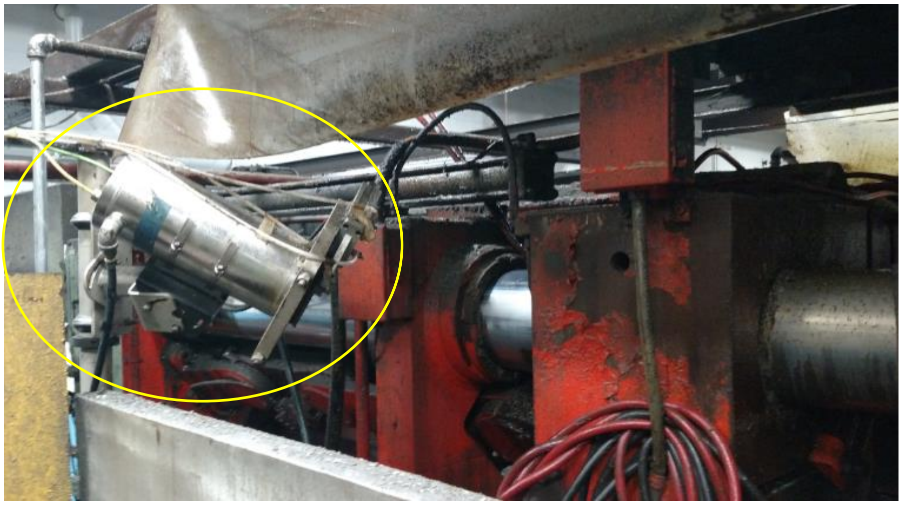


Figure 61: Thermal imaging system installed on die cast machine

The thermal image system saves two types of TIFF files. One file is a standard TIFF format that shows a grayscale thermal image that can be viewed with any generic image viewer. The other file saved is a TIFF image that stores the raw temperature measurement, in degrees Kelvin, as read by the camera. This format can be uploaded into several programming languages, including *R* [215], which was used for analysis in this experiment. The key difference is the actual raw temperature reading in this second TIFF file versus only a 0 to 1 pixel intensity value that is typically stored as an image file. This upload converts the image file into a matrix of numbers based on pixel location. Figure 62 shows an example of the thermal image and the data it produces for analysis. (Note that this is not the tool used in the experiment.)

	V1	V2	V3	V4	V5	V6	V7	V8
17	309.25	309.43	309.77	310.64	311.92	313.94	316.60	319.88
18	309.65	309.73	310.48	311.79	313.68	316.24	319.36	323.59
19	309.68	310.33	311.82	313.47	315.88	318.99	322.91	326.89
20	310.51	311.76	313.47	315.68	318.64	322.39	326.41	330.17
21	311.71	313.29	315.54	318.39	321.87	325.96	328.43	331.90
22	313.57	315.50	318.28	321.62	324.88	328.95	331.84	333.41
23	315.92	318.46	321.34	324.43	326.00	329.92	332.18	333.89
24	318.50	321.30	324.62	326.36	327.01	331.74	337.79	339.05

Figure 62: Thermal image example with data conversion

Finally, in the extraction cell, a pin-stamp unit was used to put on a 2D barcode and a human readable 11-digit serial number to complete the part traceability. Unfortunately, a programming error at the start of the experiment left the 2D barcode with no useful information, but the human readable code was in place. Data was manually collected instead of using bar-code scanners to track what serial numbers were shipped to the machining house. Figure 63 shows an example of the serial number used during the experiment.

Figure 63: 2D barcode and serial number example

The casting experiment took approximately one week to complete in the die cast cell. The experiment spanned two work weeks and all three shifts of regular production. A fortunate set of unpredicted events occurred: the equipment on the machine experienced two issues throughout the experimental run, an operator created a third issue, and a hydraulic gate-break issue created a fourth issue.

The first issue that reoccurred several times throughout the run was a sticking hydraulic valve in the shot system. This caused a much larger-than-normal variation in slow shot speed and fast shot speed. The second issue was an intensifier pressure leak, which caused a much larger range of intensified pressure throughout the run. The third issue near the start of the experiment was caught and corrected by the operator. The die cooling water was not turned on after the warmup cycles for the die, causing the die to get extremely hot. The final issue during the experiment was the hydraulic gate break operation caused a much higher level of gate break-in issues than normal. In all cases, castings were audited and passed X-ray inspection to continue with the experiment. Machining feedback on these castings was viewed as critical to help understand and learn the impact of parameter changes and its effect on quality. Figure 64 highlights the variation that resulted from one of these operational issues. In the figure, acceptable castings were labeled as Good = 0 and castings that were either scrapped or reworked due to a foundry issue as Bad = 1.

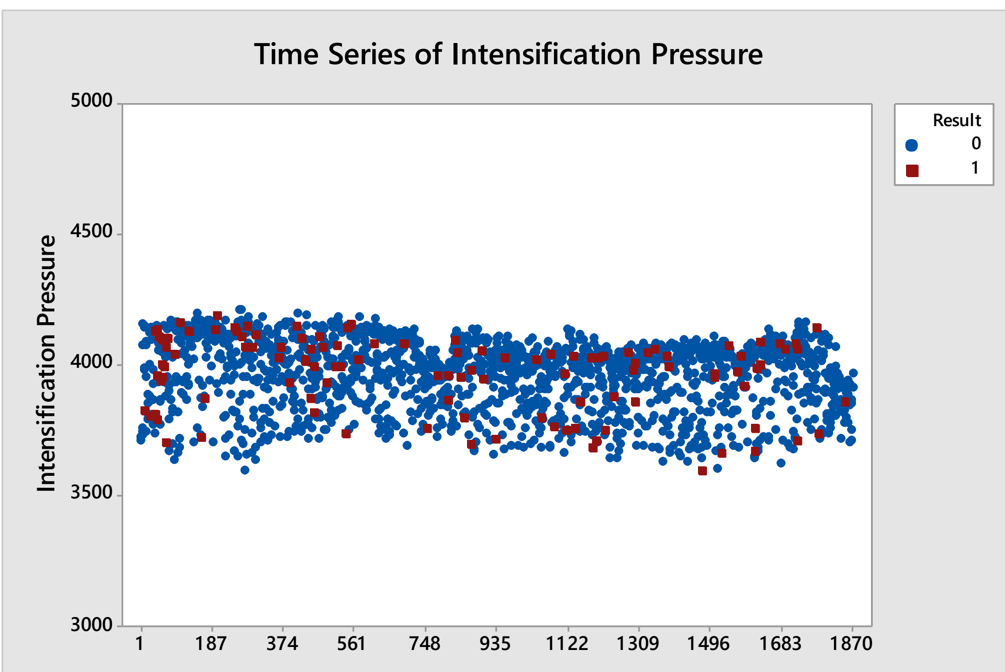


Figure 64: Intensification Pressure data over experimental run

High-Level Scrap Results

After casting, secondary trim operations were performed within the plant. Then castings were shipped to the external machine house, where they were machined and reviewed with a binary pass/fail criterion. Any casting that had a foundry defect, even those repairable, were deemed Bad in the study. Any casting that passed all porosity and leak requirements at machining were labeled Good. A total of 1873 serial numbers were recorded with results. Serial numbers belonging to castings scrapped as warm-up shots per our regular production process were not included in the analysis. Castings that were scrapped due to unrelated trim press or robotic trimming issues were also excluded from the study, since quality regarding the final machined product was not available. This scrap was insignificant compared to the total population of experimental castings. A breakdown of the scrap data can be seen in Table 30.

Table 30: High-level scrap results of experiment

Result	Quantity	Percentage
First Pass Good Castings	1761	94.02%
Gate Break - Scrap at Foundry	29	1.55%
Gate Break - Repair at Machining	55	2.94%
Isolated Heavy Section Porosity	28	1.49%
Total	**1873**	**100.00%**

Initial reactions to the yield were mixed given the known issues experienced during the experimental run. Gate break scrap and repairs were on the high side of normal, while the porosity in the isolated heavy section was a bit lower than normal. Data tables containing the critical shot information were updated with the binary machining results (Good = 0, Bad = 1), and that data was combined with the thermal image information to begin the process of applying different machine learning algorithms to the data.

MACHINE LEARNING BACKGROUND

Machine learning can be defined as the development of computer algorithms with statistical methods applied to large data sets to provide insight and help predict a future outcome [216], [217]. Google, Facebook, and Amazon use machine learning algorithms to improve the experience of their customers as well as their own bottom line [214]. The goal in machine learning is to utilize different statistical and mathematical algorithms to help detect complex patterns in data that are either too difficult for humans to understand or so large that it is too time consuming for people to complete the analysis. Algorithm are trained on big sets of data and gain insight from that data to make accurate predictions about future data.

Many textbooks have been dedicated to the study of individual machine learning algorithms. Additional books have been published regarding the programming and implementation of these algorithms in multiple software languages. Any summary on the topic of machine learning provided here should be considered an extremely high-level, basic introduction only. It is included to provide a baseline of information for the rest of the analysis.

One way to understand the process used by different machine learning algorithms is to look at a simple regression example. *Input values* along the X-axis can be used to help predict a *result* along the Y-axis. We can then use a *machine learning algorithm* to try to explain what is happening in the data and then use that explanation to make future *predictions* based on new input values. A *predictive model* is created based on the users selection of an algorithm, in this case regression, to be applied and any *tuning parameters* that need to be defined. In Figure 65 and Figure 66, the number of polynomials need to be determined for the regression. This is an example of a tuning parameter for an algorithm. Typically, data sets are split into two different, but representative, groups. The first set is called a *training set* and is used to train the machine learning model. The second set is called the *test set*, which is used to validate how well the model produced will make future predictions. Ideally a model will predict future results with good accuracy and is not subject to *over-fitting*. Over-fitting is seen when a model is highly complex and may explain all the data used to train the predictive model but does not really provide a good future prediction. See Figure 66 for an example of over-fitting.

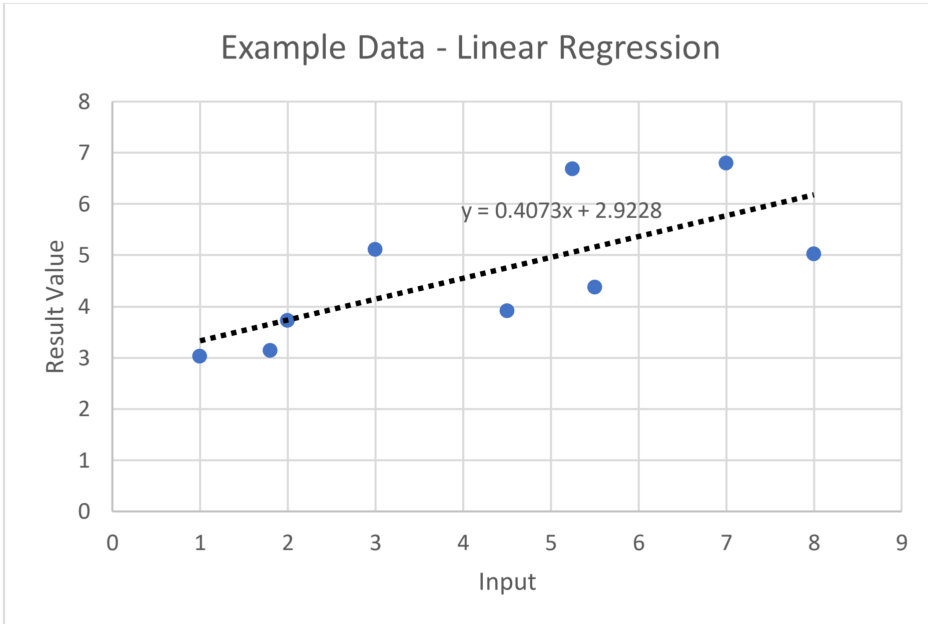


Figure 65: Example of a linear regression model

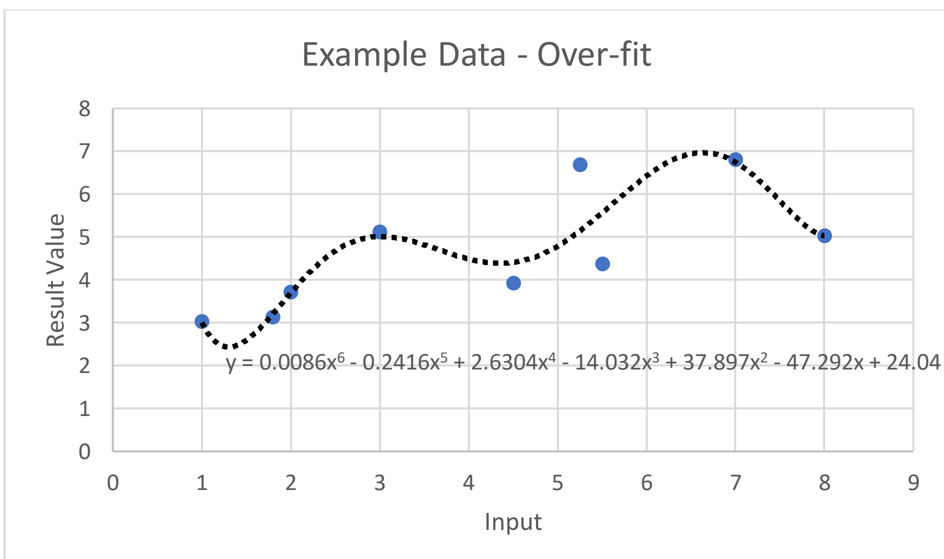


Figure 66: Example of an overfitted model

Most machine learning algorithms build on these concepts using a variety of different mathematical means and complexity. Although the approach may seem simple, the execution is not. In die casting we have multivariate inputs ranging from shot end measurements to thermal images. Sometimes this data is highly correlated, which can have an impact on certain algorithms. We also have some values measured in milliseconds and others that are measured in thousands of PSI. For many of the algorithms, the data needs to be pre-processed to be normalized so that the algorithms are not incorrectly affected by large values over small values. Normalization of the values typically follows Equation 21, but others are available.

Equation 21

$$\boldsymbol{Normalized\ Data = \frac{x - \mu}{\sigma}}$$
$$\boldsymbol{x = sample\ value}$$
$$\boldsymbol{\mu = sample\ mean}$$
$$\boldsymbol{\sigma = sample\ standard\ deviation}$$

Also with machine learning, there is an important discussion of feature selection: how do we select which features to keep in the algorithm that provide insight and which ones we can remove. Means such as Principle Component Analysis (PCA) reduce the data features to find linear combinations that capture as much possible variance, but then lose meaning in the parameters and implementation and optimization become harder to comprehend. PCA was not used for this experiment, but could be a useful tool with highly correlated features.

Finally, the algorithms themselves can vary greatly. One big distinction is the definition of supervised versus unsupervised learning algorithms. In supervised learning, the results are known on the data set, and the algorithm is trained to that result. In unsupervised learning, the result is unknown, and the algorithm is grouping items together based on their input variables. Table 31summarizes some of the algorithms by describes groupings of algorithms, names of individual algorithms, and the type of learning used. This table is only a sample of the algorithms available.

Table 31: Summary table of machine learning algorithms

Algorithm Group	Learning Type	Examples
Clustering	Clustering	k-Means Clustering, EM Clustering, KNN
Trees	Classification and numeric	Decision Trees, Regression Trees, Random Forest, Bagged Trees, Boosting Trees
Regression	Numeric	Linear Regression
	Classification	Logistic Regression
Support Vector Machines (SVM)	Classification and numeric	Classification SVM, Regression SVM, Class One SVM (unsupervised)
Neural Networks (NN)	Classification and numeric	Forward Feed NN, Back Propagation NN, Recurrent NN, Bayesian NN

Each of these models has its own set of strengths and weaknesses. There is no right model when trying to complete predictive analytics. Often multiple models need to be set up and evaluated to see which one works best for the data at hand. This becomes an iterative process, which was learned first-hand during this experiment.

ALGORITHM IMPLEMENTATION

With the 1873 results available, the data from the shot monitoring system (24 process parameters) was combined with the furnace temperature and the thermal images of both die halves. Also combined into this data set was a calculated count from the last machine stop. This data would determine if the shots taken soon after a start-up had a higher potential for being scrap. All data and thermal images were pulled from SQL tables or network locations and combined into one data-frame in *R* using custom programming scripts. Thermal images were also cropped and scaled as needed in *R* to reduce the number of pixels and therefore input parameters. The data was normalized with scaling and centering using the *caret* package in *R* [218].

The goal was to utilize machine learning algorithms to properly classify acceptable castings compared to any type of scrap or rework caused by casting defects. The hypothesis was that a combination of two or more parameters was trending to one section of their natural variation all at the same time. This interaction between the multiple inputs would cause the defect, which would occur infrequently throughout the course of the production run. Therefore, if the critical input variables could be understood, a process could be targeted to help stay in the zone that produced only good castings.

Complex Algorithm Results

Given the complex nature of the die casting process, testing initially started on the more complex machine learning algorithms. The data was randomly split into two sets, a training set and a test set. Different splits were tested ranging from 70%/30% to 95%/5%.

Neural nets (NN), which are designed to mimic the brain neuron activity, were first tested. Neural nets have an input layer, multiple layers of hidden neurons or nodes, and then an output layer. Weights are summed and calculated through an activation function. Weights are then calculated and modified based on the entire training set of data. Figure 67 provides a visual of neural net structure.

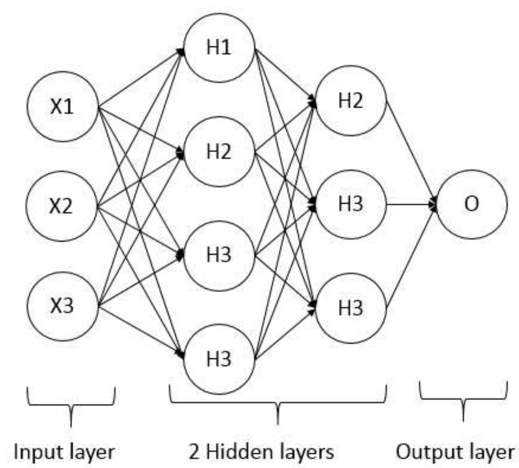


Figure 67: Example of a neural network structure

R packages of *nnet* [219] and *neuralnet* [220] were used to test forward feed and back propagation NNs. The networks of multiple sized hidden layers and neurons per layer were tested with different activation functions. The results of these trained models were the first indication of a problem with the experimental data sets and machine learning. The predictive model created by any training set always predicted 100% Good product. Although this model was correct to a very high percentage (given the yield rate of the castings), it produced a false positive result for all Bad castings in the testing sets.

With this data set and the initial testing completed, the neural net models could not accurately classify the castings produced.

The next algorithm tested was a support vector machine (SVM) with the *e1071* package in R [221]. An SVM uses vectors to create divisions between different classes of items. These vectors can be non-linear to help solve highly complex data-sets. The vector used maximizes the distance between data classes and can be tuned with a cost function that can allow the user to program the amount of cost associated with an incorrect classification. An example of a simple SVM is seen in Figure 68.

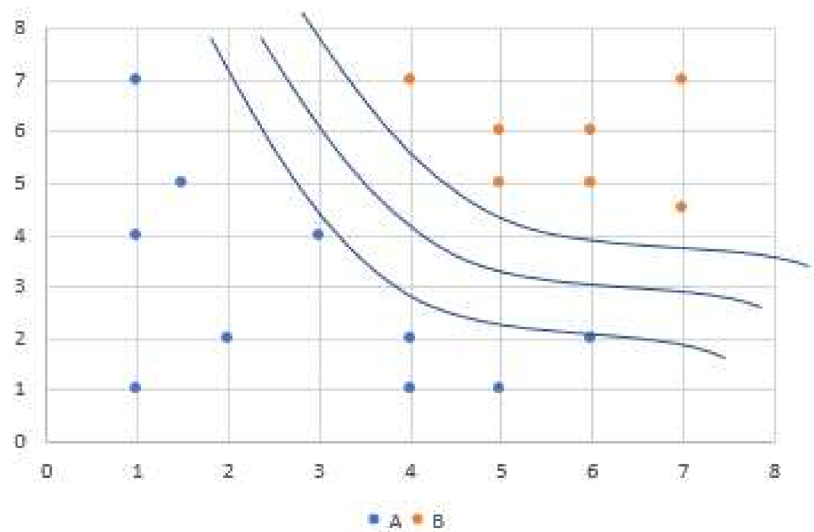


Figure 68: Example of a SVM classification in two dimensions

The result of the SVM was similar to the neural net. Everything was predicted as Good castings thereby poorly classifying Good and Bad results in the test data set. Unfortunately, comparable results were seen when applying Random Forest and CART tree models as well as regular Logistic Regression. None of the complex models could correctly classify the Bad castings from the Good castings with the data that was collected with the experiment. After this point, the approach in analyzing this data was changed.

Clustering Algorithm

Clustering analysis is typically done on data whose results are typically unknown, which is also called unsupervised learning. The purpose of clustering algorithms is to group related items based on all the input parameters. The k-means clustering algorithm works by dividing a sample of size *n* into *k* different clusters, where each sample is measured to the nearest mean value that serves as an average or representative sample of that cluster. The number of clusters the sample is to be divided into is a user-defined value or tuning parameter. Observations are assigned to each cluster to reduce the within cluster sum of squares or variance. Mathematically this is defined in Equation 22.

Equation 22

$$\begin{aligned}
&\boldsymbol{Given}: \\
&\boldsymbol{observations} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3} \ldots \boldsymbol{x_n}) \\
&\boldsymbol{number\ of\ clusters} = \boldsymbol{k} \\
&\boldsymbol{Cluster\ sets} = (\boldsymbol{S_1}, \boldsymbol{S_2}, \boldsymbol{S_3} \ldots . \boldsymbol{S_k}) \\
&\boldsymbol{Objective}: \\
&\mathbf{arg\,min} \sum_{i=1}^{k} \sum_{x \in S_i} \|\boldsymbol{x} - \boldsymbol{\mu_i}\|^2 = \mathbf{arg\,min} \sum_{i=1}^{k} |\boldsymbol{S_i}|\ \boldsymbol{Var\ S_i} \\
&\boldsymbol{Where}: \\
&\boldsymbol{\mu_i\ is\ the\ mean\ of\ S_i}
\end{aligned}$$

The k-means algorithm needs an initial set of cluster mean values. There are multiple ways to provide this. One option, which was used in this experiment, is to randomly select *k* number of points in the data set and use them as the initial cluster means. An iterative process of two steps then takes place:

- **Assignment Step**: Assign a sample to the cluster with the nearest Euclidean distance to the cluster mean.
- **Update Mean Step**: With samples assigned, calculate a new mean value for that cluster.
- **Repeat** steps until there are no new assignments of points to new clusters; at this point the algorithm has converged.

A k-means algorithm will converge to a local minimum, but there is no guarantee that the solution found will be the global minimum because of the randomly assigned initial starting values. To help overcome this issue, k-means clustering algorithm is typically ran multiple times with new random initial weights at each start to help find the global minimum [222]. Because Euclidian distance is used for assignment, data needs to be normalized before applying the algorithm.

For this experiment, the Elbow Method was used to help determine the number of clusters that would be used in the analysis. The Elbow Method involves plotting the ratio of the within cluster variation compared to the total variation of all samples. As more clusters are added, additional variation is explained within the clusters and therefore approaches the total variation of the sample. With the Elbow Method, the ratio of within variation to total variation is plotted among clusters sizes to see approximately where increasing the number of clusters provides a diminishing return on additional variation explained, the “elbow” of the graph [216].

Figure 69 shows the Elbow plot created with all the experimental data. The *R* package *stats* [215] contains the *kmean* function used to complete this analysis. The parameters were pre-processed to normalize them. Clustering was based on the data set containing Good castings and porosity in the isolated heavy section. Castings with gate break-in issues were removed due to known issues with the casting extraction process. Analysis was ran with a maximum number of step iterations to be 4000, and each cluster size used 400 iterations of randomly selected starting points to find the best local minimum. The variance ratio was calculated through a value of k = 30.

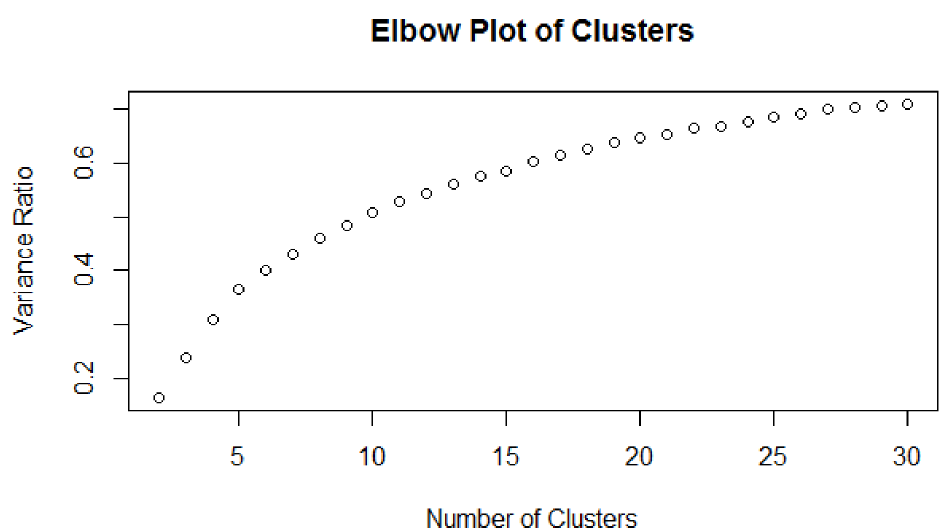


Figure 69: Elbow Method for determining number of clusters

Because there was no strong elbow point in the plot, individual analysis of clusters was performed ranging from 5 to 12 clusters. The analysis performed on the clusters combined the unsupervised clustering algorithm with the known results of the experiments. A table was created for each cluster listing the number and percentages of Good and Bad castings. The intent was to understand if any clusters performed significantly better than others and then determine what those clusters had in common. Also, it was the goal to keep the number of samples within each cluster statistically significant, with ideally several hundred samples falling within each cluster. Based on these requirements, a final number of 7 clusters was selected.

This cluster arrangement yielded several interesting results. Cluster 4 had only 4 samples in it. These were all castings with cycle times significantly longer than normal by a factor of 3 or 4 times the mean cycle time, driving a unique cluster just for that parameter. Another interesting result was Cluster 2 had 329 samples with zero Bad castings. Cluster 5 and 6 show the highest level of scrap percentages. Cluster 1 and 3, each with a large number of samples, showed good performance: more than 99.2% yield. Table 32 provides the details on the count and percentages of the cluster assignment. A graph of the cluster assignments can be seen in Figure 70, showing when the clusters were assigned throughout the time-ordered sample run.

Table 32: Cluster Assignment Counts and Percentages

Cluster	**Count**		**Percentage**	
	Good	**Bad**	**Good**	**Bad**
1	255	2	99.222%	0.778%
2	329	0	100.000%	0.000%
3	416	3	99.284%	0.716%
4	4	0	100.000%	0.000%
5	341	12	96.601%	3.399%
6	362	10	97.312%	2.688%
7	54	1	98.182%	1.818%

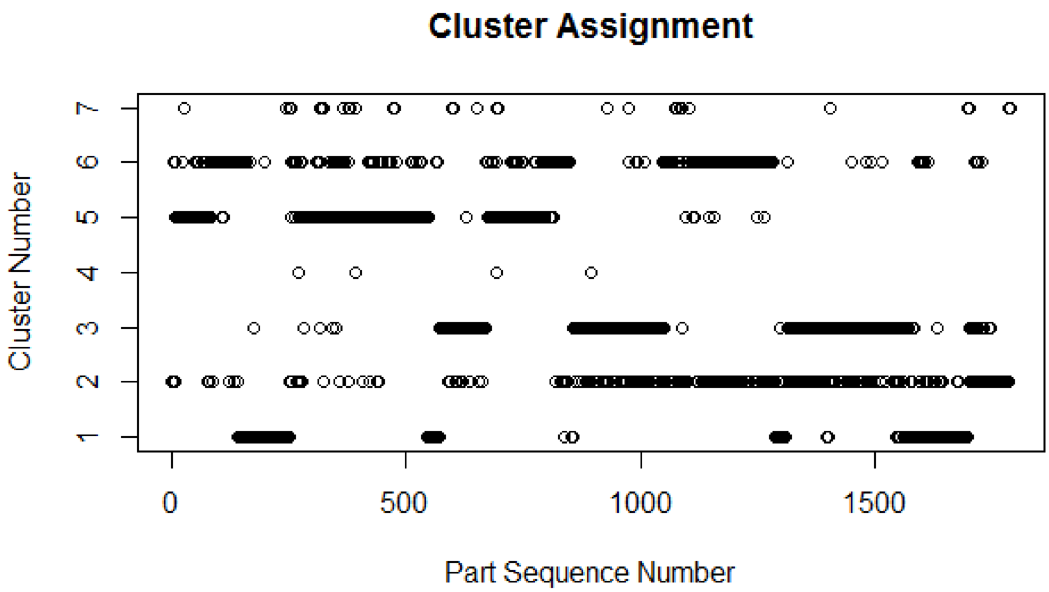


Figure 70: Cluster Assignment with 7 Clusters

An interesting realization found in Figure 70 is that during any 150-piece consecutive run, castings from at least 4 of the 7 clusters will show up. 150 castings represent approximately 4 hours of production. The clusters identified by the algorithm are independent to parameter drift or process problems that were experienced during this sampling. Using traditional SPC methods on individual process parameters is inefficient at identifying when multiple parameters may be moving together or interacting with each other. Utilizing the clustering algorithm quickly provided process insight that did not previously exist.

Cluster Visualization

With a final cluster model complete, visualization on the input parameters was completed using box-plots, individual value plots, and scatter diagrams to identify differences in the parameters per clustered group. From the box-plots, differences between cluster 2, 3, and 1, the best quality clusters, can be seen when compared to 7, 5, and 6, the worst quality clusters (numbers are ordered with increasing scrap percentages). Values for the Average Slow Shot Velocity and Intensification Pressure can be seen in Figure 71.

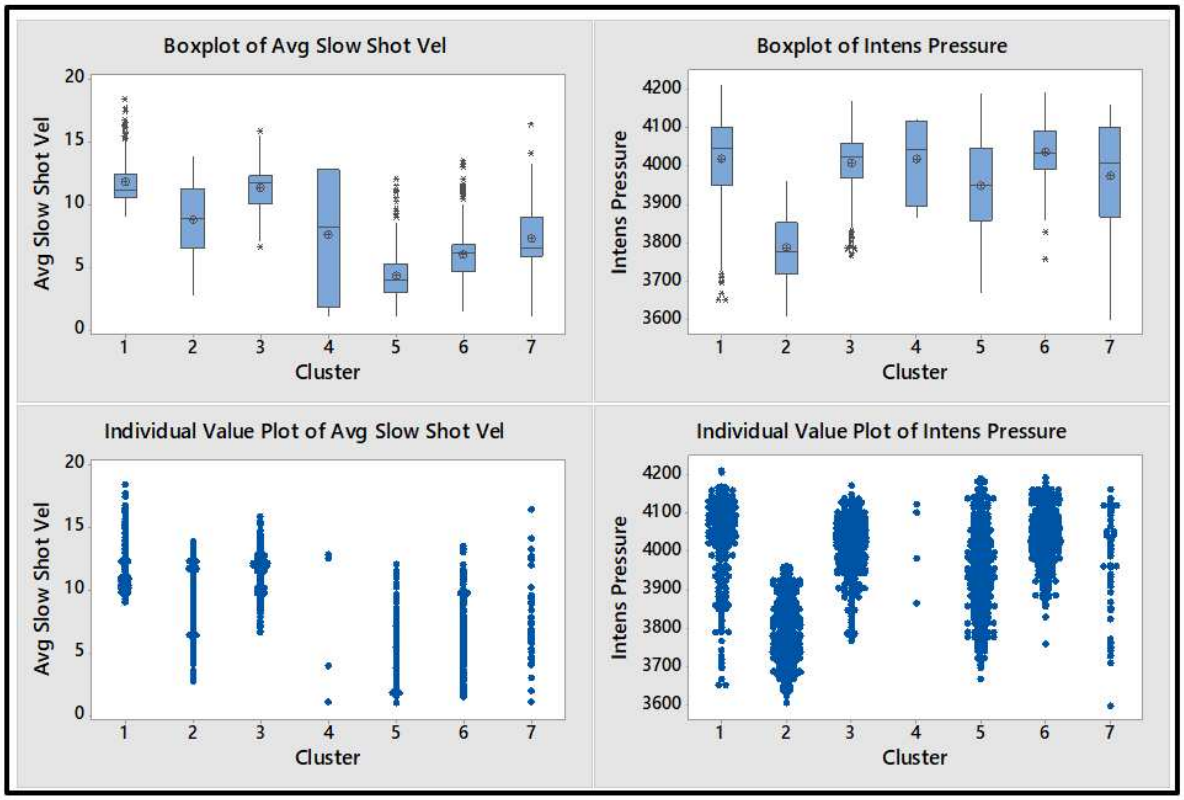


Figure 71: Boxplot and Individual Plots of Average Slow Shot Velocity and Intensification Pressure

Figure 71 shows Cluster 2 (the best quality cluster) has a significantly lower mean intensification value when compared to the other clusters. However, intensification pressure itself is not the key between good and bad, since Cluster 1 and 3 with around a 99.2% yield are approximately equivalent to Cluster 5, 6, and 7 with the worst yields. Reviewing the box and individual value charts of the input parameters, it is clear there are interactions between the parameters that result in better quality castings. A scatter diagram matrix was produced with the key input parameters. Two diagrams were created. Figure 72 shows the different clusters each identified in the matrix. Figure 73 shows the same matrix but highlighting Good = 0 versus Bad = 1.

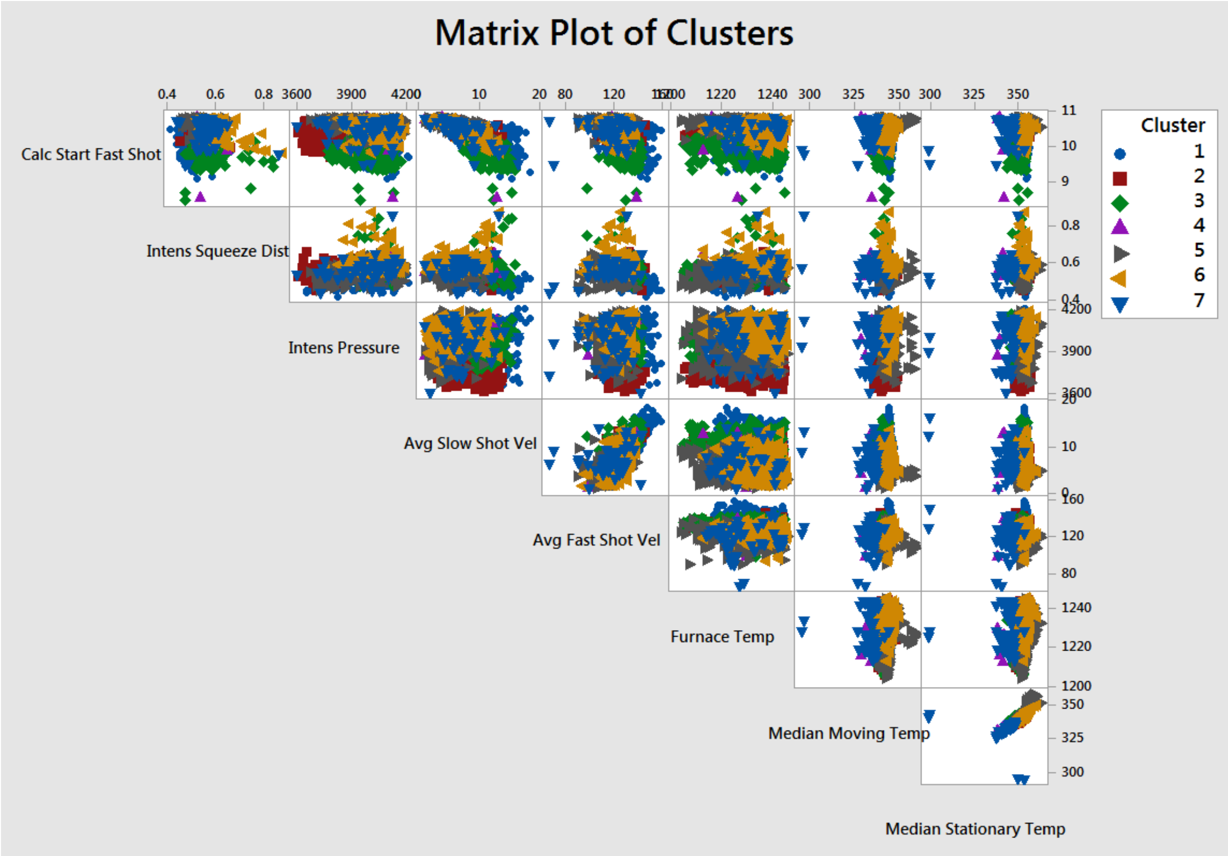


Figure 72: Matrix scatter plot of key parameters by clusters

With these matrix scatter plots, a clear area of interest became evident. An L-shape of Bad castings show up in the scatter plot of Intensification Pressure versus Average Slow Shot Speed. Inside this L-shape is a region that yielded 100% Good castings, absent of porosity in the isolated heavy section, as detected by machining results. This area is comprised mostly of Cluster 2 (the best cluster) as well as Cluster 1 and 3. Figure 74 shows the scatter plots discussed. The other parameter scatter plots did not show any other significant areas of interest. Average Slow Shot Velocity and Intensification Pressure became the two key parameters to validate in this experiment.

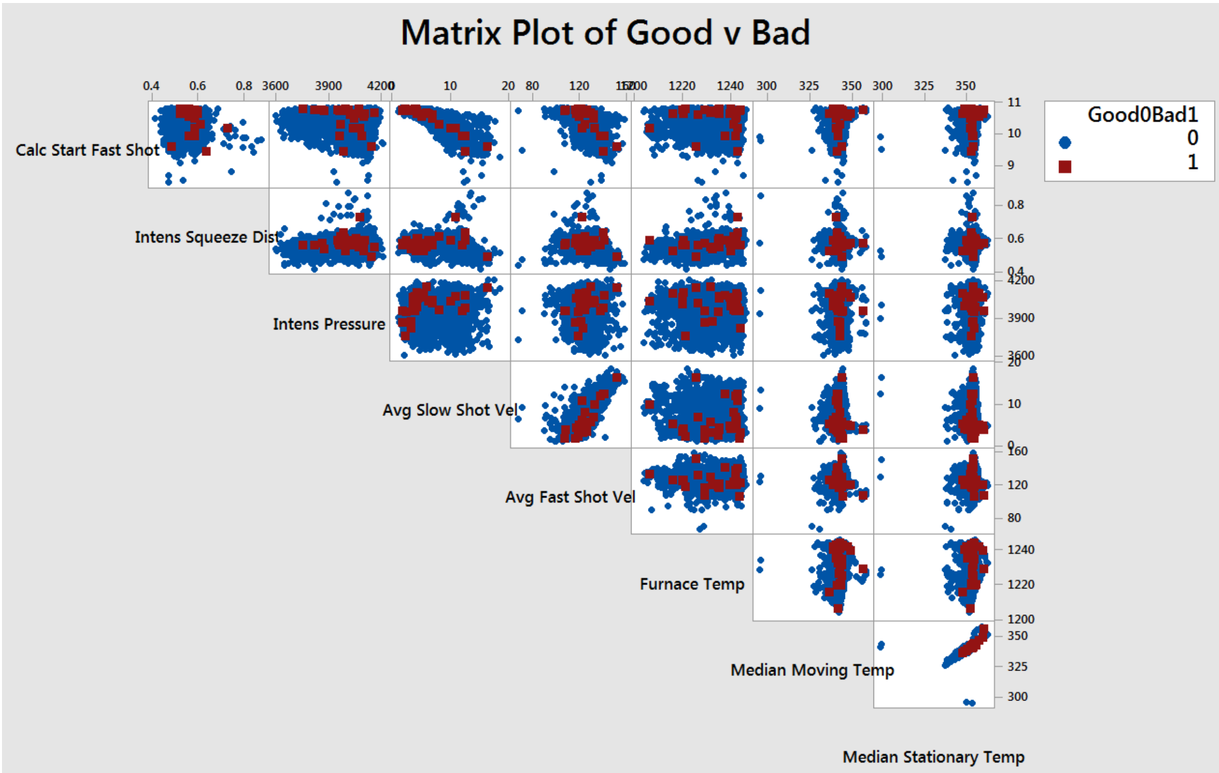


Figure 73: Matrix scatter plot of key parameters by result (Good v. Bad)

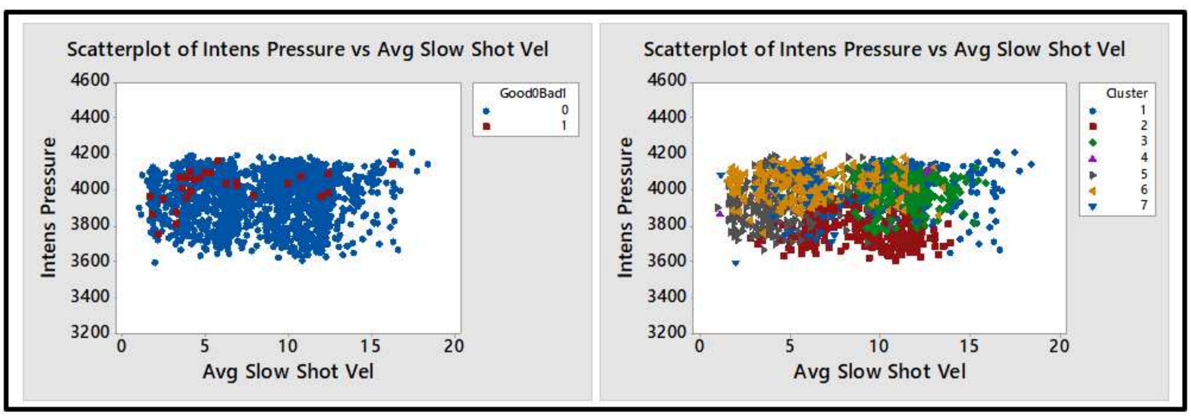


Figure 74: Scatter plots of intensified pressure versus average slow shot velocity

<u>RESULT VALIDATION</u>

Two different means were used to help validate or provide directional guidance on the results from the cluster analysis. Initially a computer simulation was performed to test the relationship between the Average Slow Shot Velocity and Intensification Pressure. Then a historical review of the data from those castings found to have the same defect was compared to process parameters of the different clusters.

Casting Simulation

A casting simulation design of experiments was performed varying the Average Slow Shot Velocity and the Intensification Pressure between simulations. Based on mean values of the clusters, the Average Slow Shot Velocity ranged from 5 inches per second up to 11 inches per second. The Intensification Pressure ranged from 3800 PSI hydraulic pressure to 4100 PSI.

The results of the simulations showed an improvement in the Fraction Solid results as well as the Hot Spot result with the increased slow shot velocity. There were no quantifiable differences when looking at the increased intensification pressure in the given area between experiments. The positive influence on the faster Average Slow Shot Velocity follows standard practice given the higher range brings the speed up closer to the Critical Slow Shot Velocity [70]. Overall, it was found to be difficult understanding the results of the simulations showing 97% yield rates versus 100% yield rates, since changes in the results were minor. The fidelity of the simulation may be lacking at this detail without significant knowledge from actual castings. Additional steps were taken to look at actual casting data since the results of the simulations were not conclusive.

Historical Review

A review of historical data for a previous six-month period was completed to see if additional insights could be gained regarding the process. The current practice in place is to record all serial numbers that are scrapped at the external machine house. This information is stored in a database with the defect reason. The current gap with this process is only scrap castings are recorded. Assumptions could be made about other serial numbers being machined and passing as Good based on a duration from the cast date. In this case, this assumption could be very inaccurate, since the inventory levels and work-in-process changes significantly with this product.

The database was filtered to find the historical data for just the isolated heavy section porosity to create a new data set. A total of 432 scrap castings were identified. With these serial numbers, the shot data for intensification and average slow shot speed was pulled and combined with the cluster analysis data. The box plots and individual value graphs were created with the 7 clusters and an 8th group added for historical scrap (HIST). These graphs can be seen in Figure 75.

The graphs show similar speeds and pressure used in the HIST scrap group when compared to the worst yielding clusters 5, 6, and 7. Graphically, the historical results confirm the findings in the clustering algorithm, in that both average slow shot velocity and intensification pressure impact the quality of this isolated heavy section. What appears to be academically interesting is that it requires a lower intensification pressure to produce the best castings, provided the slow shot speed is within an acceptable range.

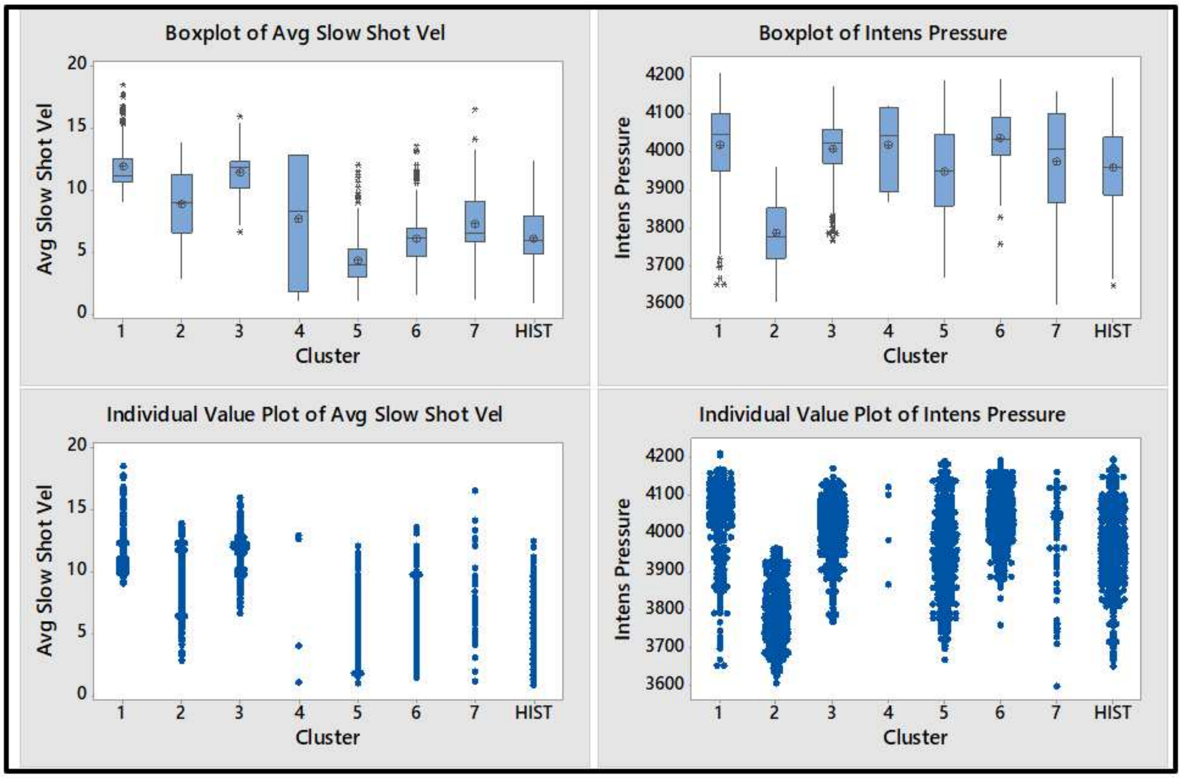


Figure 75: Boxplot and Individual Plots of Average Slow Shot Velocity and Intensification Pressure comparing clusters with historical scrap data

FUTURE STEPS

Machine learning algorithms were applied to big data seeking scrap rate improvements. Results have provided many learning opportunities for Mercury Castings and additional work and research areas have been spawned by this project. Many of the lessons can be applied for those interested in leveraging big data for die casting. This section will review the eight key points from this experiment.

Data Collection and Processing

The data collection and data processing portion of this type of project is not to be underestimated. This effort is worth mentioning again since the data is what drives any of the machine learning algorithms. Manual processes for data collection are prone to human error. Pin stamped 5 can often look like a 6, and a 6 can look like an 8 when a human is recording hundreds of these serial numbers on paper each day. These errors caused hours of lost time reviewing, verifying, and cleaning the data prior to using it, which itself is subject to error. The intention was to automate the collection using a 2D pin-stamped bar code, but a programming error at the start of the study left the barcode unusable. In cell verification of 2D barcodes would have immediately found this error.

One result of this experiment was the installation of 2D barcode readers in critical cells where barcoding of castings has been deemed important. When scanned, 2D barcodes provide error and contrast measurements letting the user know how easy the barcode was to scan. This additional data is now collected and stored with each casting to help identify thresholds of readability allowed and help predict, using machine learning, potential failures or maintenance needs for the pin stamp unit.

Sample Verification

It was originally intended to complete a second experiment with targeting process settings based on the results of the clustering algorithm. When the initial experiment was completed, the tool used to produce the casting was near the end of the life. Once results were compiled, the tool used in the experiment had been retired and replaced with a new model casting. The gating of the new tooling was different than the tool used in the experiment. Core changes in the area of concern also happened with the new tool design. A direct comparison between the two tools may not be accurate given these changes.

With the new tool being used, process settings were set given the lessons learned from the clustering algorithms. The process has a targeted average slow shot velocity similar to the experiment and reduced intensification pressure to the targeted 3800 PSI from the best cluster. Initial results from X-ray and machining feedback were promising. After 2.5 months with the new tool and process changes, a total of 42 castings have been scrapped due to porosity in the isolated heavy section. A 76% scrap reduction for this defect is seen in comparison to the 6-month historical scrap of 432 castings. The improvement may be contributed to a combination of new tool design and process settings. The exact breakdown is not known, but directionally the results are promising. In future experiments planned, a direct resample will be built into the analysis to provide more accurate verification.

Binary Results

In hindsight, an issue with this experiment was the binary classification of the results. Good versus Bad may explain why the complex machine learning algorithms first applied failed to yield any classification results. The true results, especially with the porosity in the isolated heavy section, are only partially known with this binary approach. We learned which castings did not exhibit the porosity after they were machined versus the ones that did. However, as the experiment was set up, it failed to show if the porosity was eliminated with the different process settings or simply just shifted out of the area the casting is machined.

Additional data needs to be collected to eliminate this binary result. Initial thoughts include X-ray images or CT models to have a continuous variable for the measurement of quality. This could provide the needed information to help more advanced algorithms perform well on the input data. Developing ways to quickly collect and assess these items are difficult, but manageable. X-ray images could be processed in software to provide quantitative measurements of porosity size and location. Regardless of the solution, it leads to more data collection and processing requirements to apply the machine learning algorithms. Finding ways to automate this will be critical for success.

Probabilistic Outcomes

Building off the binary results, another key lesson learned was to understand ranges in process parameters should be seen as changes in the likelihood of producing 100% acceptable castings. Within the multiple dimensions of input parameters, there are interactions happening that impact casting quality. Even if process limits are set using statistical process control (SPC), parameter values that produce 100% good product at target values may result in lower yields as they approach the limits. This reduction may be limited to 1% to 2% scrap increases, or it could be a cliff, increasing scrap rates more than 20%. It is not economically feasible for any die caster to test all these limits when setting up the process for a new tool.

Utilizing clustering algorithms provides a potential insight into the likelihood of scrap based on past results. More on this topic is in the next section on K-Nearest Neighbor. The future way of thinking about process parameter changes may be understanding that the process parameters is producing at X% yield, and by moving the parameters to this new target, the operation will now yield Y%. This does assume that nothing else in the process changes. Water lines that leak or cooling lines not turned on may result in step changes that may not be predicted with any learning algorithms, unless this could be detected through thermal image anomaly processing.

K-Nearest Neighbor

Tied to the previous section, applying a machine learning algorithm called K-Nearest Neighbor (K-NN) may be helpful to understand the probabilistic nature of yield rates based on process settings. K-NN is similar to the k-means clustering analysis that was completed here. In this algorithm, the training data set is used as a benchmark with known results. A new sample is then put into the algorithm and applies a distance measurement to its *k* number of points that are dimensionally closest to the new sample. The *k* value is selected by the user creating the algorithm. Based on the outcome of these *k* closest results, a prediction is made about the results of the sample [195].

This approach could be used to produce an item that may be useful for die casters. If one chooses a large *k,* say close to 100, what is being calculated is a yield rate within that sample point using a statistically significant number of nearly identical sample parts. With a large value, we could look at the probability of yield percentage based on the sample around it. If K-NN is applied on the factory floor, every shot could have a yield prediction associated with it. Another way of thinking of this is that it could create a profile or gradient of scrap rates throughout the parameter limits. Parameters could be selected to reduce the likelihood of scrap or, at a minimum, a sound business decision could be made as to why the change is necessary.

Thermal Image and Shot Profile Learning

During this experiment, statistics of the shot profile and thermal image were used. One of the critical lessons learned from this was the variation of slow shot speed that was experienced. In the typical position versus velocity shot profile, the difference between an average of 5 inches per second compared to 10 inches per second is not visually as significant as when a fast shot velocity scales the entire graph. Looking at these results on a time versus velocity graph will show considerable time delays in the slower speed. One of the future areas of investigation is applying the entire shot profile through a machine learning algorithm to better detect changes that may have an impact on the process. One area where data is lacking in the standard statistics used from shot end-monitors is the application of "low impact" or "braking the shot rod" at the end of the filling cycle. This can be visually seen on a profile trace, but lacks a parameter statistic. Use of the entire profile may help with future predictions.

The same logic can be applied to the thermal images. In this study, the images were summarized with a median statistic in the cavity area. Additional development in programming and means to detect image anomalies is needed to make this additional data even more robust. Those types of programming scripts were not a major focus here but present many opportunities for further improvement.

Future Machine Learning – Complex Algorithms

Although the more complex machine learning algorithms like neural nets and support vector machines failed to produce useful results in this experiment, this does not mean they should be ignored for future work. These algorithms can handle highly complex data sets and provide excellent predictions and insight. Additional images and time-series data, as discussed in the previous section, are prime sources of data sets that may provide the insight needed to make accurate predictions.

A user of machine learning should become familiar with all these different tools and be able to apply them as needed. This was an introduction of the topic to a die caster with no familiarity with the topic. The amount of open source information, knowledge, and programs on machine learning can be overwhelming. This is a result of the vast use of machine learning algorithms by large, technology based companies.

Lower Pressure = Better Castings?

Perhaps the most interesting casting process learning from this experiment was the results of the intensification pressure on Good versus Bad castings. The results from the experimentation showed that lower intensification pressure created a higher yielding casting after machining. When dealing with a porosity in an isolated heavy section within the casting, this approach seems counter-intuitive from everything discussed in die casting and feeding shrink areas.

This concept is one of the areas where additional research and experimentation will be completed to see if this hypothesis holds true over multiple casting designs that struggle from an isolated heavy section. Further analysis is needed to understand why this may be happening form a metallurgical standpoint or if it is just an anomaly of this casting and tool design. Regardless, this result does present an intriguing finding from the experiment worthy of greater study.

CONCLUSIONS

This experiment has demonstrated that using machine learning algorithms can provide additional insight into the process. Machine learning was shown to assist with finding patterns in the data when they otherwise are unclear. Gaining this insight and moving beyond world-class scrap rates was the motivation for this work. Improving scrap is a challenge because die casting is a highly complex, multivariate, stochastic process of transforming liquid metal into a near net shape casting. This type of problem would have been already solved if it only involved a singular parameter that could be applied to any machine and tool combination. Instead, there is an infinite number of parameter choices to make as different castings have their process developed. The issue is finding a way to efficiently cycle through all this data to make sure die casters consistently make the best choices for the process. Machine learning algorithms can help solve this problem. This work should spark additional conversation, research, and development within the industry to continue to improve die casting with a more robust process.

CASE STUDY II: UNSUPERVISED MACHINE LEARNING AND STATISTICAL ANOMALY DETECTION APPLIED TO THERMAL IMAGES [4] [9]

Processing and analyzing continuous data is a challenge in the die cast industry. Standard practice is to collect discrete performance data on a small number of variables. Decisions are often made from a portion of this small data set on a variable-by-variable basis. Analytics become exponentially more challenging with two-dimensional (2D) data sets, such as thermal images of the dies or X-ray pictures of castings. Instead of tens or hundreds of variables, these 2D data sets can create more than a million data points for each casting. Significant changes in the images are easy to detect, but slight anomalies or trending can be challenging or time consuming for people to identify. It is possible to efficiently evaluate this type of data through machine learning.

This chapter will review an unsupervised clustering, semi-supervised training, and a statistical analytics algorithm to automate the analysis of images data sets. Thermal images of die halves taken before the die is closed for a shot will be used in this analysis. A three-step algorithm will be implemented that reviews every thermal image to detect and identify anomalies without requiring detailed results of thousands of images. This algorithm could also be applied to X-ray and other 2D data sets within the industry.

INTRODUCTION

Die casting is a hydraulic, mechanical, and thermal process [1]. Arguably, all three have a significant importance in producing high quality castings. The hydraulic and mechanical portions of the process have an advantage, since data collection systems monitoring these processes have existed for decades. The industry is familiar with reviewing data such as cycle time rates, tie bar tonnages, velocity profiles, and pressure profiles. Dedicated suppliers provide hardware and software for this data collection that is standard on new equipment investments.

[4] This section is an edited version of a conference paper, included with permission from the North American Die Casting Association. D. Blondheim, "Unsupervised Machine Learning and Statistical Anomaly Detection Applied to Thermal Images," 2018 NADCA Congress and Exposition. Indianapolis, IN, Oct. 2018

The thermal portion of the die casting process has lagged in data collection and analysis. Measuring temperature and flow rates of cooling lines provides information on the heat removed from the die by cooling lines, but this does not directly translate to die surface temperatures or incorporate the spray cycle impact. Embedded thermocouples in dies have been used to detect temperature changes within the internal die steel at specific locations. However, they fail to collect the surface temperatures that will be in direct contact with the liquid metal at injection. Hand-held infrared cameras can provide a more complete picture of the die surfaces, but they require stopping the machine, so the die can remain open for the images to be taken. The impact to the cycle drives these hand-held thermal images to be infrequent audits of the process. The continuous data the industry is familiar with, such as velocity profiles, has not existed in the past for thermal images.

More recently, there has been some early adoption in mounted thermal cameras that take thermal images during every cycle [45], [223]. Use of this type of mounted thermal camera system in a production setting highlights a few difficulties with the typical approach of data analysis on images. Typically, software provided with these thermal camera systems allows users the ability to create a region (square, circle, rectangle, etc.) on the image and trend the minimum, maximum, and average temperatures within that area. For some features, like a critical jet cool core, it is easy to understand and select this region as one to track and trend with the software. What is more difficult to identify are other, more subtle critical areas within the thermal process. It is near impossible for a user to know where to select regions and tie them to specific water cooling channels or die lube spray nozzle paths, especially with hundreds of potential water lines in a die and spray nozzles within a spray manifold.

Another difficulty is understanding the variation that occurs within the die naturally. Steady state conditions in the die may be better classified as "quasi equilibrium," since small temperature changes within the die may accumulate over time [44]. The inputs to the thermal process are also not the constants that are assumed in most simulations. Furnace temperature of the liquid metal typically follows a sine-wave pattern. Cooling water flow rates vary based on usage throughout the plant. Die lube temperatures may be impacted by the external climate. These variables, as well as countless others, all impact the ability of the die to hit a steady state surface temperature condition. This variation makes it difficult to understand which temperature changes seen on the thermal image are part of the normal process variation (common cause variation) and which are due to a failure within the thermal cycle (special cause variation). If this understanding does not exist, setting up thresholds for warnings in the regions the user selects is exceedingly difficult and therefore often left incomplete.

The final difficulty to mention is not an issue with the thermal camera system itself but a problem of part tracking and data collection systems. In the ideal situation, all the parts produced would be properly tracked through the supply chain and accurately documented with rejection and acceptance data. With these results, advanced supervised machine learning algorithms can be applied to the image files to ideally predict potential failures based on the thermal images. Getting this ground truth data on the thermal images is costly, time consuming, and difficult to implement due to the amount of effort.

Large image analysis projects are a challenge for human vision. Sudden changes in a small area of an image can be easily detected by humans. However, gradual changes of an image are not easily observed due to our ability to adapt to brightness changes. It is impossible for humans to extract specific measurement temperature data out of thermal image files [224], whether the image is in grayscale or shades of color. Because of these difficulties, an innovative approach in image analysis is needed, reducing the effort required for the die cast industry to utilize these types of technologies. Outlined here is a three-step anomaly detection algorithm that can be used on thermal images. The approach is general and can also be used on other repeatable image files such as x-ray images.

ANOMALY DETECTION ALGORITHM

A three-step anomaly detection algorithm was developed to detect and identify regions within the thermal image that are statistical anomalies from a known condition. The first step in this algorithm is to complete an unsupervised k-means clustering on a representative training data set. A minimum of several hundred images should be used. Ideally, several days or weeks of production should be reviewed, which would represent thousands of parts. Once clusters are identified through the first step, they would be reviewed by subject matter experts. The semi-supervised training occurs when the expert selects the cluster grouping that represents the steady state condition of the die. With the steady state cluster selected, a statistical model is created to which any future image can be compared for anomaly detection. Figure 76 shows the anomaly detection algorithm, the needed inputs, and the output.

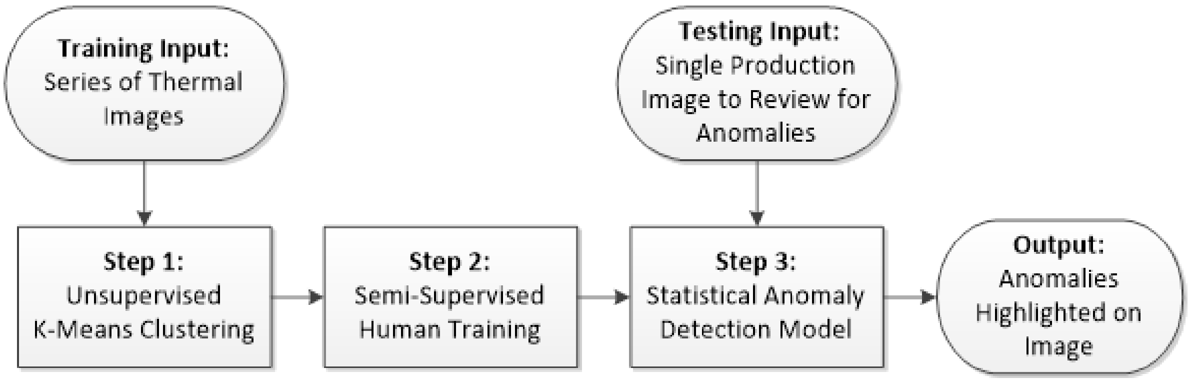


Figure 76: Three-step algorithm for anomaly detection

Clustering Algorithm

K-means clustering algorithm is an unsupervised clustering algorithm that can be used to group similar unclassified data sets. Other clustering algorithms exist, and in some situations, may be more appropriate. However, they were not reviewed here because the simple k-means clustering algorithm proved to be successful in the testing done during the algorithm's development. The k-means clustering algorithm works by dividing a sample of size *n* into *k* different clusters, where each sample is measured to the nearest mean value that serves as an average or representative sample of that cluster. The number of clusters the sample is to be divided into is a user-defined value. Observations are assigned to each cluster to reduce the within cluster sum of squares or variance. Mathematically this is defined in Equation 23.

Equation 23

$$\begin{aligned}
&\boldsymbol{Given}: \\
&\boldsymbol{observations} = (\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3} \ldots \boldsymbol{x_n}) \\
&\boldsymbol{number\ of\ clusters} = \boldsymbol{k} \\
&\boldsymbol{Cluster\ sets} = (\boldsymbol{S_1}, \boldsymbol{S_2}, \boldsymbol{S_3} \ldots . \boldsymbol{S_k}) \\
&\\
&\boldsymbol{Objective}: \\
&\arg\min \sum_{i=1}^{k} \sum_{x \in S_i} \|x - \mu_i\|^2 = \arg\min \sum_{i=1}^{k} |S_i|\, Var\ S_i \\
&\boldsymbol{Where}: \\
&\mu_i\ is\ the\ mean\ of\ S_i
\end{aligned}$$

The k-means algorithm needs an initial set of cluster mean values. In this example, a random set of k points is used as this initial set. An iterative process then takes place:

Assignment Step: Assign a sample to the cluster with the nearest Euclidean distance to the cluster mean.

Update Mean Step: With samples assigned, calculate a new mean value for that cluster.

Repeat: Repeat steps until there are no new assignments of points to new clusters; at this point, the algorithm has converged.

A k-means algorithm will converge to a local minimum based on the initial assignment step. There is no guarantee that the solution found will be the global minimum because of the randomly assigned initial starting values. To overcome this issue, the k-means clustering algorithm is ran multiple times with new random initial values at each start. This usually ensures that the global minimum is found [222]. In this work, 100 sets of random initial values were used.

The difficult decision within the k-means algorithm is to determine optimum cluster size k. The method used here is called the F-Score Approach. The F-Score is similar to the F-test completed during ANOVA analysis. The Sum Square Error (SSE) is calculated both within the clusters and between the clusters. Using the SSE and the degrees of freedom (DF), a ratio is created to provide an F-Score. This score is calculated using Equation 24 by looping through the k-means calculations for multiple cluster sizes. The cluster size that produces the maximum F-score is selected as the optimum and would be used for the analysis [225].

Equation 24

$$\boldsymbol{F = \frac{SSE_b / df_b}{SSE_w / df_w}}$$

Where:

SSE_b = *Sum Squared Error between the clusters*
df_b = *degrees of freedom between the clusters*
SSE_w = *Sum Square Error within the cluster*
df_w = *degrees of freedom within the cluster*

Semi-Supervised Training

The second step is referred to as "semi-supervised training." This step requires a subject matter expert to review the cluster assignment data to determine which of the clusters best represents the steady state condition of the process. This is the step that requires human input into the algorithm. There are several types of data to be reviewed to assist with determining the steady state cluster.

The output of the clustering step is an assignment list of the thermal images to assigned clusters. Through programming, these images can be turned into image stacks, which represent all the images for each cluster. A stack becomes a three-dimensional array of data with the first two dimensions being the image width and height and the third dimension being the set of images that belong to that cluster. With the images in a stack, mean and standard deviation matrices can be calculated for all the pixels within the image. Statistics and histograms calculated from these matrices are used to help identify differences between clusters and which cluster represents the steady state. For example, a production cycle during the warm-up process will have a significantly lower mean temperature distribution than a steady state production part.

Another set of data that can be reviewed to identify steady state is cluster size and occurrence. In the ideal process, the cluster size for steady state would be significantly larger than all the other clusters identified because it should be the default condition of the tool in production. Clusters should identify warmup conditions, since the dies are significantly cooler during these times, and potentially a transition period from a cold die at start of warmup cycle to a moderately hot die that is not yet fully at steady state. The time series occurrence of the cluster assignment is another indication as to which cluster represents the steady state condition. Ideally, the steady state condition will appear consistently throughout the entire series of data being reviewed.

A third option to review would be trended points on the die surface across the series of images. A graph could be produced that shows the temperature readings with different colored points to represent the different cluster assignments. Since the k-means algorithm does not supply direct output describing which input values were used to cluster, this approach may be a bit more "guess and check" than desired.

Detailed knowledge of the part and tool may eliminate this “guess and check” method. It may be preferred to complete analysis of the standard deviation matrix and its corresponding histogram to gain insight on variation of the steady state cluster.

These are just a few examples of data to review to determine the steady state cluster. Other data such as time between shots may also be reviewed and used to assist with the selection. The key understanding is that the steady state cluster selection is a critical step in the algorithm to ensure the rest of the algorithm is built to properly detect anomalies. It is reliant on input from a human with domain knowledge in processing of die castings.

Statistical Anomaly Detection

With a steady state cluster identified, the last step in the algorithm is to produce an anomaly detection model, which can be used to review any new images that are produced in production. This anomaly detection model was based on a Shewhart statistical process control (SPC) chart [226]. The typical SPC chart plots one variable over time. This approach would be too time consuming and difficult to comprehend with a 2D image file that contains thousands of pixels requiring analysis.

With a thermal image, a creative new way to implement an SPC approach was developed for 2D data sets to visualize where anomalies occur on the image. Mean and standard deviation matrices are calculated from the steady state image cluster. These matrices would reflect the average and typical levels of variation in the temperature readings of each pixel within the thermal image. Figure 77 provides a visual example on the mean and standard deviation matrices.

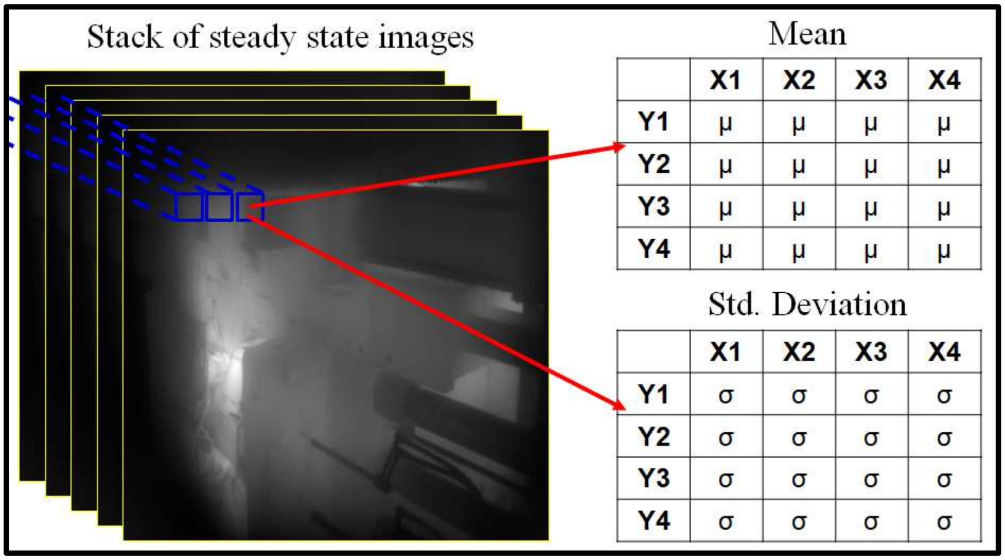


Figure 77: Visualization of steady state image stack with mean and standard deviation matrices

With these matrices in place, a statistical z-score can be calculated on a pixel-by-pixel basis for any new image input into the algorithm. The z-score represents the number of standard deviations a sample is from the mean and is calculated from Equation 25.

Equation 25

$$\mathbf{z}_{x,y} = \frac{(\mathbf{T}_{x,y} - \boldsymbol{\mu}_{x,y})}{\boldsymbol{\sigma}_{x,y}}$$

Where:

$z_{x,y} = Z\ score\ of\ sample\ image\ at\ pixle\ location\ (x, y)$

$T_{x,y} = temperature\ of\ sample\ image\ at\ pixel\ location\ (x, y)$

$\mu_{x,y} = mean\ temperature\ of\ steady\ state\ cluster\ at\ pixel\ location\ (x, y)$

$\sigma_{x,y} = standard\ deviation\ of\ temperature\ of\ steady\ state\ cluster\ at\ pixel\ location\ (x, y)$

A z-score value that is zero would represent a pixel that is at the mean of the steady state cluster and would be considered similar to the steady state. Statistically speaking, the null hypothesis that the new image is the same as the steady state cluster could not be rejected. As the new pixel temp value moves away from the mean, either positively or negatively, it would represent a larger magnitude z-score. Therefore, it would be less likely the new pixel temperature is from the same distribution as the steady state image. Outside certain z-score boundaries, the null hypothesis that the new image and the steady state are the same would be rejected.

Much like the traditional SPC charts, a threshold needs to be set to represent the number of standard deviations from the nominal that is allowed before the value is considered out of control or an anomaly. Often in SPC charts, a threshold of 3 standard deviations (also commonly referred to as 3σ) is used. ±3σ from nominal represents 99.7% of the area under a normal distribution curve. In the code developed, this threshold is programmable by the user so that he or she can test different threshold ranges. The results here are based on a threshold z-score of ±3.

The result of this step is a matrix of z-scores representing each pixel of the thermal image. This alone does not provide value to visualize where the anomalies occur since an image of z-scores will look nothing like the die in question. Additional mathematical modification is required to convert the original thermal image into an output image that highlights the anomalies.

Output Mathematics

The output of this final step needs to combine the z-score matrix produced in the previous step with the original tested thermal image to provide an image file a human can view. The required output is a grayscale thermal image based on the temperature readings that would highlight areas of the die that were statistically hotter (in red) or colder (in blue) than the steady state cluster. To do this, the original grayscale image needed to be replicated into three layers to represent the Red, Green, Blue (RGB) used in color image processing. Then all three layers layers need to be modified with the z-score values above or below the threshold value to create an image that remains grayscale for areas within the threshold limits, but blue in areas below the negative threshold and red in areas above the positive threshold. This would produce an image that can easily be reviewed and understood by operators.

Colored images require matrix values between 0 and 1. The first step is to convert the temperature readings from the new test image to this 0 to 1 range using an assigned minimum and maximum value from the entire range of images as seen in Equation 26 .

Equation 26
$$\boldsymbol{pixel\ range\ [0,1]} = \frac{(\boldsymbol{test\ image\ pixel\ temp - minimum\ global\ temp})}{(\boldsymbol{maximum\ global\ temp - minimum\ global\ temp})}$$

If both die halves are used and combined into one image, which was done for this work, separate minimum and maximum values may be required for the moving and the stationary die half images to provide the correct visualization of the die halves. If one of the die halves has a narrow temperature range, it will appear washed out or undefined compared to the other half. Different minimum and maximum values between halves can fix that problem. It is important to note that the image produced is no longer an actual temperature output. The images are converted for a visual image for the operator only. Once this range adjustment is made, this matrix, which is called **[GS]** (original Gray Scale), with the z-score matrix, which will be called **[Z]**, is used for remaining matrix math to create the RGB layers. An "identity type" matrix of 0s and 1s is developed based on the different z-scores for the different layers. This is done to turn on or off the coloring to get the red and blue anomaly highlights as described in the previous section. Entire matrices will be labeled in brackets [ ], an individual location within that matrix will be listed without the brackets (ie: Z versus [Z]). When there is a multiplication, it is assumed to be an elementwise multiplication between matrices (also called Hadamard product or entrywise product – represented in several programming languages as: .*) and not matrix multiplication. This mathematical code is seen in Figure 78.

```
For the Green layer matrix [G], the math to complete the section follows:

Let [ZG] = [Z]
IF ZG is > threshold, THEN change ZG value to 0
IF ZG is < (-1*threshold), THEN change ZG value to 0
IF ZG is >=(-1*threshold) AND ZG is <= threshold, THEN change ZG value to 1
[G] = [GS] .* [ZG] (elementwise multiplication)


For the Blue layer:

Let [ZBNeg] = [Z]
IF ZBNeg >= (-1*threshold), THEN change ZBNeg to 1
IF ZBNeg < (-1*threshold), THEN change ZBNeg to 100,000
Let [ZBPos] = [Z]
IF ZBPos <= threshold, THEN change ZBPos to 1
IF ZBPos > threshold, THEN change ZBPos to 0
[B] = [Z] .* [ZBNeg] .* [ZBPos] (elementwise multiplication)
IF B > 1, THEN change B to 1


For the Red layer:

Let [ZRNeg] = [Z]
IF ZRNeg >= (-1*threshold), THEN change ZRNeg to 1
IF ZRNeg < (-1*threshold), THEN change ZRNeg to 0
Let [ZBPos] = [Z]
IF ZRPos <= threshold, THEN change ZRPos to 1
IF ZRPos > threshold, THEN change ZRPos to 100,000
[R] = [Z] .* [ZRNeg] .* [ZRPos] (elementwise multiplication)
IF R > 1, THEN change R to 1
```

Figure 78: Mathematical code for different anomaly identification per color layer

The final image file is created by the combination of the [R], [G], [B] matrixes.

Data Collection and Pre-Processing

The thermal images used to develop this algorithm were gathered using an infrared camera system created with the assistance of a third-party integrator. Two cameras are mounted on opposite sides of the die cast machine in an enclosure, that opens for the images to be taken immediately before the die closes. Figure 79 shows the camera system mounted on the moving half side, pointed at the stationary half die.

Figure 79: Thermal image camera system install on die cast machine [45]

The camera system stores a TIFF file for both die halves. This file saves the actual temperature reading of the die instead of a scaled 0 to 1 grayscale value typical of image files. This prevents the TIFF file from being opened in most Windows based image viewers but allows a full matrix of temperature readings to be uploaded into analytics software such as *R* [215]. *R* and other software can automatically scale a matrix from 0 to 1 based on values of the matrix to allow it to be viewable. Figure 80 shows an example of the gray scale image produced from the temperature matrix.

297.91	297.95	297.92	297.88	297.90	298.02	298.00	298.03	298.06	298.02
297.92	297.89	297.98	297.98	297.99	297.97	298.09	298.04	298.07	298.01
297.88	297.91	297.96	297.97	298.02	297.95	298.05	297.99	297.97	298.02
297.88	297.85	297.88	297.99	298.00	297.97	297.96	298.08	297.95	298.02
297.94	297.92	297.95	297.96	297.99	297.95	298.07	298.00	298.04	298.05
298.00	297.86	297.97	297.98	298.01	298.00	298.01	298.01	298.08	298.05
297.99	297.98	297.94	297.99	297.97	297.98	298.02	298.04	298.06	298.06
297.87	297.91	297.97	297.99	297.94	297.96	298.01	298.02	298.22	298.06
297.90	297.95	297.99	297.83	298.00	297.99	298.02	298.00	298.23	298.03
297.92	297.85	297.83	298.00	298.00	297.92	298.00	298.07	298.04	298.20
297.97	297.99	297.86	297.99	298.02	297.85	298.03	298.06	298.04	298.04
297.91	297.95	297.93	297.99	297.95	297.97	298.08	298.06	298.04	298.04
297.95	297.95	298.06	297.92	297.92	297.92	298.06	297.92	297.99	298.05
297.93	297.93	297.94	297.99	297.96	297.98	298.09	298.14	298.13	298.06

Figure 80: Thermal image temperature matrix and grayscale image

To be successful with the image analysis, it is critical to understand that the camera system taking the images must be fixed and produce a repeatable image. The clustering algorithms used are simplistic and do not search through an image like a convolutional neural net (CNN) would. This makes it significantly easier to train and implement but requires dedication to the stability of the images. Image registration, or alignment to a target image, can help with drift or repositioning of the camera, but it was not required within this application due to the camera mounting.

Before the clustering algorithm was ran, the image files were preprocessed as follows:

1. Cropping of both stationary and moving half images individually
2. Pixel scaling of both stationary and moving half images individually
3. Combining the two cropped and scaled images into one matrix file
4. Vectorizing the 2D matrix to a 1D vector labeled with a serial number tag

Original image files were 480 pixels x 640 pixels x 2 images = 614,400 data points per casting. Due to lenses required and the camera mounting locations, a portion of the image file does not contain useful cavity temperatures. These areas were cropped out to reduce the amount of data to be processed. An example of this cropping is seen in Figure 81.

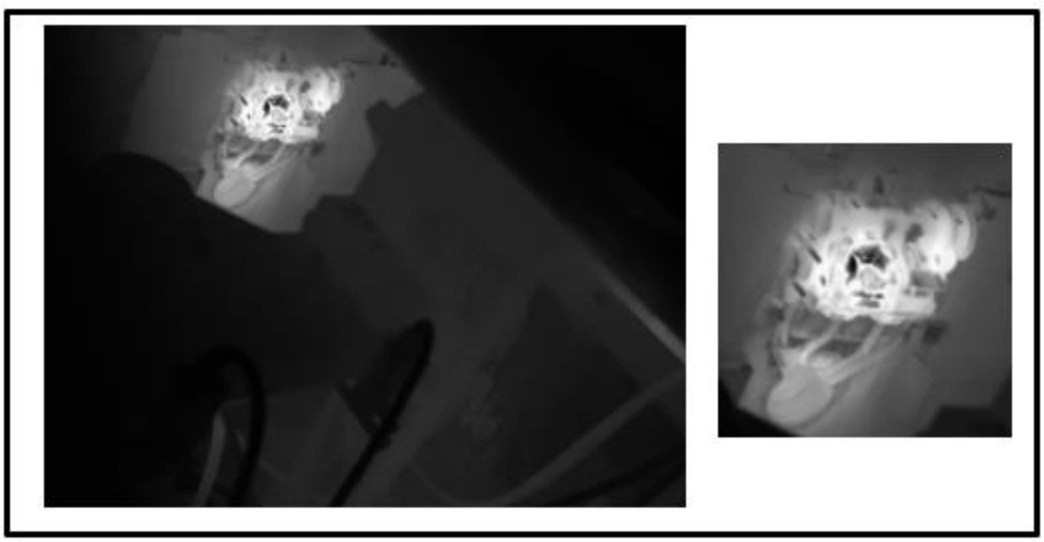


Figure 81: Pre-crop versus post-crop thermal image (stationary half)

After cropping, the images were scaled to 90 x 90 pixels. A 3x3 or 2x2 average was calculated on the image to produce the scaled pixel image. Image pixels were scaled to allow for data dimension reduction, but more importantly, this would also average out any small image shift that could occur image to image. Figure 82 shows an example of pixel scaling used.

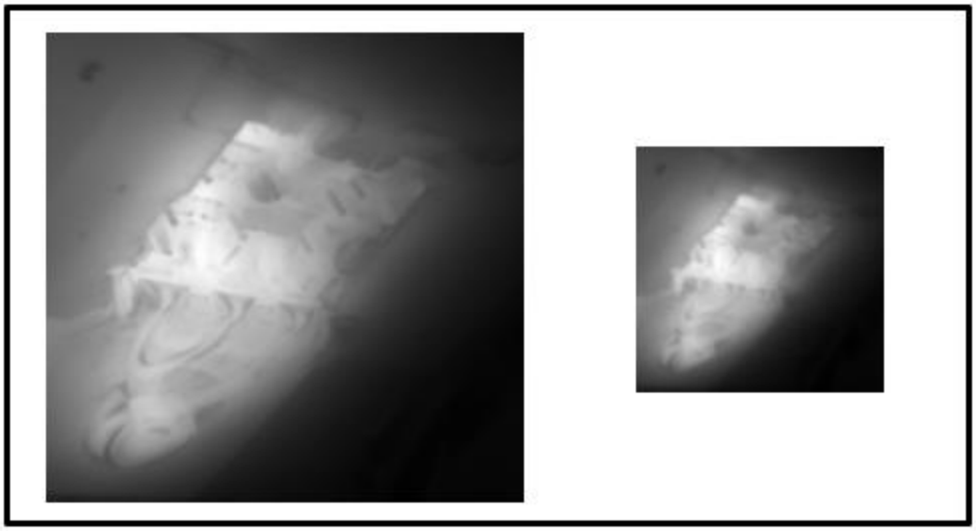


Figure 82: Comparison of scaled pixel thermal images (moving half)

Finally, the two images were merged into one matrix of size 90 x 180 pixels or 16,200 data points. This was a significant data reduction from the original image files, but the critical cavity temperature data was maintained. A *for-loop* was programmed to cycle through all the images to create this combined image, which was then vectorized into a 1D data string and merged with all the other images. A data set was created that contains a serial number label and a vectorized temperature list.

With this data set, the k-means clustering algorithm is ran to assign serial numbers to a cluster. In a clustering algorithm, the data columns typically represent multiple input variables with different value ranges. These different variables are required to be normalized before running the k-means algorithm. This is done because the k-means clustering is dependent on distance calculations and would be incorrectly influenced on large values. Normalization does not need to happen with this application of 2D images. The thermal image is one set of temperature inputs located across a 2-dimensional surface of the die. Normalizing would affect each pixel location individually, and it would lose its relationship with other areas on the die.

For this algorithm, an *R* software script was created to complete the operations of loading the images, cropping, pixel scaling, combining, and clustering.

RESULTS

To validate this algorithm, thermal image data was pulled from the production system and analyzed through the steps. A total of 894 images were used to develop the algorithm. This represented production over multiple shifts on a Friday and Monday. During this time, there were several downtime events during production, including a shutdown over a weekend. As a result, several warmup cycles happened during this time range.

Clustering

The *R* package *stats* [215] contains the *kmeans* function that was used to complete the clustering analysis. Initially, cluster sizes of k = 2 to k = 30 were looped through the package to cluster the images. The k-means clustering algorithm was set up to run with 100 random starting points to help find the global minimum. The algorithm was also programmed to stop at 500 iterations of the algorithm if an optimal solution was not found. After the clustering was ran, the algorithm found a stopping point typically within 6 to 9 iterations, so the 500 maximum iterations was adequate. With the *kmeans* function, the data required to complete the F-score for each cluster size is extracted and saved. A graph of the F-score per cluster size can be seen in Figure 83. Based on this, an optimal cluster size of k = 4 was selected.

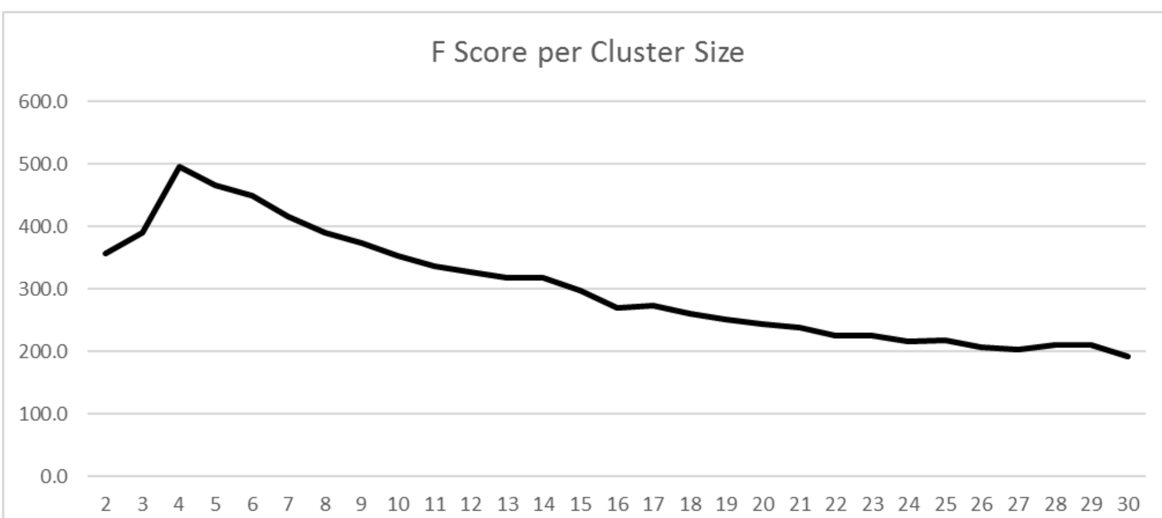


Figure 83: F-Scores per cluster size

Semi-Supervised Learning

The next step is to identify the cluster that best represents the steady state temperature profile of the die. The clustering process will produce statistically unique groups based on minimizing the distance within the cluster of the input variables. In this example, these input variables are the thermal picture pixel location temperature measurements.

Since there were no major, purposeful, thermal changes made to the process during this run, the clustering algorithm worked extremely well identifying a steady state cluster based on size of the cluster assignment alone. Table 33 shows the complete breakdown of the 894 images to each cluster assignment, with cluster #1 having the largest assignment of 738 parts.

Table 33: Cluster Assignment Sizes

Cluster Assignment Number	Number of Samples in Cluster	% of Total Population
#1	738	82.6%
#2	121	13.5%
#3	21	2.3%
#4	14	1.6%

Beyond the number of occurrences, it is useful to look at the sequence of cluster assignments. Ideally one would like to see clusters grouped in a meaningful way, such as a cluster of cold images showing up every time there is a downtime event. Figure 84 shows a straight sequence based cluster assignment (1 to 894), while Figure 85 is the cluster assignment based on a date and time stamp of when the image was taken. Figure 85 helps highlight downtime events, such as the weekend, and how clusters are assigned near these events.

To validate correlation to warmup shot conditions, a table showing the number of warmup cycles per cluster was created. Warmup shots are the immediate three castings produced after a downtime event. A downtime event is defined as any time the cycle took more than 10 times the standard cycle time. Table 34 shows this data. Cluster #3 had almost 81% of its cycles classified as part of the warmup process after a downtime event. Cluster #1 had two shots classified as warmup. Further review of these two shots found they both were shot number 3 during a shorter downtime event period. This has led to further investigating of the heuristic rules regarding the warmup process.

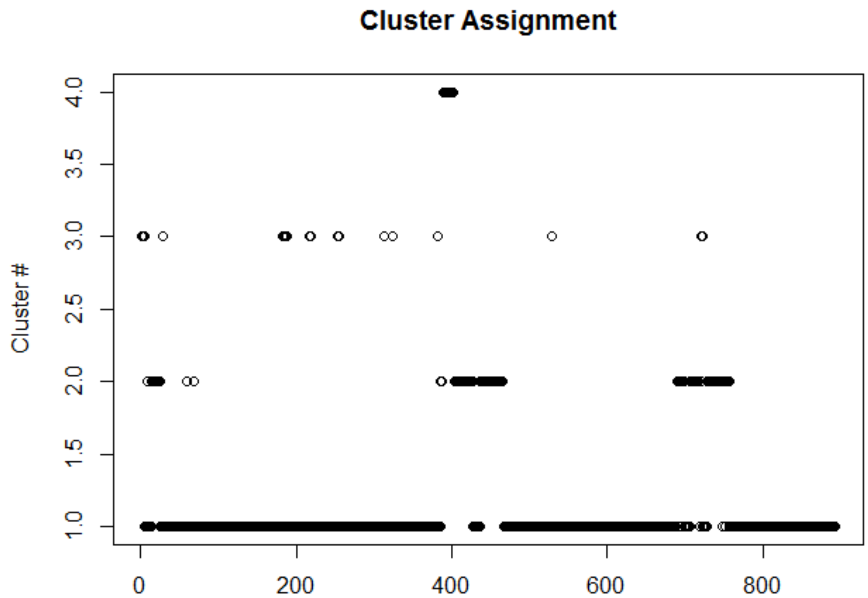


Figure 84: Sequence of cluster assignments

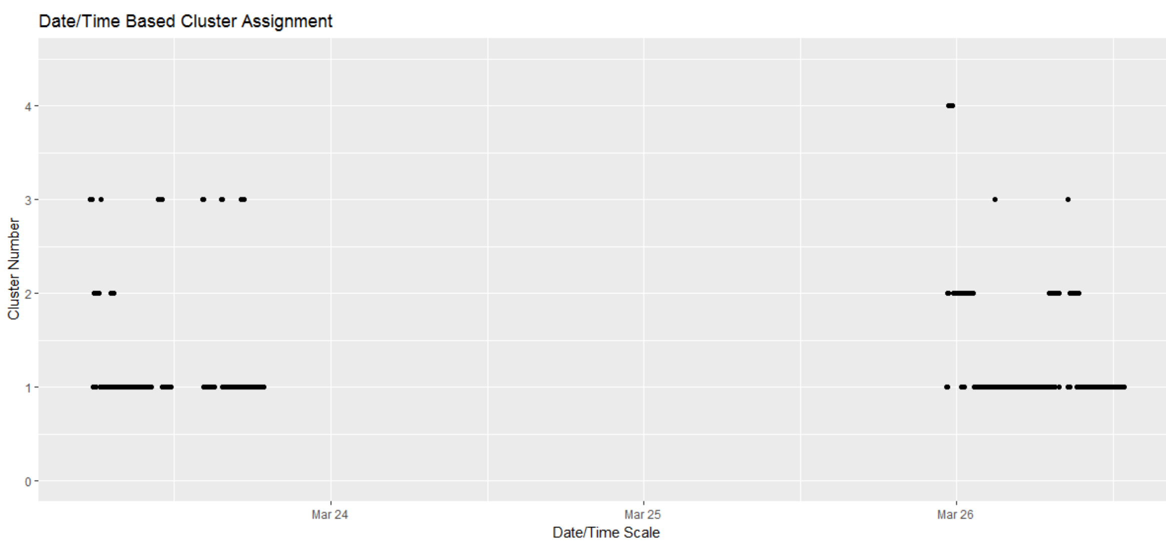


Figure 85: Time-based sequence of cluster assignments

Table 34: Cluster Assignment Versus Warmup Shot Counts

Cluster Assignment Number	Number of Samples in Cluster	Number of Warmup Shots	% of Warmup Shots
#1	738	2	0.27%
#2	121	0	0.0%
#3	21	17	80.95%
#4	14	0	0.0%

From this information, a convincing argument can be made that cluster #1 is best defined as the steady state cluster. To further validate this claim, image stacks of this cluster can be created, and a grayscale image of the mean and standard deviation of each pixel location can be calculated. By plotting these as a grayscale image, additional visual insight can be gained when compared to other clusters or differences between clusters. The critical point to remember is that a grayscale image will have the lowest number assigned as "black" and the highest value as "white." The rest of the image is scaled gray between these, so it is not possible to compare directly two different grayscale images. Figure 86 shows the mean and standard deviation matrix of the cluster #1. The mean image will highlight the general temperature gradient, while the standard deviation image will highlight areas on the die that see the largest temperature variation in lighter shades of gray to white.

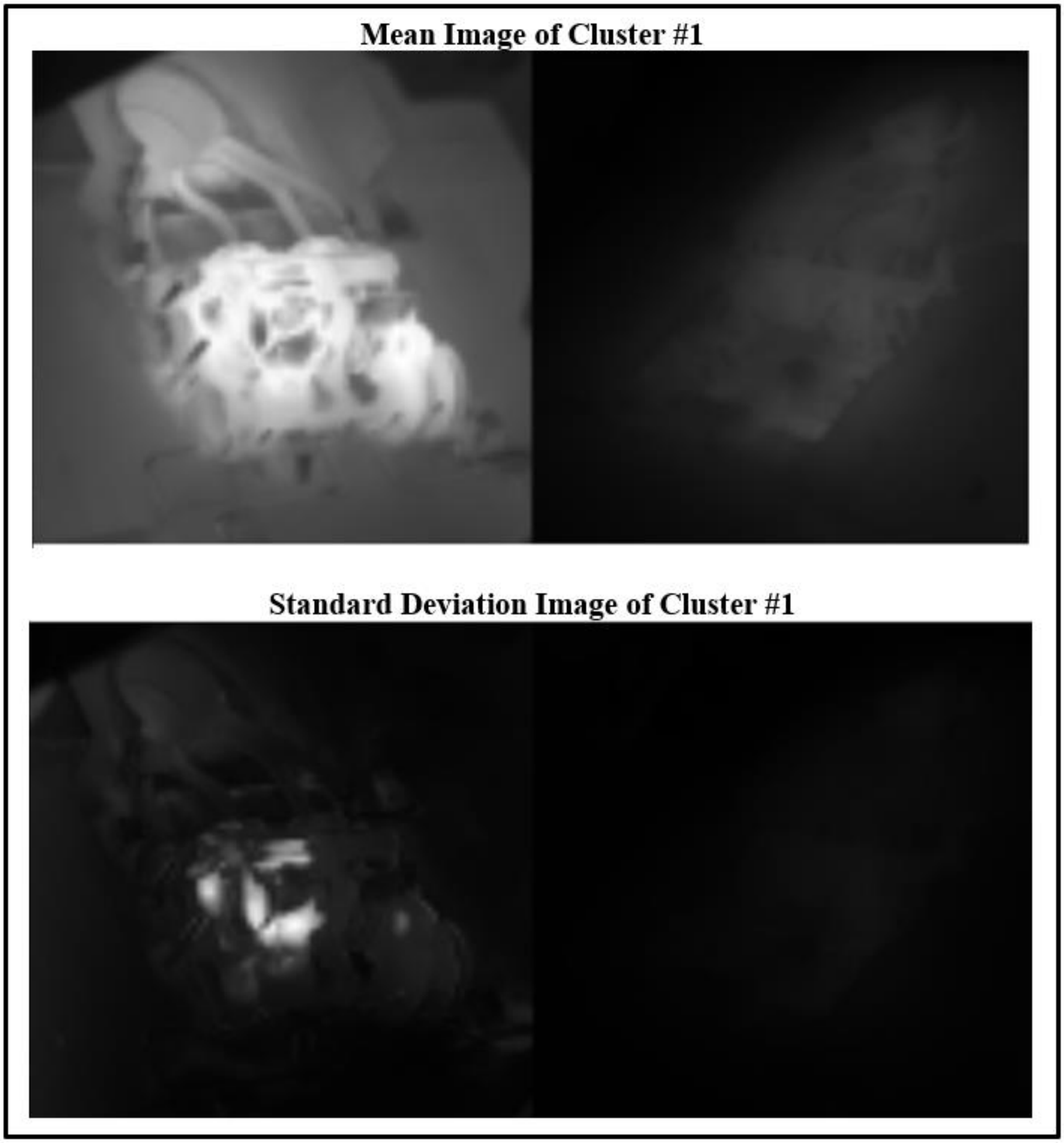


Figure 86: Mean and standard deviation images for cluster #1

Another useful way of comparing cluster mean images is with a histogram of the different temperature points. The shape and distribution of the histogram can provide insight that image visualization loses due to the scaling properties of images. In Figure 87, the histogram of cluster #1 has significantly more points at the hotter temperature range when compared to the histogram of cluster #3.

Cluster #3's histogram has most of its temperatures near the 320 to 330 range. Given that cluster #3 is tied to the warmup process (as seen in Table 34), this shift in temperature readings is expected.

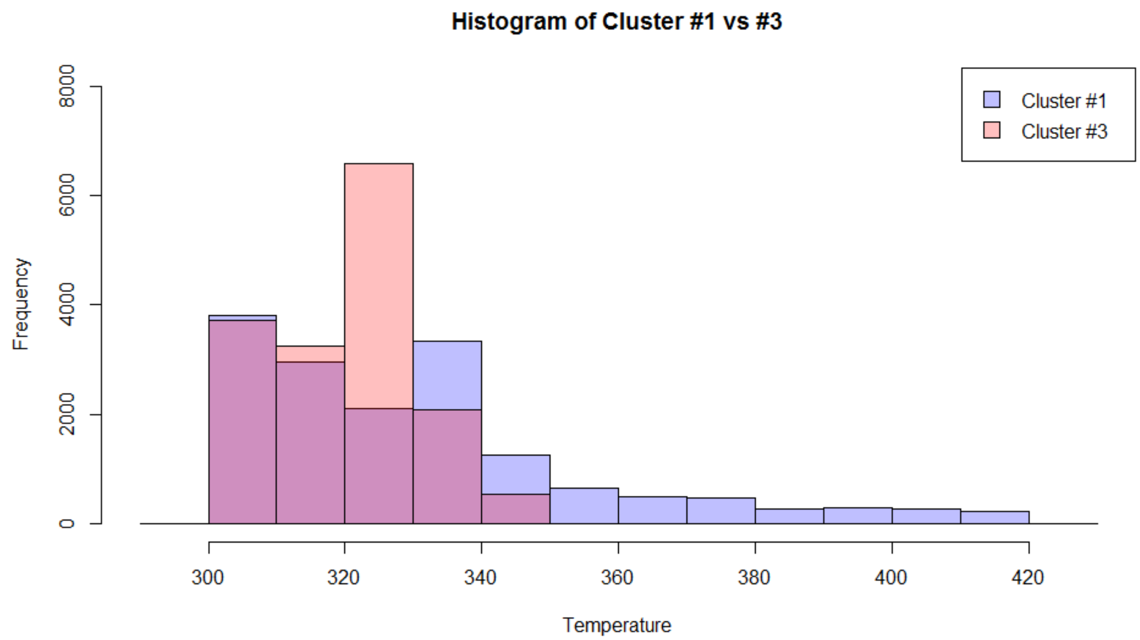


Figure 87: Histogram of mean image of cluster #1 versus cluster #3

As previously mentioned, the graphs and images used in this section are some ways the data available can be used to select the steady state cluster. This assignment is a critical step in the algorithm. Cluster selection becomes the basis for the model used to compare all future thermal images. Selecting a poor representative cluster will mean incorrect statistical anomaly detection. In the multiple sets of images ran through the algorithm in preparation for this work, it has been straightforward to find the steady state cluster for the set of thermal images in review. This may not be the case in all situations.

Along with the steady state condition (cluster #1) and the warm-up cycle condition (cluster #3), the unsupervised algorithm also detected two other clusters. Cluster #4 was a situation where the die got extremely hot. A manual water valve was accidently turned off prior to starting up. The result was an extremely hot die. This failure was identified through shift notes that were captured the day of the testing. Cluster #2 presented a different challenge, since it was not readily identifiable as to why this group existed. This highlights the importance of having this process real time on the production floor to better troubleshoot these types of anomalies.

Statistical Anomaly Detection

With the steady state cluster now identified, the last step of the algorithm can be completed. An image stack is created from the thermal images assigned to the steady state cluster. The mean and standard deviation matrix for each of the pixel locations within the image are calculated. This creates two matrices of data, which will be used to calculate z-scores. Each pixel location is assumed to be normally distributed within a cluster, and a z-score calculation will provide the statistical probability that a new image would belong to the distribution of the steady state cluster. If there is a large z-score, beyond the defined threshold, it is assumed the given pixel location is an anomaly to the process. It will then be appropriately flagged in the image.

The z-score calculation is straightforward. Creating the color image from an input temperature matrix and the z-score matrix created is more difficult due to the multiple layers required for a color image. Following the equations defined earlier will ensure all the color layers are appropriately created. With a script created containing the mean matrix, the standard deviation matrix, and the layer color code, one can almost instantaneously select an image file and produce a thermal image with anomalies colored red or blue to highlight statistical temperature differences. Figure 88 shows a standard warmup shot after an extended downtime when compared to the steady state cluster. Figure 89 shows a casting from cluster #4 where the water was identified as being turned off. The thermal camera is quick to pick up the temperature changes within the die, while the algorithm was able to correctly classify this as a separate group of parts.

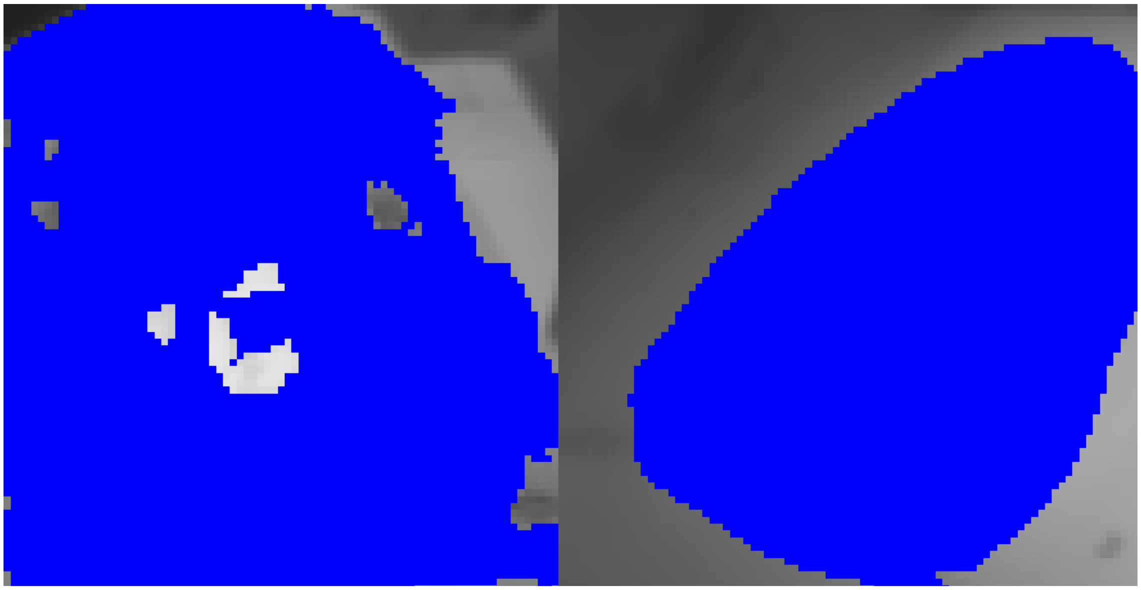


Figure 88: Anomaly detection on warmup shot

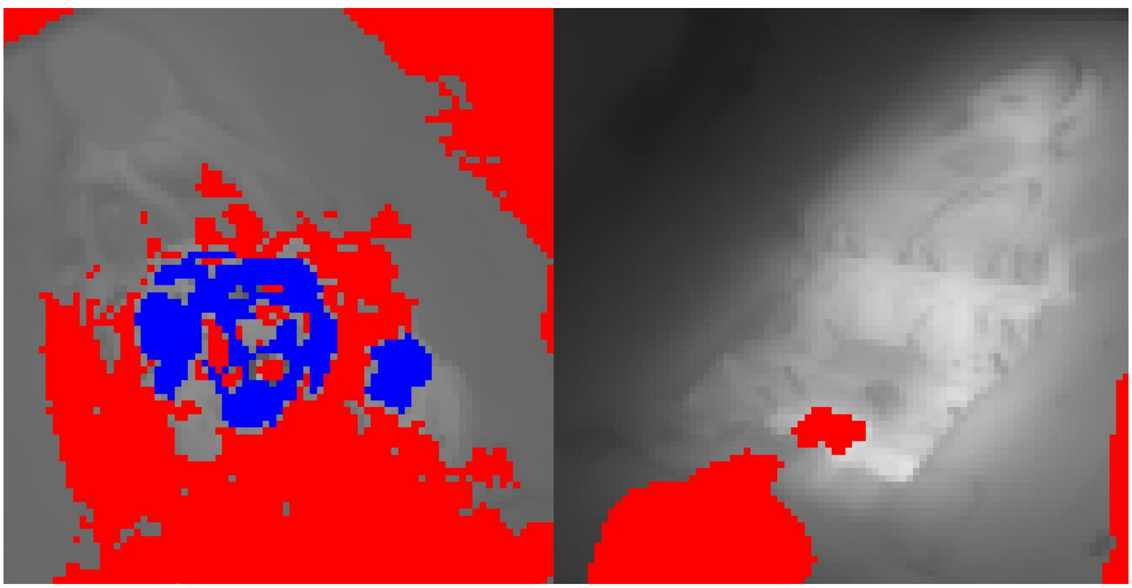


Figure 89: Anomaly detection on die with water turned off

It is important to note in Figure 89 is that emissivity becomes a critical component in understanding the thermal images. At certain temperatures, oxidation amounts of die surfaces, and polish levels, the infrared thermal camera may not be able to accurately read the temperature. NADCA publications have discussed emissivity in thermal images in more detail [223], [227]. In this case, the reflective areas near the extreme high-temperature zones in Figure 89 appear to be significantly colder (in blue) than typical. Although this is not a correct measurement by the camera system, this still highlights the power of the anomaly detection algorithm to identify changing temperature regions.

Benefits

The goal was to walk through this three-step anomaly detection algorithm to show how it can be utilized in an effective manner to gain valuable insight on complex, 2D, unsupervised data sets, such as thermal images in high pressure die casting. There is an initial commitment required for the image collection system and programming, but once it is in place, the processing time to review anomalies on new thermal image is almost instantaneous. Minimal effort is required to make modifications of the script for new parts (mainly to preprocess the image files with cropping and pixel scaling). This script becomes a powerful image processing tool that can allow a user to visualize and clearly communicate anomalies on image files.

One of the significant benefits of this approach is there is no requirement to have pre-labeled image files. Part traceability and data collection throughout the supply chain is time consuming. This approach provides valuable data regarding the process without the need for traceability of castings. Additionally, this algorithm requires domain knowledge to select the proper steady state cluster. Several types of analysis can be used to assist a user in selecting the steady state cluster, but it still relies on human input. The benefit is that the human is only looking at a significantly smaller portion of the data that the machine learning algorithm compiled. In this example, the human is required to understand the four unique clusters found via the clustering algorithm versus the original 894 images. This automation of data analysis while providing insight needs to be the goal of applying machining learning to manufacturing.

FUTURE APPLICATIONS

With this algorithm developed and proven on thermal images, two areas of future applications are currently under development. The first is the industrialization of this algorithm on the shop floor. The second is the use of this three-step anomaly detection algorithm on other images data sets, specifically x-ray images.

Development is underway with software providers to industrialize this algorithm on hardware located at the die casting cell. The goal is to make the algorithm "real time" to detect the anomalies identified within the die casting thermal process. Hardware will be implemented on the floor and interfaced with the thermal imaging system. Thermal images will be stored long-term on the network for future analytic modeling, as well as pushed to the hardware for real time analysis. The system will have programmed scripts to execute the image preprocessing and the SPC anomaly detection model calculations. The output image, with the blue and red anomaly highlights, will be displayed for the technician at the die cast machine. This image provides a visual tool to aid in troubleshooting any thermal issues within the die. Finally, a feedback loop from the hardware back to the die cast machine will provided an alert or stop the machine if a certain threshold of anomalies is detected. The idea is to prevent the machine from producing suspect product before making the casting. Plans are in place to have this system operational by mid-2019.

Along with industrializing the solution for thermal images, the three-step anomaly detection algorithm can also be utilized in other image analysis applications. Detection of changes in porosity on x-ray images was the next potential use investigated. Because x-rays are interpreted by humans, there is subjectivity on the interpretation of the images. The use of the three-step anomaly detection algorithm helps removes this subjectivity.

Initial experiments with the x-ray images yield some noteworthy results. First, it was quickly found that the x-ray fixturing used to locate parts was not repeatable enough. Image files were shifted significantly one part to the next. This prevented the immediate implementation of the algorithm. Image registration needed to be completed on x-ray images to align them prior to clustering. At the time of this publication, experimentation with the anomaly detection portion was still underway. However, several cases of creating a mean and standard deviation matrix and grayscale image were completed. This was done to understand the condition of the average part and where variation in porosity tended to form. A second application of the mean and standard deviation images was a comparison between two different process settings to understand if there was a visual improvement in the average porosity from the x-ray. Although neither of these cases utilized the full algorithm, they highlight the additional insight gained form image stack data. With further development and an improvement in image repeatability, the same colored anomaly detection algorithm can be applied to the x-ray images. Figure 90 shows an example of a mean and standard deviation grayscale image created using a registered image stack of x-ray images.

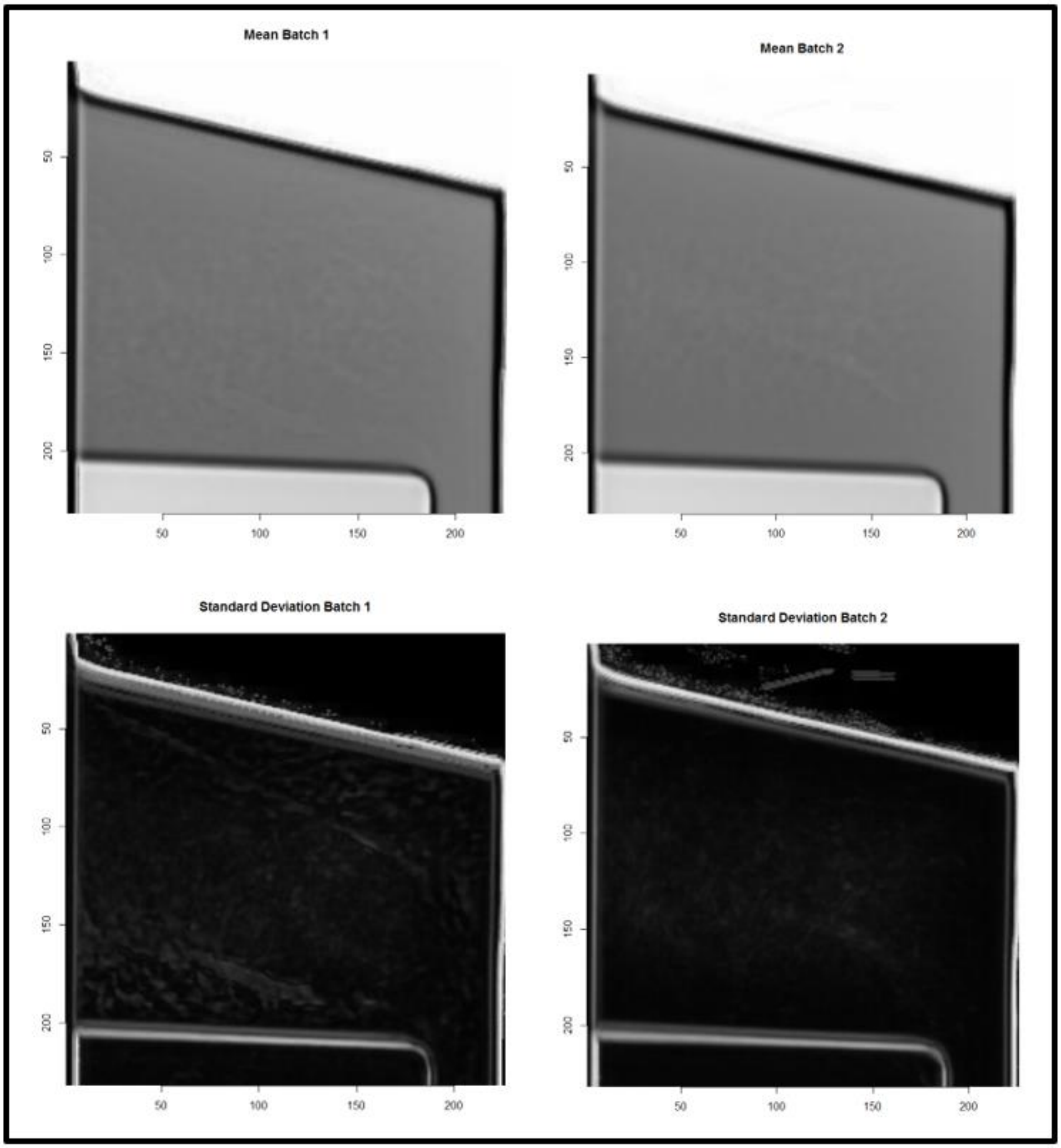


Figure 90: Mean and stand deviation images of two process settings

CONCLUSIONS

The three-step anomaly detection algorithm provides an extremely useful means of analyzing large volumes of images with relatively low human effort. Instead of manually reviewing and classifying thousands of images, a clustering algorithm creates a small number of groups to review. A subject matter expert can invest a small amount of time understanding and classifying these clusters. Once an optimal cluster is selected, the image stack is used to create mean and standard deviation matrices. These matrices become the SPC model for anomaly detection in any future image when compare to this optimal cluster. Areas beyond the selected threshold value can be highlighted and represent anomalies that are easily understood and visualized by humans.

As more data becomes available to manufacturing with the additions of sensors and cameras, it is becoming increasingly difficult to process this data in a timely matter to solve manufacturing problems. Machine learning must be put into practice within the industry to aid this data analysis. If anomalies can be detected and responded to in real time, the cost of poor quality will be significantly reduced. The goal of machine learning applications must be to automate the vast majority of data analysis. The algorithm proposed here helps achieve this goal for repeatable image data sets.

CASE STUDY III: TIME-SERIES ANALYSIS AND ANOMALY DETECTION OF HIGH-PRESSURE DIE CASTING SHOT PROFILES [5] [169]

For decades, the die cast industry has boiled down the thousands of data points collected in the time-series shot profiles to a handful of statistics such as average fast or slow shot velocities. These statistics are then trended and provide valuable process control information but averaging loses valuable data stored in the large set of time-series data. The approach of only using statistics from shot velocity and pressure profiles is no longer appropriate with the technology and computing power available today.

[5] This section is an edited version of an article published in Die Casting Engineer, included with permission from the North American Die Casting Association. D. Blondheim and S. Bhowmik, "Time-Series Analysis and Anomaly Detection of High-Pressure Die Casting Shot Profiles," Die Casting Engineer, p. 14-18, Nov. 2019.

A collaborative effort between Mercury Castings and Mercury Digital Services created an innovative approach to storing and analyzing shot profile data on Mercury Casting's BuhlerPrince die cast machines. The end solution was a real-time Azure App that can be accessed by employees via mobile devices on the foundry floor. This application provides both anomaly detection of a time-series profile as well as all the traditional statistics the foundry is used to reviewing.

This case study will review the benefits of analyzing time-series data, provide an overview for anomaly detection, and show an operational application solution.

CRITICALITY OF SHOT PROFILE

The importance of the shot profile or shot trace to produce quality castings has been a technical topic covered by many through the years. Slow shot velocity, fast shot velocity, fast shot deceleration, and intensification pressure are parameters people have researched and published on. Each impact the overall casting quality. These parameters are found in the velocity and pressure time-series profiles associated with high-pressure die casting.

In the 1990s, much work was done by Thome at Ohio State University showing the importance slow shot speed plays in cold chamber die casting. The research presented how a controlled, parabolic acceleration to a critical slow shot velocity reduced turbulence and minimized the air entrapment during the casting process [76]. Fast shot velocity is another critical parameter to the creation of high-quality castings. Fast shot velocity is calculated from flow rates established in PQ^2 calculation. As a result, fast shot velocity ties into the atomization and filling time of the die casting die [15]. In 2017, Miller and Monroe showed the benefits and limitations of varying fast-shot profiles to create prefill and reduce average cavity air pressure [90]. The importance of fast shot deceleration and the potential reduction of the impact pressure is discussed in both overflow design [228] and hydraulic control of the die cast shot system [87]. A rapid deceleration at the end of the shot can reduce the impact spike and prevent flashing of the die. This could lead to a reduction in downtime issues that improves the thermal cycling of the tool. Savage *et al* also discussed the impact pressure as well as the static and intensification pressures in a 2001 NADCA paper in producing quality castings [88].

CURRENT MEASUREMENT ISSUES

Decades of research in die casting show the injection velocity and pressure profiles are critical to the overall casting quality. Yet, when publications associated with process control during this same time are reviewed, a focus exists only on statistical parameters from this time-series data [19], [20], [23], [45], [89]. Often, the shot profile parameters generated by the data collection system ignores the acceleration, deceleration, and impact pressures, which have been shown to influence casting quality.

Most of the high volume of data collected during the injection process is reduced to a few averages. Technicians set points that act as ranges for average calculations in the shot profile collection system. An example of ranges set for slow, intermediate, and fast shot velocities are seen in the yellow circles of Figure 91. The need for consistent averages often drives the selection of these points. Setting a point too close to a transition would create averages viewed to be inaccurate or inconsistent. A small subset of the data available is used for a parameter average.

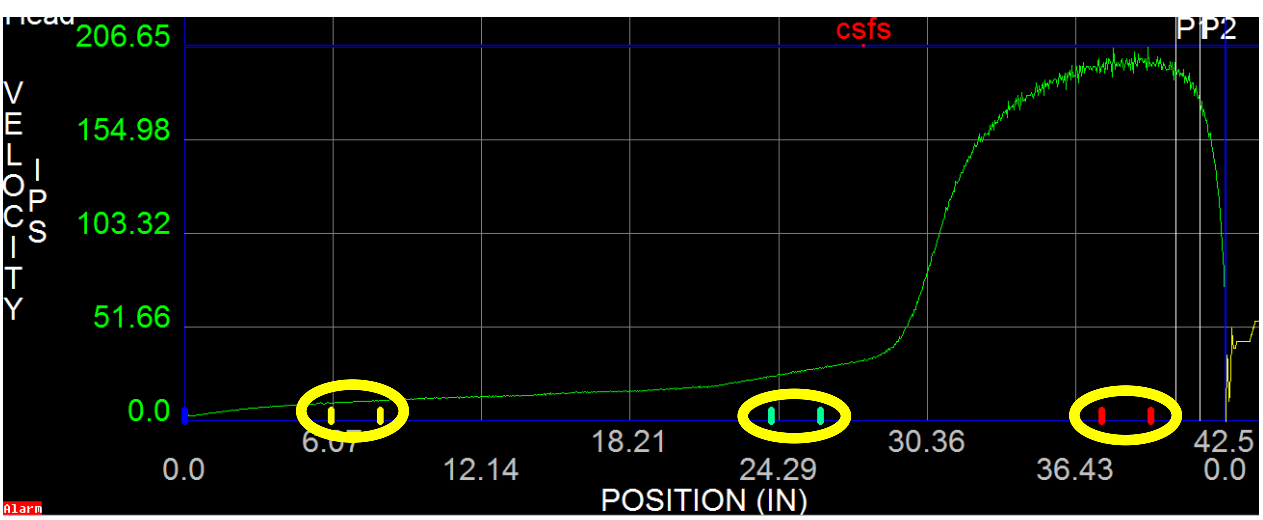


Figure 91: Typical range point settings for slow and fast shot velocities

Between these points, averages are calculated and shown as parameters used for process control. Lost in this average is anything associated with the change of rate associated with the profile. An average does not show if a planned constant acceleration in the slow shot is occurring. Information is lost unless one investigates the entire time-series string of data, often visually one shot at a time. Potential information loss is highlighted in the four different alternatives presented in Figure 92. Here, all four profiles have the same exact average of 152.5 inches per second (IPS) throughout the profile; however, the profiles are all significantly different: constant acceleration, constant deceleration, velocity step, and noisy or repetitive signal.

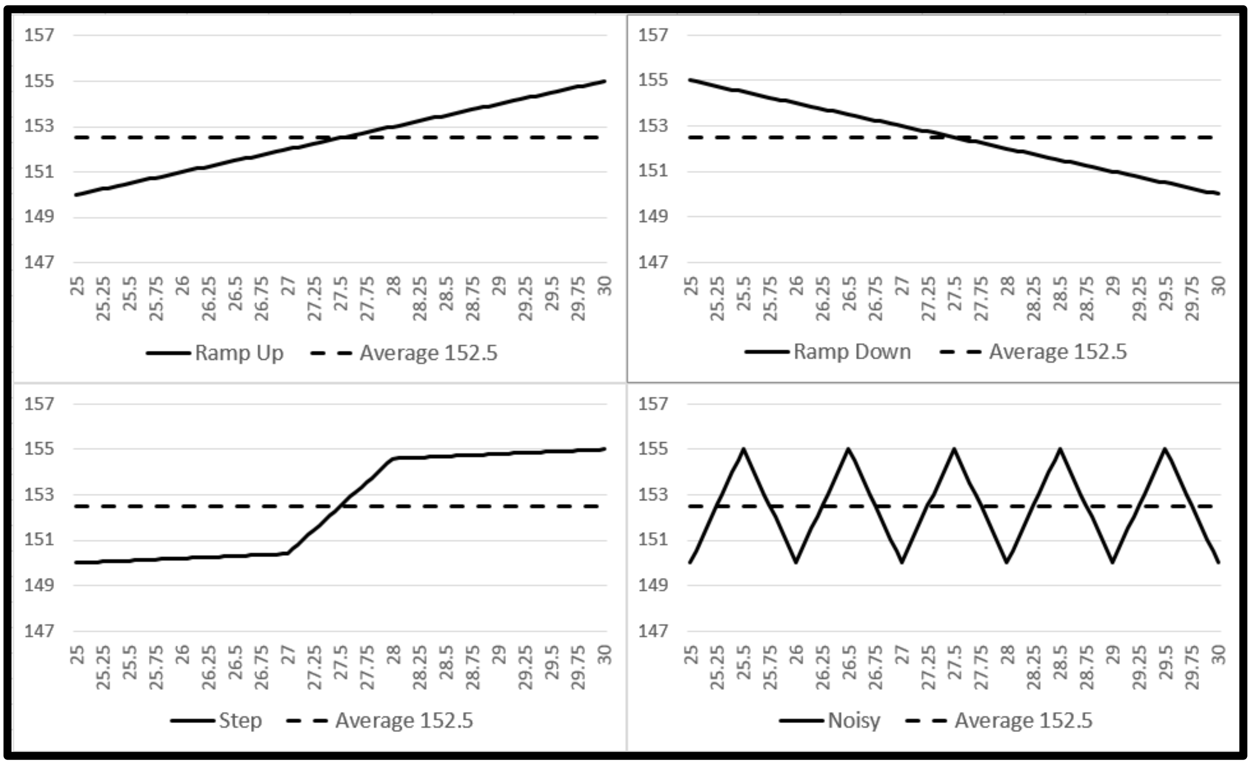


Figure 92: Four example profiles with the same average

Along with missing rate changes within the averages the shot monitoring systems provide, a typical commercially available shot monitoring system will also lack parameters for some of the critical process settings used. Impact pressure plays a significant role in casting quality as well as the ability to run the die. If a foundry is using hydraulic brakes to help control the impact pressure at the end of the shot, it would be critical to watch this. Visually, it is easy to see on a pressure profile. However, the data collection system lacks parameters around this impact spike to trend with the other parameters it collects. The parameter requires human review to ensure the process does not change over time.

The motivation within Mercury Castings to develop a new approach to review and analyze the entire time-series shot profiles led to the collaborative effort between Mercury Castings and Mercury Digital Services to store, process, and run this data analysis.

TIME-SERIES ANOMALY DETECTION METHOD

Anomaly detection methods are used to identify unusual patterns in data. Introduced here is a method to detect anomalies within time-series data including velocity and pressure profiles. Additionally, the process discussed can be used to find anomalies within other domains containing multidimensional data.

The velocity or pressure of a shot through time can be represented as a one-dimensional vector of real numbers. Consider a vector **A**, representing the shot velocity over *n* steps, such that $\mathbf{A} \in \mathbf{R}^{\mathbf{n}}$. For *k* castings, the mean velocity vector can be calculated. Per Equation 27, let **M** be the mean velocity vector such that:

Equation 27
$$\mathbf{M} = \frac{\mathbf{1}}{\boldsymbol{k} \sum_{i=1}^{n} \boldsymbol{A_i}}$$

Having found the mean velocity vector from a set of *k* velocity vectors, an algorithm can be used to find the anomalous velocity vectors. Each new velocity profile is compared to the mean velocity vector. Their similarity is calculated per the procedure below, and the least similar profiles are identified as anomalies.

Similarity Measure – Cosine Similarity

Cosine Similarity is commonly used as similarity measure for a vector of continuous numbers. The magnitude of shot profiles can vary significantly, but the overall pattern detected similarity is determined by the magnitude at each position of the vector and not the magnitude shift. Cosine Similarity is suited for this application as it is magnitude invariant but is sensitive to the overall pattern [229]. Equation 28 shows the standard Cosine Similarity calculation used for multidimensional vectors.

Equation 28 $$\boldsymbol{Cosine\ Similarity} = \cos(\boldsymbol{\theta}) = \frac{\boldsymbol{A \cdot M}}{\|\boldsymbol{A}\|_2 \|\boldsymbol{M}\|_2} = \frac{\sum_{i=1}^{n} A_i M_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} M_i^2}}$$

A set of sample velocity data was reviewed to determine fitness of the standard Cosine Similarity equation. The sample data was then passed through the anomaly detection algorithm, which is based off the Z-Score of the Cosine Similarity values. The Z-Score calculation can be seen in Equation 29.

Equation 29 $$\text{Z-Score} = (x - \mu) / \sigma$$

This anomaly detection procedure is explained in detail in the next section. Figure 93 shows most data identified as normal with a few profiles with patterns not consistent, marked as anomalies.

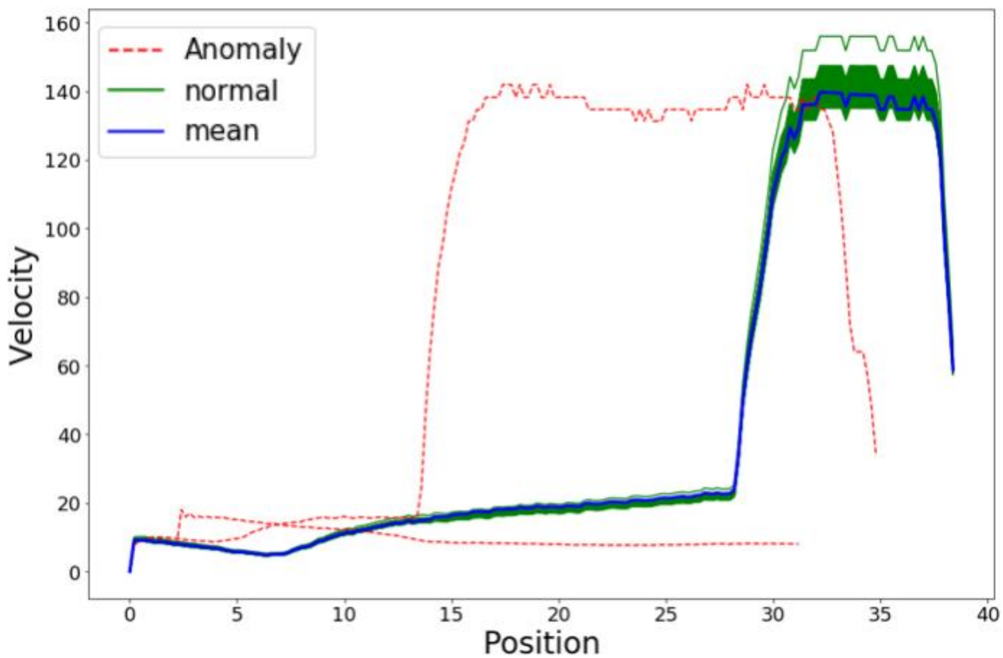


Figure 93: Anomaly detection using Standard Cosine Similarity

A velocity profile, which has the same pattern, but a difference in magnitude, was not identified as an anomaly. To help detect these types of anomalies, a modification was made to the standard Cosine Similarity formula. A penalty term was added to penalize large magnitude shifts. Equation 30 shows the updated formula with this penalty term. In the equation, β is a very small constant. Selection of β is dependent on the nature of the time-series data. For the data set reviewed here, it has been found $\beta < 0.0001$ works best.

Equation 30 $$Modified\ Cosine\ Similarity = \cos(\theta) = \frac{A \cdot M}{||A||_2 ||M||_2} - \beta||A - M||_2$$

Using the proposed modified Cosine similarity as the similarity measure, the same sample results were ran through the anomaly detection algorithm. Figure 94 shows the results for the modified equation, which now accurately detect the similar pattern, but magnitude shifted profile, as an anomaly.

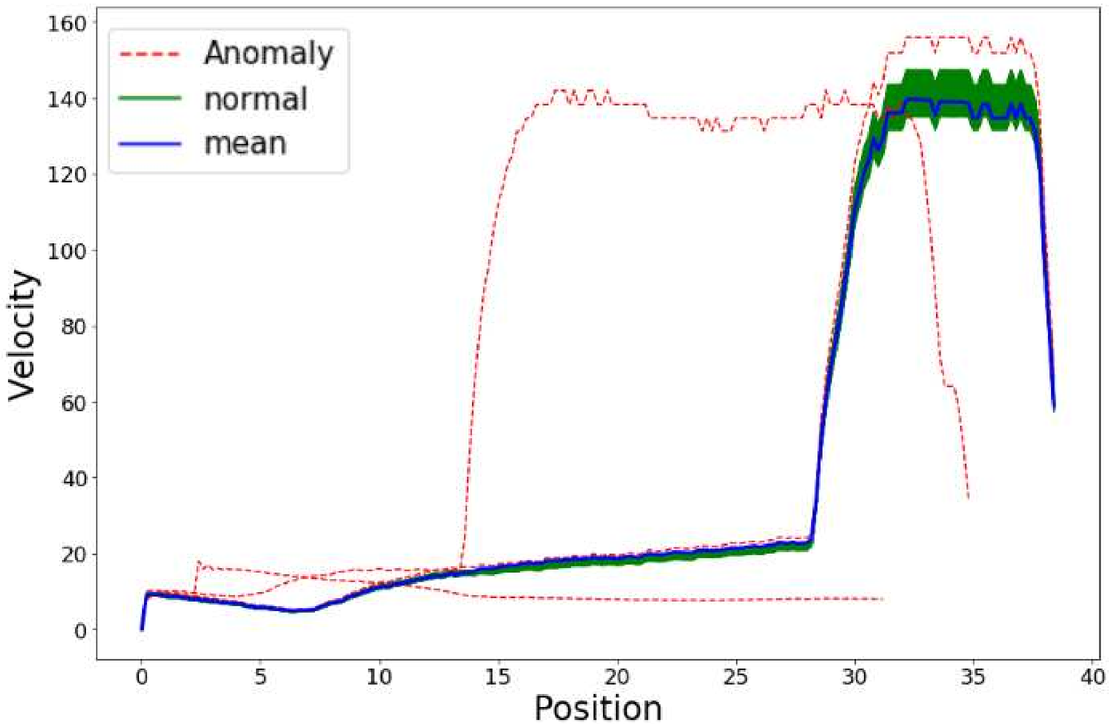


Figure 94: Anomaly detection using Modified Cosine Similarity

Unsupervised Anomaly Detection Method

Upon comparing each velocity vector **A** with the mean velocity vector **M**, a vector of Cosine Similarity scores is obtained. Since it is unknown up front which profiles are normal and which ones may be anomalies, an unsupervised approach is needed for anomaly detection. The first step is to calculate a Z-Score based on the Modified Similarity Cosine calculation with each vector **A** to the overall mean **M** of all vectors. The distribution of Z-Scores will likely be non-normal and skewed like the 0 to 1 cosine similarity values. This distribution should be analyzed to determine where appropriate Z-Score thresholds should be to represent a statistically significant difference in a profile. In a normal distribution, a Z-score of -3 to 3 would represent 99.7% of the population. With a skewed data set, this percentage does not directly apply. As a result, the Z-Score needs to be determined through analyzation of the calculated Cosine Similarities. Figure 95 shows a distribution of Cosine values used in this example problem.

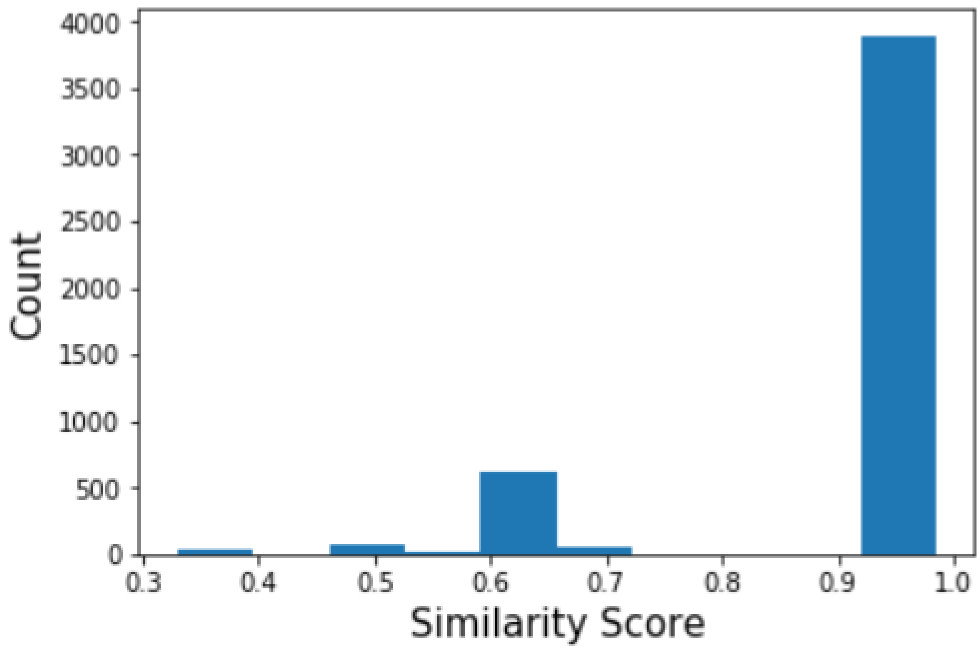


Figure 95: Distribution of Cosine Similarity calculation values

The analysis done here took a two-phase approach. Initially, it was determined that any Z-Score less than -5 would be removed from the list of cosine similarities and the corresponding shot profiles are identified as anomalies. With the new values, a second iteration of Z-Scores is calculated. Now any vector with a Z-Score less than -3 is defined as an anomaly. Only the negative Z-Score threshold is used, since the positive threshold would represent vectors that are almost identical to the mean vector. These positive anomalies can be ignored. If a significant number of shot profiles are identified with Z-Scores outside the identified threshold value, an investigation of the profiles should be completed to identify if any process shifts or changes may have happened during the data collection.

With **M** and Z-Scores established using unsupervised methods, it can now be compared to future shot profiles. If the Z-Score of the new cycle is below the negative threshold, the profile would be identified as an anomaly. This unsupervised method of anomaly detection works well in finding anomalous shot profiles and is computationally inexpensive allowing a web-based application to be easily built to run stored shot profile data.

APPLICATION IMPLEMENTED

An automated method to collect, store, and analyze time-series shot profiles was needed to successfully launch this advanced analytics method in a production environment. The initial focus of this project was relatively new BuhlerPrince die cast machines installed at Mercury Castings. Current work is underway to connect to other machine formats and create a single web-application that would function regardless of die cast manufacturer.

With this initial development, the shot profile data was already being stored into a network SQL database using BuhlerPrince's Pro-Supervisor system. The time-series data was stored in a compressed format, so the first step was to develop a Microsoft Azure Function Application that runs a script to decompress the time-series data and store it as a JSON file in Microsoft Azure Blob Storage. Azure Blob Storage is unstructured cloud storage, which is often used for video and image storage. The unstructured nature of blob storage differs from traditional structured storage of SQL databases. JSON files are a type of file format that uses text to store data often in array or time-series format.

With the time-series profiles decompressed and stored as JSON, Databricks service was used to build the anomaly detection model. This model is deployed on a Python microservice and callable via HTTP calls. A web application is written to display the data on a browser and make use of the anomaly detection microservice to detect anomalous shot profiles in real time. Also, the web application provides additional features such as multiple shot overlay, display of traditional shot parameters, and the use of defining a visual "master profile" that can be used as a visual comparison to each new shot profile. By using Azure Web Application and Office 365 login, any approved user can have access to the "Shot Profile App" on a device connected via Office 365. This approach provides true mobile connectivity to the die cast operations.

Figure 96 is the home screen of the application showing multiple machine profiles. Figure 96 shows an intensification pressure anomaly detection identified. The application compares the four different profiles with the anomaly detection algorithm and communicates which of the profiles are anomalous to the user.

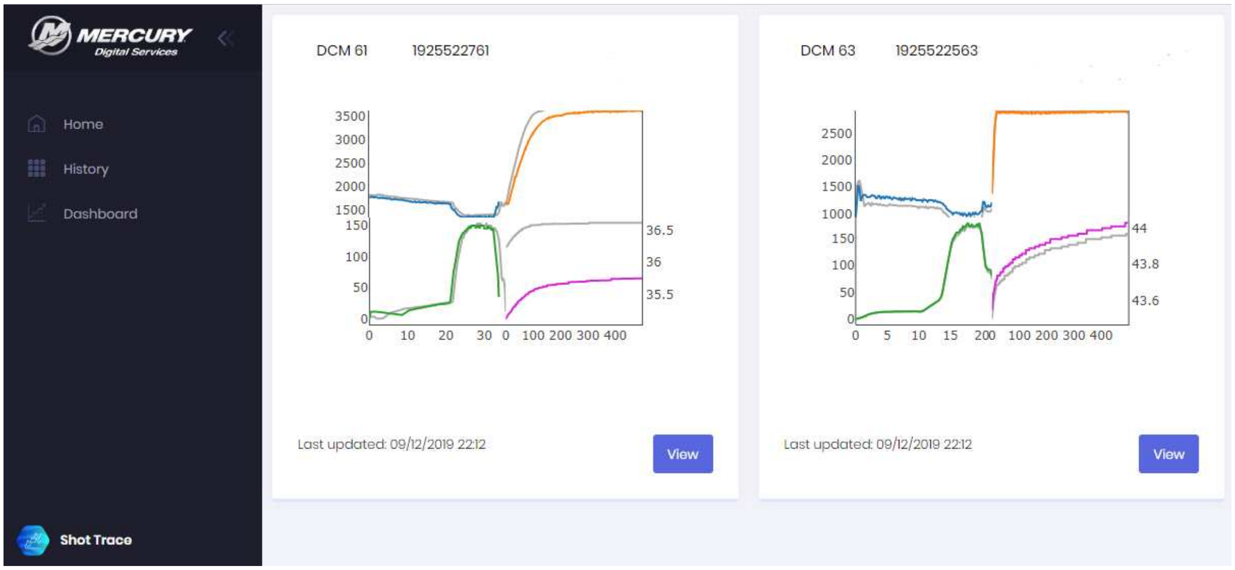


Figure 96: Home screen of Azure web page showing multiple machines running

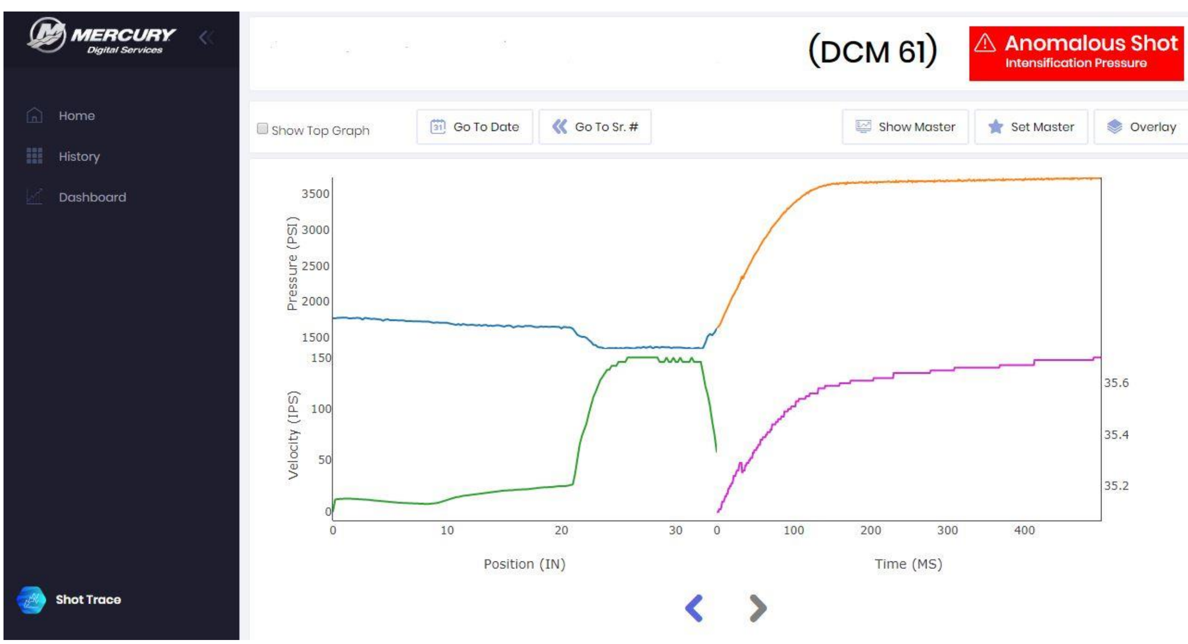


Figure 97: Detailed shot screen with anomalous shot detected

CONCLUSION

Time-series data collected during the die casting injection cycle provides significantly more data than traditionally used within the industry. Utilizing averages pulled from the profile does not provide the entire story of what is occurring with the shot. Accelerations, steps, and hydraulic brakes can all be missed with this approach. Published research shows that these can have a significant impact to overall casting quality, yet this is largely ignored in the data collection and processing systems that exist today.

The new approach presented utilizes the increase in data processing power and cloud applications to create user-friendly application that executes a sophisticated anomaly detection system on time-series data. By doing so, the entire shot profile is analyzed and compared to determine anomalies versus only averages calculated from the profiles. The application and algorithm provide many benefits to the foundry and will be further improved as additional research continues to determine the sensitivity of the anomaly detection algorithm in a production environment.

Technologies exist today to push the envelope on how the industry defines "Process Control." As data becomes easier to collect, store, and analyze, the industry must change its approach in how it analyzes data to generate quality improvements in the die casting process.

CASE STUDY IV: UTILIZING MACHINE LEARNING AUTOENCODERS TO DETECT ANOMALIES IN TIME-SERIES DATA[6] [179]

INTRODUCTION AND BACKGROUND

High-pressure die casting (HPDC) is a manufacturing process capable of producing high-speed, large-volume data sets [2]. The industry is familiar with and has history of collecting the velocity and pressure profiles during the injection process of the casting. Because the data is stored as a paired time and measurement, it is defined as a time-series data set. Initial shot monitoring systems from the 1960s

[6] This section is an edited version of a conference paper, included with permission from the North American Die Casting Association. D. Blondheim, "Utilizing Machine Learning Autoencoders to Detect Anomalies in Time-Series Data" 2021 NADCA Congress and Exposition. Indianapolis, IN, Oct. 2021

included using transducers, signal conditioning equipment, and a physical recording device using pen on graphing paper collecting date at half-second intervals [78]–[80]. Improvements to data collection came in the 1980s, and 1990s, as the adoption of PCs and PLCs digitized the data collection process [81], [82]. By the early 2000s, most North American die casters utilized shot monitoring systems [230]. The access of the data then brought the use of advanced analytics. Publications on the use of machine learning on shot profile data became more evident in the late 2000s, 2010s, 2020s [21], [22], [45], [83]–[85]. Additionally, the acceptance of Industry 4.0 or Smart Manufacturing has driven the availability of sensors and data collection in all manufacturing [231]. The technology of data collection and monitoring has significantly improved, creating a great opportunity for advanced analytical tools.

Although the shot profile is a critical aspect of the die casting process, it is not the only controlling factor of how well the equipment runs and the quality of parts produced. A comprehensive review from a manufacturing system perspective generates many other time-series data streams that can be monitored and used to improve process control in die casting [2]. Table 35 contains a list of processes or equipment where additional time-series data can be collected. By monitoring these processes in a similar fashion as the shot monitoring systems, additional insight, control, and process understanding can be achieved within the industry.

Table 35: Additional time-series data in HPDC

Process/Equipment	Data to Be Collected	Potential Benefits
Die Cast Machine Motor	Amp Draw, Vibration, Temperature	Downtime Prevention, Equipment Maintenance
Hydraulic Pump	Vibration, Temperature, Pressure	Downtime Prevention, Equipment Maintenance
Nitrogen	Pressure	Recharge Rates, Tank Leaks
Closing Cylinder during Open and Close	Pressure	Equipment Maintenance, Prevention of Tool Damage
Thermal Management: Hot Water Units, Hot Oil Units, Jet Cooling Units, Cold Water	Temperature, Flow Rates, Pressure	Part Quality, Thermal Management, Equipment Maintenance, Process Consistency
Ejection Cylinder	Pressure	Part Quality/Part Warp, Downtime Prevention/Stuck Part
Slide Pulls	Pressure	Part Quality/Part Warp, Downtime Prevention/Stuck part
Die Spray System: Lube and Air	Pressure, Flow Rate	Part Quality, Lube Reduction, Process Consistency
Furnace	Temperature, Amp Draw, Fill Level	Part Quality, Downtime Prevention, Equipment Maintenance
Sawing Equipment	Amp Draw, RPM, Feed Rate	Downtime Prevention, Equipment Maintenance
Trim Press Equipment	Hydraulic Pressure	Downtime Prevention, Equipment Maintenance

After data is collected, some level of analysis must be completed to create the value in collecting the data [29]. Two approaches for analyzing time-series data [23] are:

- Trending of descriptive statistics via statistical process control (SPC)
- Visual comparison to another profile, completed by the operator

Both these approaches have shortcomings.

Average fast shot velocity is an example of a descriptive statistic that is familiar to the industry. Average fast shot velocity is an average of the time-series data calculated between two set points selected by the user. These points are selected in a shot monitoring system by specifying locations within the chamber to start and stop the averaging. An example of the set points for average slow shot velocity (red points) and average fast shot velocity (blue points) can be seen in Figure 98.

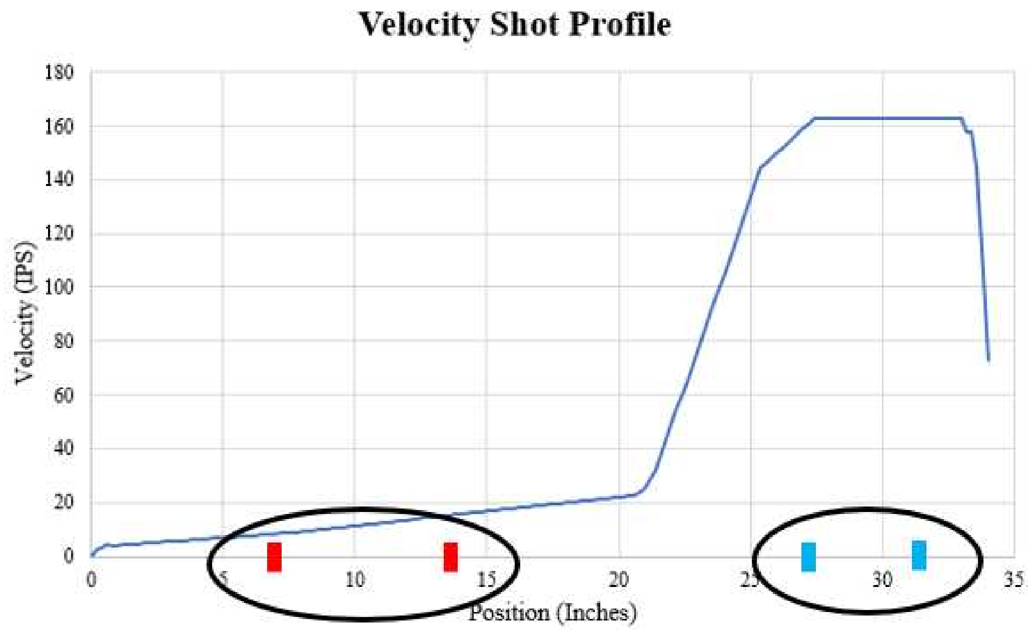
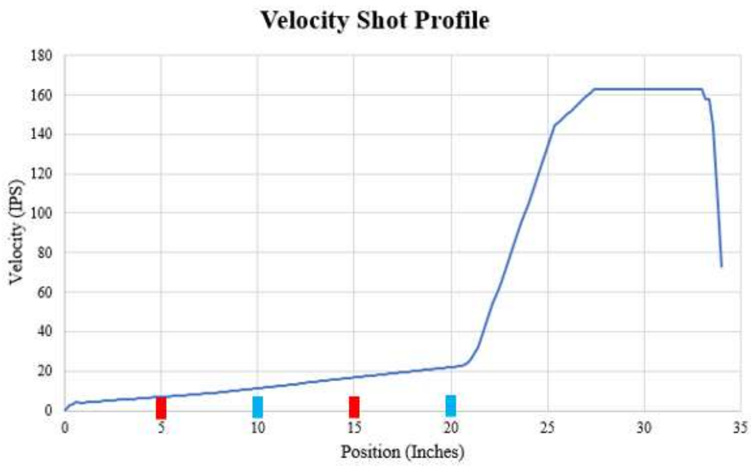


Figure 98: Set points for shot profiles

Based on the set points, an average of the time-series data is calculated and used to describe that larger data set. This statistic creates a simplified view of the complicated data. It is easy to trend over time with the use of statistical process control (SPC) tools [23], [226]. Unfortunately, this statistic also represents a significant loss of data. Additional data should be collected to provide it context, such as the set points to see if they change over time as well. A shift of the velocity average may not be caused by a process change, but instead it could represent the set points being modified by a technician. The knowledge of this change is lost without the additional details being stored and analyzed along with the average. An example of this issue can be seen in Figure 99. This figure contains a theoretical ramped slow shot velocity profile with different set points, which result in different averages. These differences are especially evident in ramped portions of the shot. Ramped velocity is often used during the slow shot to help reduce defects [76]. Similar average differences can also occur if the flag is near a transition point, such as the end of the fast shot.



Range	Average Slow Shot Velocity (IPS)
5 to 15 (Red Flags)	11.53
10 to 20 (Blue Flags)	16.76
5 to 20	14.17

Figure 99: Changing setpoints on a ramped slow shot velocity profile

Additionally, by using a descriptive statistic, the data values collected between the set points are not fully utilized. An example of this can be seen in Figure 100 where four significantly different velocity profiles all produce the same statistical average between the set points of 5 and 15 inches. Relying only on descriptive statistics can lead to incorrect assumptions about the process.

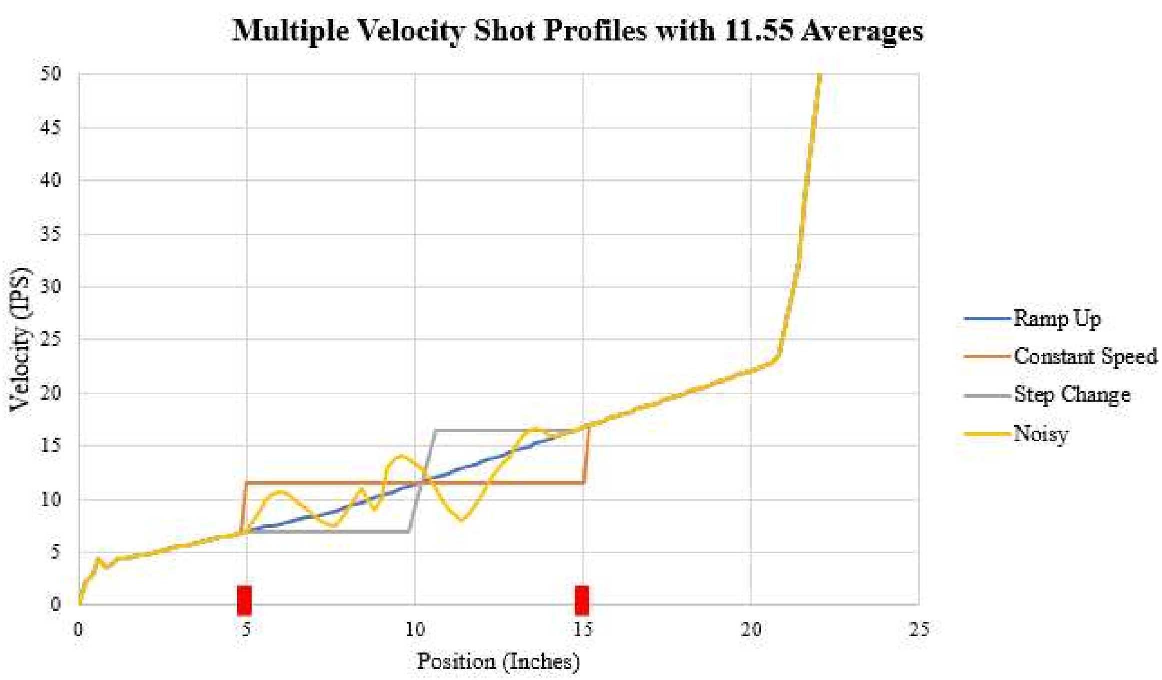


Figure 100: Examples of identical slow shot velocity averages with different profiles

One way to overcome these incorrect statistical assumptions is for an operator or technician to complete a visual review of each time-series profile. This is done regularly within the industry when developing and troubleshooting a process. It is one of the reasons why shot monitoring systems are popular. Differences can be visually determined and can trigger investigation and troubleshooting into the process or equipment. Additionally, there are aspects of the process like low impact at the end of the shot or tip wear showing up in the slow shot that can be missed in the descriptive statistics, but easily recognized by a well-trained technician visually inspecting the profile.

The issue with this visual inspection is two-fold. First, humans tend to perform poorly on visual inspection tasks that are repetitive and mundane. Studies have shown that humans miss 20 to 50% of the manufacturing defects during visual inspections they perform [120], [142]–[146]. When an operator is required to monitor every shot profile, in addition to the job requirements associated with HPDC, this miss rate will likely worsen. The second issue faced with human inspection is the skill level of the operators themselves. The shortage of skilled and knowledgeable HPDC operators exasperates the situation [232]. In addition to training on equipment, operators will need to spend time training on process control, monitoring, and inspection of velocity and pressure profiles. This is far from an ideal situation, given the current shortage of workers and the amount of training a new worker needs [233].

The HPDC industry is at a crossroads from the data perspective. Traditional data collection systems like shot monitoring exist and are accepted by the industry, but they only provide meaningful value for a portion of the overall HPDC process. These systems rely on descriptive statistics or visual operator inspection and have the potential to miss changes in the process that can affect part quality or equipment failure. The growth of Industry 4.0 or Smart Manufacturing provides a catalyst for foundries to collect and store additional time-series data without commercialized systems. Once the data is collected, it can quickly feel like an insurmountable task to analyze, given its volume and velocity. The systems that have been used to date will not be sufficient for the data that will be generated moving forward. The shot profile, along with the flow rates in spray, the slide pull pressures, the ejection force, the vibration in the motors and pumps, and many other data streams must also be analyzed. Foundries will start to collect these additional sources to gain better process knowledge but will need tools to analyze the volume and velocity of the data. Machine learning approaches are needed to analyze the volume of time-series data generated within HPDC.

One method for analyzing time-series data in HPDC is the use of a modified cosine similarity function to determine anomalies [169]. The traditional cosine similarity function uses the dot product of two vectors and the vectors' magnitudes to calculate the degree the vectors overlap or are similar. The drawback of the traditional cosine similarity function is that although it can detect differences based on vector orientation, it has limited ability to find anomalies based on its magnitude [234]. This was addressed by modifying the cosine similarity function with a penalty term for the magnitude differences [169]. The penalty term included a tuning parameter, beta (β), the user would have to set based on the data set. The modified cosine similarity function comparing new vector A with a mean vector M is seen in Equation 31 [169].

Equation 31 $$\textit{Modified Cosine Similarity} = \cos(\theta) = \frac{A \cdot M}{||A||_2 ||M||_2} - \beta ||A - M||_2$$

Because the beta value in the modified cosine similarity function requires a degree of interaction with the processing of the data, other approaches should also be considered for anomaly detection. Machine learning neural networks have been used for time-series data analysis [140], [175], [235]. Typically, neural networks are used to make predictions of results (supervised machine learning) [175]. Autoencoders are a special case of neural networks that do not require results to be trained (unsupervised machine learning). This work will explore the use of autoencoders for anomaly detection in time-series data sets generated in HPDC.

Since the use of machine learning is novel to most foundries, a theoretical example using hand-written digits will first be reviewed to explain the process of building an autoencoder. This hand-written digit example provides images to allow someone unfamiliar an easier means of visualizing the process. After the theoretical example, a case study on time-series flow rates of the die casting spray cycle will be reviewed. This will show how two weeks of data can be used to build and test an anomaly detection autoencoder.

In the end, the goal of this work is to eliminate the notion that implementing machine learning within a foundry is an insurmountable task. The process to train an autoencoder for time-series data is straightforward. Although purchased software packages exist, all the work completed here was done using open-source programming languages and packages. Applying machine learning analytics is an attainable objective for foundries. This work will provide a framework for utilizing autoencoders to detect anomalies in time-series data.

AUTOENCODER

Background

An autoencoder is an unsupervised machine learning technique that can be utilized to detect anomalies within time-series data [175], [235]. Machine learning (ML) can be defined as a subset of Artificial Intelligence (AI) in which algorithms are used to learn patterns within the data without explicitly being programmed [140], [236]. Two common types of ML are supervised ML and unsupervised ML. In supervised ML, both the input variables and the results are known. It is the goal of the supervised ML to learn the patterns between the input variables and the results in the training data to accurately predict the result or classification on future data sets. In unsupervised ML, there are no results available. Training only occurs on the input variables. Unsupervised ML focuses on clustering and anomaly detection within data sets [140]. An autoencoder is one of several types of unsupervised ML algorithms and utilizes the power of a neural network to develop models for complex, non-linear data [235].

A neural network is a type of ML algorithm that is based on the connection of neurons within a brain. Neural networks are made up of interconnected neurons that attempt to learn and detect patterns in the data that is presented [235]. A simple neural network with one input layer containing four variables, one hidden layer of three nodes, and two-class output layer is shown in Figure 101.

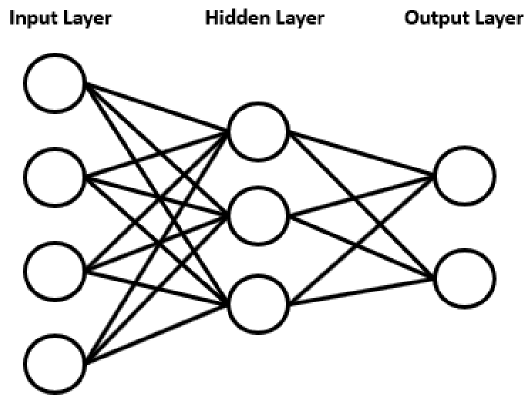


Figure 101: Simple neural network

As computing power has increased, the depth and complexity of neural networks has followed. Deep learning (DL) has become widely used in ML for image classification [237] and autonomous driving [198]. DL is a neural network which contains multiple hidden layers, thereby making it deep when compared to a traditional neural network [175]. Figure 102 shows an example of a DL neural network.

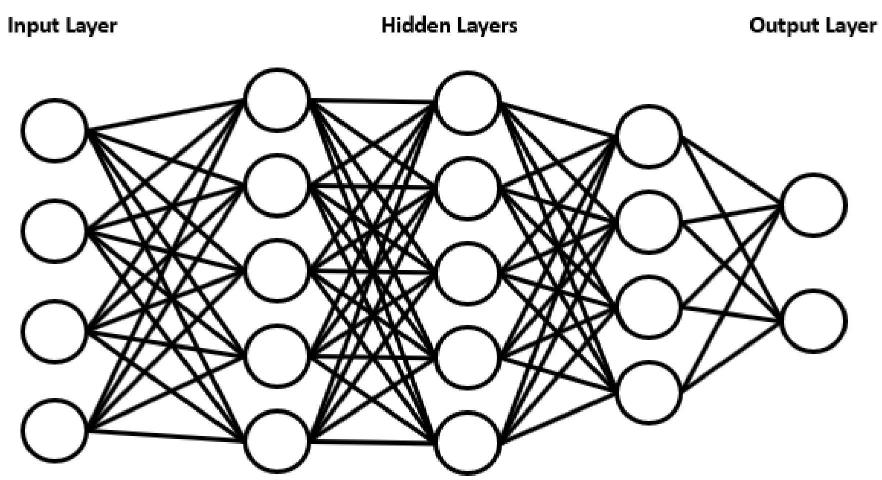


Figure 102: Deep learning example

An autoencoder is a specialized type of neural network with a specific function. The goal of an autoencoder is to learn the pattern within a given data set that allows the algorithm to recreate a given instance of data through an encoding and decoding process. An autoencoder becomes a compression algorithm. Interestingly, an autoencoder is purposefully designed not to be perfect at reconstruction. If it did learn to be perfect, the algorithm would be able to copy data, but not learn. Instead, an autoencoder is forced to learn the key patterns of the input data that will generalize a reconstructed output [175]. Learning these patterns is how an autoencoder becomes useful for anomaly detection. Autoencoders are unsupervised since there are no results associated with the training. The goal is simply to recreate the output data to minimize the error with the true input data.

An autoencoder is comprised of five key components: input variables, compression layers (encoder), code layer, decompression layers (decoder), and recreated output layer [175], [238]. These components can be seen in Figure 103. Multiple layers within the encoder and decoder can exist, giving an autoencoder the multiple hidden layers like deep learning. There will only be one input layer, output layer, and code layer. An autoencoder will often have a mirrored number of layers and nodes associated with the encoding and decoding process, but this is not required [175].

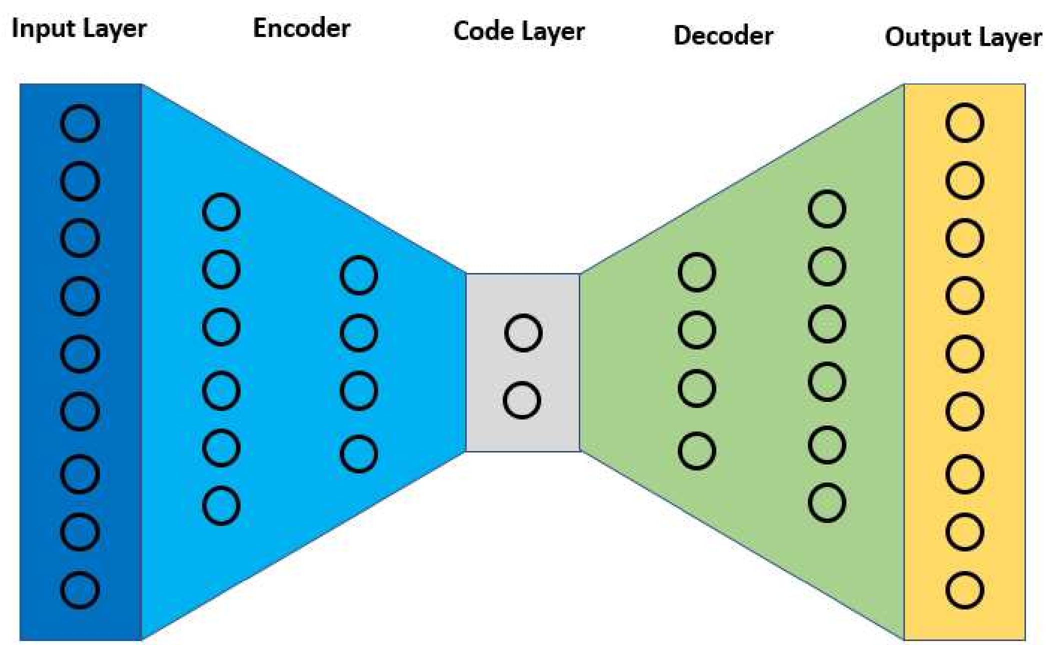


Figure 103: Autoencoder components

An autoencoder learns and adjusts the weights within its neurons based on the error produced between the input data set and the output data set. The algorithm's goal is to minimize the error or difference between the input and output through the entire batch of training examples it is provided. Entire textbooks are dedicated to the mathematics and programming behind neural networks [175], [235].

These references should be reviewed if there are additional questions regarding terminology used in this work including topics like activation functions, epochs, batch-sizes, optimizers, and loss functions.

Application and Theoretical Example

Creating, training, and implementing an autoencoder can be completed through many different software packages. Some softwares like MatLab, Wolfram Mathematic, and Neural Designer require licenses to be paid to utilize the software. Many other packages, such as Keras, PyTorch, and TensorFlow are open-source software requiring no financial investment. Likely the largest hurdle for most HPDC foundries to implement this technology is the programming skillset needed to complete the code. Python, C++, and R are programming languages commonly used by these packages.

It is highly probable that someone new to the ML field will become confused with how these packages work and interact. There is no one right approach to take to implementing this type of technology within manufacturing. Success can be reached through many different programming languages and software packages. A walk through of the software used for this paper helps highlight the complexity and interconnectedness of this software for those not familiar. This work utilizes the R programming language and the Keras package for R. Keras provides a neural network library through a Python interface while using the TensorFlow library. The benefit of Keras is that it provides a user-friendly programming interface to train and run a neural net. The approach used here should not be considered superior or inferior to other approaches. This path was selected because of the author's familiarity with the R programming language and the desire to easily implement an autoencoder using modern deep learning technology. A different path through other programming languages and software packages will follow a similar process and should achieve comparable results.

To explain at a high level, the programming associated with an autoencoder can be broken down into four key steps. These steps include:

- **Step 1 – Prepare Data:** The data must be prepared to be accepted by the machine learning packages
- **Step 2 - Define the Autoencoder:** Parameters of the autoencoder including number of inputs, number of layers, size of layers, and activation functions must be determined.
- **Step 3 - Train the Autoencoder:** Training data is passed into the algorithm with the required loss function and optimization method.
- **Step 4 - Test the Autoencoder:** Test data, which the algorithm was not trained on, is used to compare the performance of the trained model.

A theoretical example utilizing the MNIST handwritten data set [239] will be reviewed ahead of a HPDC case study to show the four steps. The MNIST handwritten data set is commonly used in the ML algorithm testing and development. This data set is comprised of both training and test data. The training data is 60,000 handwritten digits between 0 and 9. The testing data set is 10,000 additional digits, not included in the training data set. All digits are stored as a 28 by 28 pixel image or matrix, for a total of 784 data points per digit. Images are stored with a 0 to 255 value in a format typical for grayscale images [239]. The labels of the images are not needed with an autoencoder as the goal is to simply replicate the original input image. Figure 104 shows examples of random MNIST handwritten digits used in training.

Figure 104: Example MNIST handwritten digits

Step 1: Prepare Data

The initial step in creating an auto-encoder is preparing the data. For this MNIST example, the data needed to be both scaled and vectorized as described below. Some of these steps may not be needed for time-series sensor data. The format of the data will depend on the software packages used, so this is just an example.

As stated earlier, the images are stored with values from 0 to 255. It is easier to work with raster image files in R when the values are scaled from 0 to 1, instead of 0 to 255. This is easily done by dividing the training and testing data by 255. Now 0 represents a pure black pixel and 1 represents a pure white pixel. With this conversion, the process to plot the image is a one line code as seen in Figure 105.

Figure 105: Plot raster example

In addition to the scaling, the MNIST data is a three-dimensional array. The two dimensions of the image are straightforward to understand. The third dimension is the stack of 60,000 training and 10,000 testing images that exist. For this simple example in Keras, the individual example data needs to be vectorized into a one-dimensional data stream. Then they need to be combined as multiple rows to be consumed during the training of the algorithm. This is the data preparation work that is specific to different languages and packages.

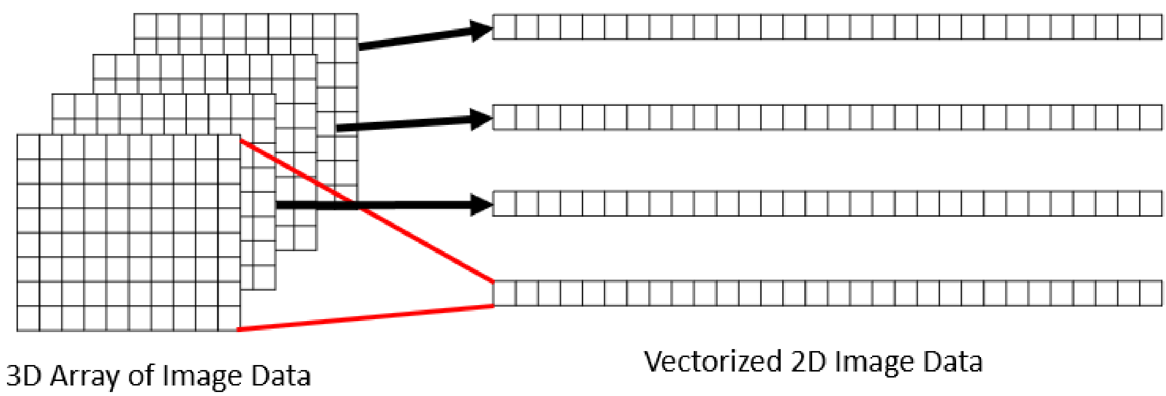


Figure 106: Data preparation

For time-series data, this type of data preparation may not be needed, yet other steps may be required, such as vectorizing a compressed field or transferring the data from one file type into the required shape. The exact level and type of data preparation varies by languages and packages. However, it is not overly complicated, nor should it discourage a foundry from to analyze data with ML.

Step 2: Define the Autoencoder

With the data prepared for the package, the next step is to define the structure and activation functions of the autoencoder. For those unfamiliar with autoencoders and neural networks, this step could seem overwhelming. Decisions made at this point will drive the length of training and the overall accuracy of the model prediction. The benefit is that this is done electronically, so even if a poor decision is made, the testing of the algorithm will identify poor performance.

To provide several examples, a range of structures for the autoencoder was selected to show the impact on training time and test results. These ranges included different number of layers, the neurons per layer, and neurons in the code layer. The Leaky ReLu activation function was used for all hidden layers, and the Sigmoid function was used at the output layer to recreate a value between 0 and 1. The input and output layers are set based on the input size of 28 by 28 pixels or 784 variables. Table 36 shows the ranges used within this theoretical example.

Table 36: Ranges of structure used for autoencoder example

Example Number	Example 1	Example 2	Example 3	Example 4	Example 5
Input Layer Size	784	784	784	784	784
Output Layer Size	784	784	784	784	784
Structure of Encoder	1st layer = 300	1st layer = 400 2nd layer = 100	1st layer = 500 2nd layer = 300 3rd layer = 100	1st layer = 500 2nd layer = 300 3rd layer = 100	1st layer = 600 2nd layer = 500 3rd layer = 400 4th layer = 300 5th layer = 200
Code Layer	10	20	20	50	100
Structure of Decoder	1st layer = 300	1st layer = 100 2nd layer = 400	1st layer = 100 2nd layer = 300 3rd layer = 500	1st layer = 100 2nd layer = 300 3rd layer = 500	1st layer = 200 2nd layer = 300 3rd layer = 400 4th layer = 500 5th layer = 600
Hidden Layer Activation Functions	Leaky ReLu	Leaky ReLu	Leaky ReLu	Leaky ReLu	Leaky ReLu
Output Activation Function	Sigmoid	Sigmoid	Sigmoid	Sigmoid	Sigmoid

Step 3: Train the Autoencoder

Once the autoencoder is defined, the next step is to complete the training for each example. Because this is an unsupervised autoencoder, the input data will also be the target data used to measure. In other words, the neural network is being trained to compress and decompress the input value to minimize the difference between the predicted final output. To do this, again several decisions must be made by the user. For this example, a batch size was set to 300 samples. After some testing, 50 epochs were used in the training for all samples. The RMSPROP optimizer was used along with a loss minimization function of Binary Cross Entropy. The training was completed on the CPU of a laptop computer with 32 GB of memory and an Intel Core i7-6820HQ 4 core at 2.70 GHz CPU. The training time it took for each example is seen in Table 37.

Table 37: Training times for each example

Example Number	Example 1	Example 2	Example 3	Example 4	Example 5
Training time (minutes)	2.28	3.36	4.89	5.16	9.58

Step 4: Test the Autoencoder

The testing of the model is the last step in building an autoencoder. In the testing, a data set that the model was not trained on will be processed by the algorithm to show how well it can do at predicting. An autoencoder is different than traditional supervised ML testing, as instead of a predicted value, the predicted output is compared to the actual input. The algorithm tries to minimize the difference between the two. The unsupervised ML autoencoder cannot use traditional accuracy or quality metrics of supervised ML, such as Confusion Matrix or ROC graphs. As such, a graphical display of random digits selected will be shown along with a mean absolute error calculation between the input image and the output image.

The visual display of the testing, as seen in Figure 107, highlight that an autoencoder loses information when it goes through the encoding and decoding process. The algorithm is not designed to be lossless compression. In Example 1, several of the digits have blacked out areas where information was lost between the input and output. Visually, the autoencoder does a good job of replicating the input to the output in all the examples.

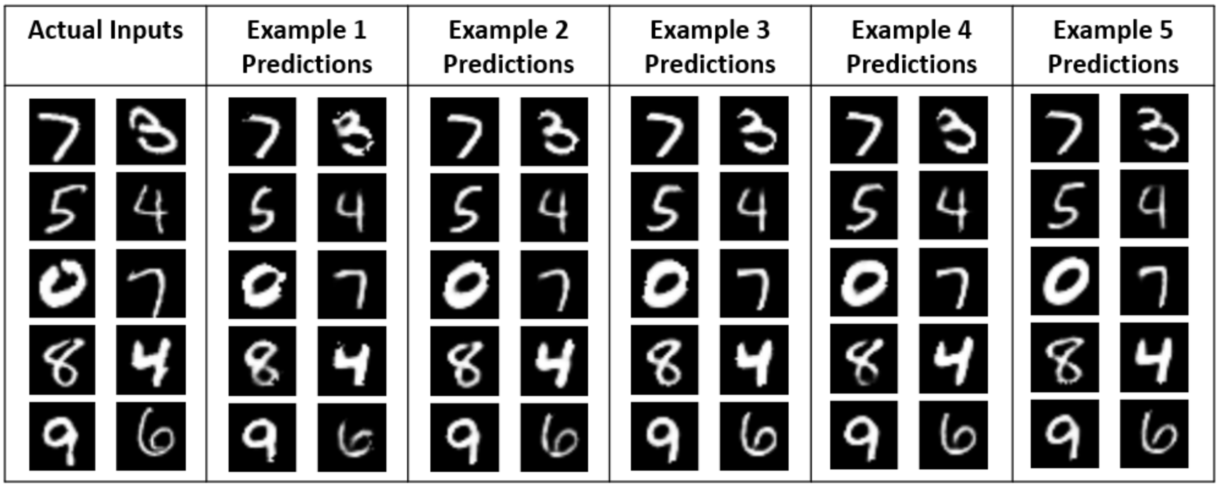


Figure 107: Digit predictions per example

For a calculated error value, the mean absolute error (MAE) was calculated based on Equation 32. Averages of those error rates by digit are seen in Table 38. Based on these averages, digits exhibit different levels of error. Typically, the digit 1 performed the best, while the digit 5 performed the poorest. The shared features between digits likely degrade the overall performance of the autoencoder. For example, the physical differences between the digit sets 3/8 and 5/6 are limited, which makes it harder for the ML. These values likely could be further improved by additional improvements of the autoencoder structure and testing of different loss functions and optimizer used. However, the goal of this theoretical example is not to make a perfect autoencoder for handwritten digits, but rather to provide the foundation of how this technology can be used to help detect anomalies in complicated time-series data.

Equation 32 $$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

Table 38: Average mean absolute error rate for examples

Average Mean Absolute Error by Digit and Overall					
Digit	**Example 1**	**Example 2**	**Example 3**	**Example 4**	**Example 5**
0	0.243	0.153	0.172	0.157	0.172
1	0.125	0.091	0.093	0.095	0.104
2	0.288	0.191	0.210	0.195	0.220
3	0.262	0.171	0.189	0.185	0.211
4	0.218	0.164	0.185	0.167	0.198
5	0.271	0.190	0.210	0.198	0.234
6	0.226	0.163	0.175	0.165	0.189
7	0.201	0.154	0.174	0.166	0.192
8	0.262	0.176	0.204	0.187	0.226
9	0.191	0.143	0.164	0.152	0.183
Total Average Error	**0.227**	**0.158**	**0.176**	**0.166**	**0.191**

Steps 2 through 4 is an iterative process. If the results of the testing are not ideal, changes in the autoencoder design may be required. This would result in new training and then new testing. Developing ML algorithms for implementation often requires much effort in this iterative space. Again, this effort should not discourage foundries from testing and implementing machine learning technologies, as there is significant value to be gained.

Anomaly Detection

A shortcoming for an autoencoders as an anomaly detection device stems from the training process. As seen in the MNIST example, the autoencoder was trained on many different digits. It could learn the pattern for all the digits in the neural network, so it could redraw the input with minimal error regardless of the digit. For an autoencoder to detect anomalies in a process, it would have to be trained on the acceptable or standard process that is expected. By training just on what is expected, the neurons will learn the pattern associated with what an acceptable input will look like and be able to recreate the output with minimal error. When an anomaly is introduced into the autoencoder, the algorithm will struggle with recreating the input, and the error rate will be higher. Based on the training and testing data, an error threshold can be selected to trigger anomaly alarms.

The tactical problem with this process is selecting a data set with only known acceptable data. In the MNIST example, the data was already labeled with the digits, which meant a person had to look at and label all 60,000 training and 10,000 test images. This is time consuming and prone to human error. It is counter-intuitive of the goals of machine learning implementation. Foundries would not want to gather 10,000 pressure profiles of the spray system, to then go through each one individually and identify if it is acceptable or normal for training or an anomaly that should not be trained.

To help solve this problem and reduce the timeframe needed to review the data, another unsupervised ML algorithm can be used. K-Means clustering algorithm can be ran to group the rows of data into similar clusters. These clusters can then be reviewed to determine which cluster best represents the standard or normal process. The autoencoder would then be trained only on the given rows of data that represent that cluster, thus making it capable of learning that process. When an anomalous data string is entered into the autoencoder, it would have a higher error rate between the predicted and input data, thereby triggering an anomaly.

This clustering and training will likely be iterative. The first clustering may not separate out all the anomalous data. As a result, after the auto-encoder is trained, the error rates on the data used for the training should be reviewed. Any error rates that are significantly higher than the rest should be reviewed and possibly eliminated. These likely represent anomalies. If the decision is made after the review to eliminate the data because it was an anomaly, then the autoencoder would need to be retrained on the cleaned data set.

To finalize the theoretical MNIST example, a new autoencoder will be trained only on the 5s of the handwritten data set. The digit 5 was selected since it typically had the highest error rates as seen in Table 38. In this example, it is easy to accomplish the separation without special ML sorting, since the digits label for both testing and training data sets were provided. The K-Means clustering step will be completed in the Case Study reviewed in the next section.

By limiting the data the autoencoder is trained on to only 5s, the model will learn how to properly encode and decode 5s better than other digits, since it will never have seen any of the digits to learn from prior to the testing. The goal with this is to build the autoencoder with the ability to reproduce the 5s well but struggle to produce other digits. If it cannot produce other digits well, the error rate when comparing the input versus output will be large, representing an anomaly. The level of accuracy of this process is highly dependent on the consistency of the data used to train. In this case, the 5s it was trained on vary considerably as illustrated in Figure 108. Ideally, in a controlled die cast process, the variation will be significantly less than handwritten digits by thousands of different people.

Figure 108: Example 5s from training data set

The autoencoder was built matching the structure of Example 2 as seen in Table 36 as it provided the lowest overall average error from the examples tested as per Table 38. The autoencoder was trained on all 5421 example 5s pulled from the 60,000 digit training set. The mean absolute error rate by digit is plotted in Figure 109. This shows that the lowest level of error is on the digit 5s, which would be expected. All other digits showed a higher level of error, although the box plots show overlap.

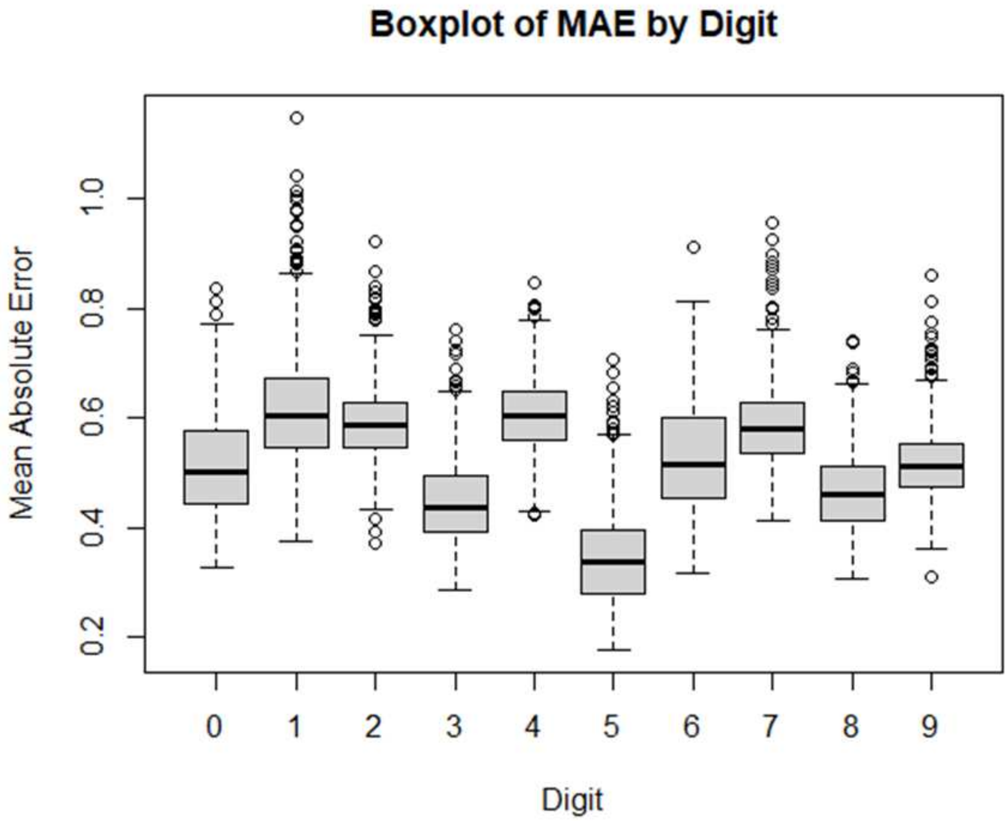


Figure 109: Mean absolute error (MAE) based on autoencoder trained on 5s only

Figure 110 shows the performance of the trained autoencoder from both a visual and error standpoint. The figure shows the original test image next to the reconstructed or predicted image the autoencoder created. The resulting mean absolute error for the prediction to the test is also shown for each example. It becomes clear that the autoencoder is trained well at producing 5s, as every input example that was tested had a prediction that visually looks like a 5, although some of them were faded and missing portions.

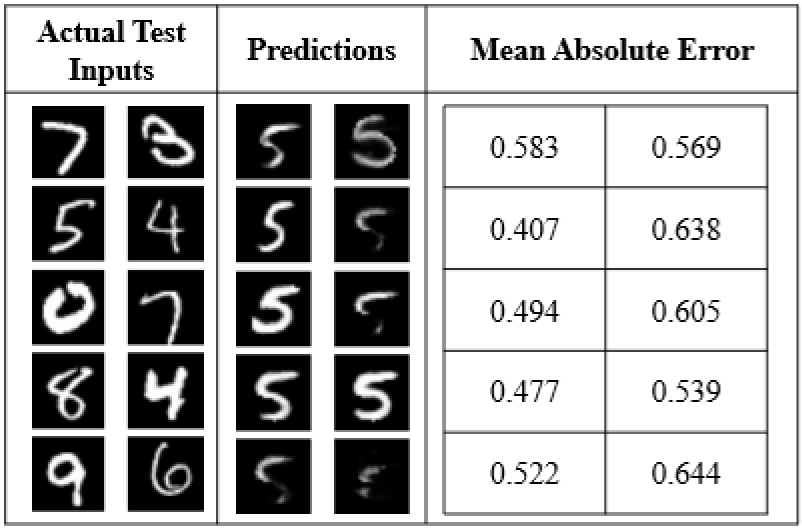


Figure 110: Visual and error examples from test data

Based off Figure 109, with the mean and standard deviation of the error rates for the performance on known digit 5s in the test data set, a threshold error rate of 0.42 was selected to see how well the autoencoder would perform in this theoretical example. If the error rate was above 0.42, code was developed to mark the image as an anomaly. Anything below that value would be marked as normal. The performance of the anomaly detection varied significantly based on the different digits. Digits like 3, 6, and 8, which have many physical features like a 5, performed much poorer than digits like 1, 2, 4, and 7, which are physically different. Table 39 shows the details of the test counts and percentages for the 10,000 test data samples.

Table 39: Anomaly Detection performance of autoencoder trained on 5s

Test Digit	Test Samples Identified as Anomaly	Test Samples Identified as Normal	% Anomaly	% Normal
0	600	380	61.2%	38.8%
1	916	219	80.7%	19.3%
2	992	40	96.1%	3.9%
3	377	633	37.3%	62.7%
4	892	90	90.8%	9.2%
5	45	847	5.0%	95.0%
6	516	442	53.9%	46.1%
7	928	100	90.3%	9.7%
8	491	483	50.4%	49.6%
9	500	509	49.6%	50.4%
Overall Anomaly	6212	2896	68.2%	31.8%
Overall True 5	45	847	5.0%	95.0%

There were 45 different 5s that performed so poorly within the autoencoder that they were also identified as an anomaly. Figure 111 shows a portion of these poor performing 5s. Upon review, it is clear the variation in hand-written digits between people increases the probability of some of these being marked as anomalies by the autoencoder.

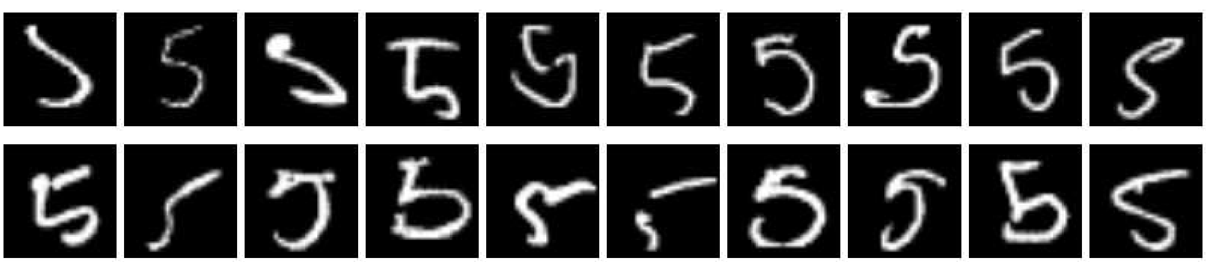


Figure 111: 5s identified as anomalies by autoencoder

Theoretical Example Conclusions

Admittedly, the performance of this theoretical example may be considered marginal. It performs better on certain digits as compared to others, but if this was a production setting, just over 30% of anomalous digits would have not been detected. The differences in hand-written digits by many people creates considerable variation that affect the training of ML algorithms. Many manufacturing processes with sensors collecting data produce more repeatable data than human handwriting. It is expected that an in-control process would be able to detect anomalies much better than this theoretical example. Additionally, selecting only one digit out of ten as a "normal" and all other digits being anomalies, means a highly unbalanced data set where normal is a small percentage and anomaly is a high proportion. This typically would not match most manufacturing examples. Finally, different error functions could have been reviewed to penalize differences more, which could lead to better performance of the anomaly detection.

This theoretical example with the MNIST digits was important to review to provide a clear, visual understanding of how an autoencoder functions. Handwritten digits are easy to understand and visually see the differences. It also provides a straightforward way for someone who may not have experience with ML to gain the basic understanding of the process of building an autoencoder.

An important take away from this example is how an autoencoder has the power to learn multiple different patterns. When training for all digits, the autoencoder was able to learn how to recreate 10 different digits all within the one encoder. This is extremely important from an anomaly detection perspective. The goal is not to recreate an anomaly with minimal error, but instead have a significant error so it is detectable. Selecting which data to train the algorithm on is a critical step when applying this to data generated within manufacturing. The next section will walk through these steps again, using an example of time-series data from a HPDC process.

CASE STUDY

This autoencoder anomaly detection approach can be performed on time-series data that is captured within the HPDC process. A case study was performed at Mercury Marine die casting plant in Fond du Lac, Wisconsin. A flow rate sensor was installed to the die lube spray system to measure the flow rate during the spray cycle. This sensor was a Keyence brand clamp-on sensor, model FD-R. The sensor was attached to the die cast machines PLC, where flow rate data was collected every half second.

Ignition software from Inductive Automation was used to pull the data from the PLC and write it into a SQL database as a comma separated value (CSV) within a database column. The first 45 seconds of the spray program are spraying, and the remainder is time used for air blow-off of the die. Figure 112 shows an example flow rate profile from the data collected from spray. The same steps of the theoretical MNIST example were followed for this case study. Additionally, some theoretical flow rate profiles were tested on the autoencoder built to see how well it would perform on potential failures within the system.

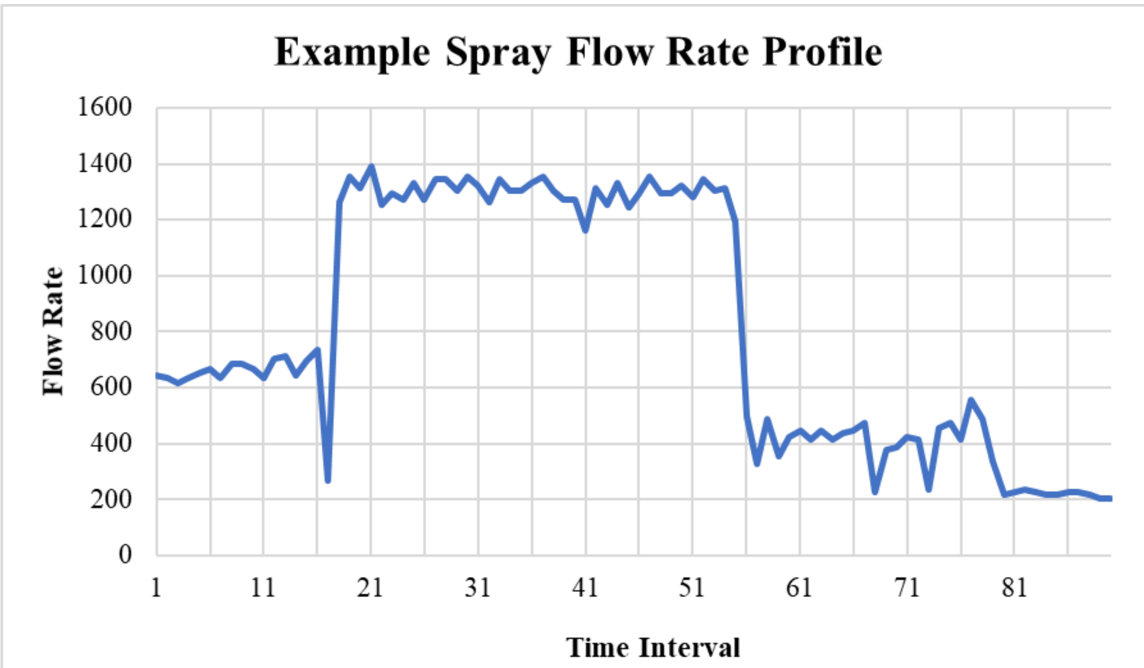


Figure 112: Example spray flow rate profile

The data used in this case study represented two consecutive weeks of production. The first week was used in the training of the autoencoder, while the second week was used for testing the autoencoder's anomaly detection capability. In total, approximately 1,500 profiles were used for training and 1,300 for testing.

Step 1: Prepare Data

The data required pre-processing and cleaning before it was ready to be used. The pre-processing of the data involved structuring it in a means that is friendly for analysis. The data was stored as a CSV string within the database, along with other information like serial number, date, time stamp, and string length as seen in Figure 113. As seen in the MNIST example, data flowing into an autoencoder needs to be separated out to individual columns. Code was written to convert this CSV string field into separate columns of data to represent each time interval. With that pre-processing completed, the data was now in a format that could be reviewed with plots.

	ID	tstamp	shotID	SeqNo	SeqLen	BString	CSVString
1	216464	2021-06-07 23:58:39	21159000163	0	200	raw(0)	227, 237, 237, 655, 715, 695, 695, 695, 705, 695, 695, 715, 70...
2	216463	2021-06-07 23:55:38	21158231963	0	200	raw(0)	217, 206, 715, 666, 675, 675, 685, 685, 695, 685, 715, 704, 70...
3	216462	2021-06-07 23:52:37	21158231863	0	200	raw(0)	217, 227, 625, 655, 685, 685, 685, 695, 685, 705, 705, 695, 70...
4	216461	2021-06-07 23:49:36	21158231763	0	200	raw(0)	227, 229, 605, 685, 685, 682, 675, 695, 695, 699, 695, 685, 69...
5	216460	2021-06-07 23:46:36	21158231663	0	200	raw(0)	217, 237, 675, 675, 675, 665, 685, 685, 685, 685, 695, 685, 68...
6	216459	2021-06-07 23:43:35	21158231563	0	200	raw(0)	220, 206, 206, 728, 675, 675, 685, 685, 695, 685, 705, 695, 69...

Figure 113: Raw SQL data storage examples

Two steps were needed to clean the data. First, the data was corrected for shifting associated with the PLC trigger point. Due to signals, some of the data stored would start collecting before the spray would begin. As a result, the profiles would appear delayed. This was corrected within the data as the first step of the cleaning process. The length of the CSV string is 200 digits, as predetermined in the PLC sampling. The next cleaning step is to trim the columns to match the actual process of blow-off for 45 seconds or 90 intervals. Based on this analysis, improvements to the data collection process are forthcoming to minimize this pre-processing and cleaning of the data. The formatting of data to prepare it for ML can be a time-consuming process. Storing and formatting the data in a consistent manner for analysis will reduce the time spent preparing the data in this stage.

To determine the fitness of the training data to remove anomalies so the autoencoder does not pattern on those examples, a K-means clustering algorithm was ran on the training data. K-means is an unsupervised ML algorithm that groups data by the commonality of its input data [222] . Cluster sizes from 2 to 11 were reviewed, and based on assignments, a cluster size of 5 was used to separate the training data. An average of the profiles for each cluster was created and then plotted in Figure 114.

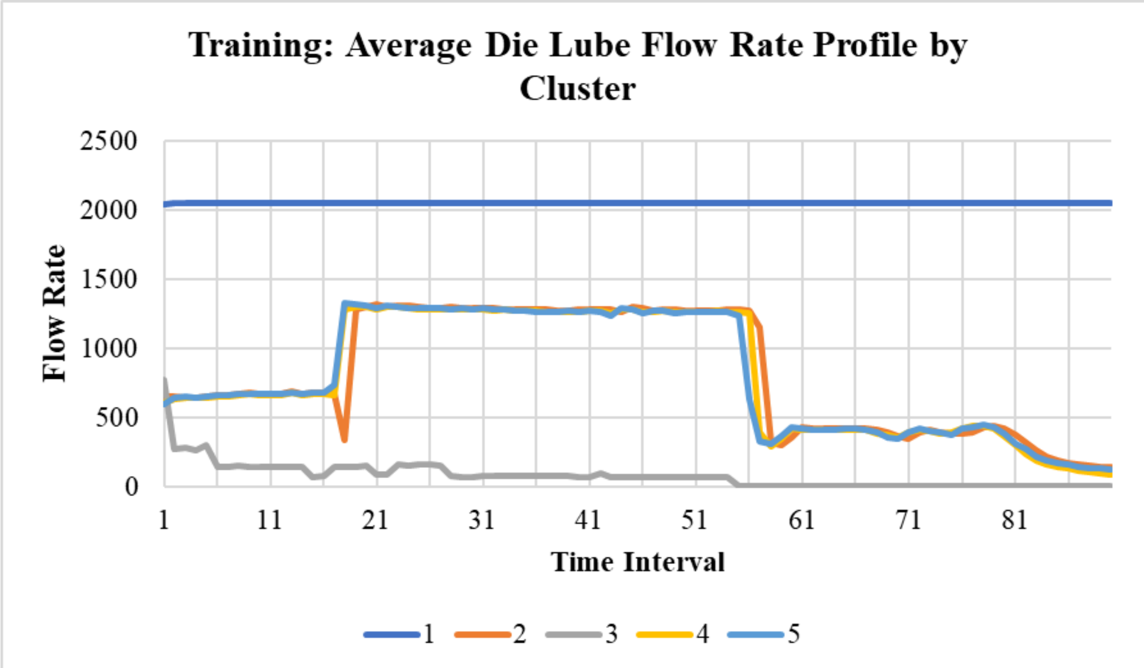


Figure 114: Training data clusters

Upon review of the cluster assignments, Cluster 1 was two cycles in which the die lube hose broke off the robot and flooded the die. Cluster 3 was a separate case where the airline to the valve broke, preventing the opening of the nozzle to spray for 3 cycles. These are precisely the types of anomalies for the autoencoder to detect. The 1556 parts comprising of clusters 2, 4, and 5 were selected for the autoencoder to be trained. The data was scaled between 0 and 1 by dividing by 2500. Max value within the training data set was 2054.

Step 2: Define the Autoencoder

A simple autoencoder structure was selected for this time-series data. The input and output layers are 90 neurons, given the 90 intervals that exist in the trimmed data set. A code layer of 20 was selected along with one hidden layer of 40 for the encoder and decoder. Sigmoid activation functions were used throughout the example. Table 40 contains the details of the autoencoder built for the spray lube case study.

Table 40: Case study autoencoder structure

Input Layer Size	90
Output Layer Size	90
Structure of Encoder	1^{st} layer = 40
Code Layer	20
Structure of Decoder	1^{st} layer = 40
Hidden Layer Activation Functions	Sigmoid
Output Activation Function	Sigmoid

Step 3: Train the Autoencoder

Because the overall number of samples was low with 1556 profiles, a batch size of 2 was selected for the data to be trained with an epoch of 100. The Keras optimizer *Adam* was selected along with a loss function of absolute mean error. The training time was 2.13 minutes on the same computer hardware that was utilized for the MNIST example. There was a significant number of iterations associated with defining the autoencoder, training, and testing to get to the published results. This would be typical of developing an ML algorithm.

Step 4: Test the Autoencoder and Anomaly Detection

To test the autoencoder and build anomaly detection, the profiles from the second week were utilized and pre-processed identical to the training data. There was a total of 1283 test samples. This testing data was processed by the autoencoder, and a regenerated time-series profile was created for each example. Figure 115 shows one the original input and the recreated profile from the testing data.

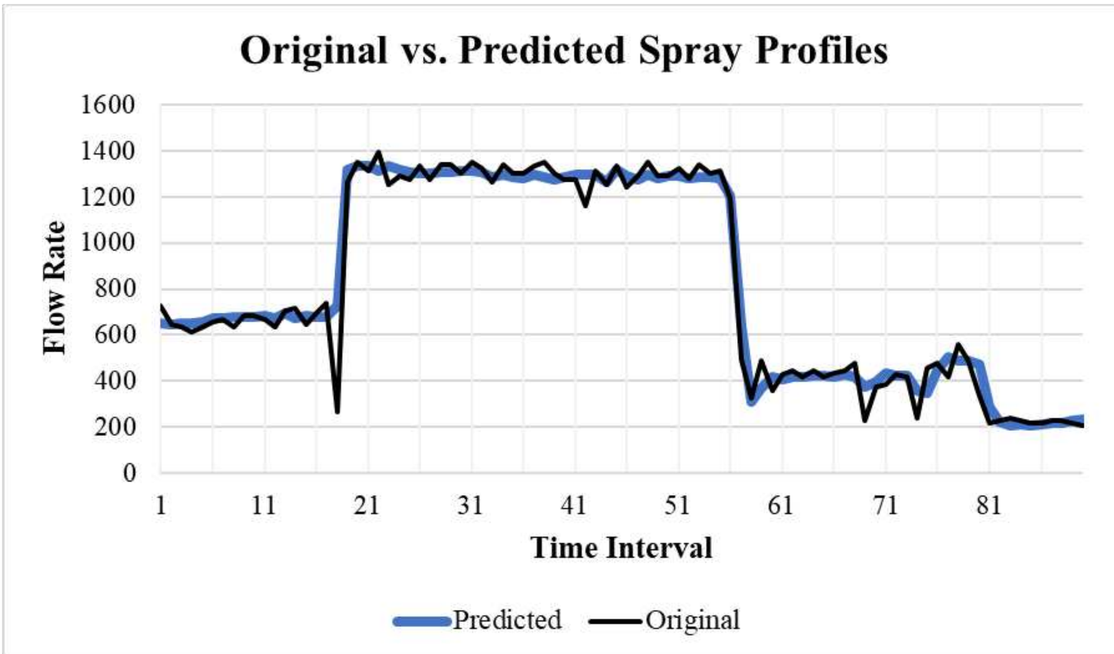


Figure 115: Original input and recreated profile

The mean absolute error (MAE) for the example in Figure 115 was 40.7. The MAE was calculated for all the test samples. A plot of the MAE over time can be seen in Figure 116.

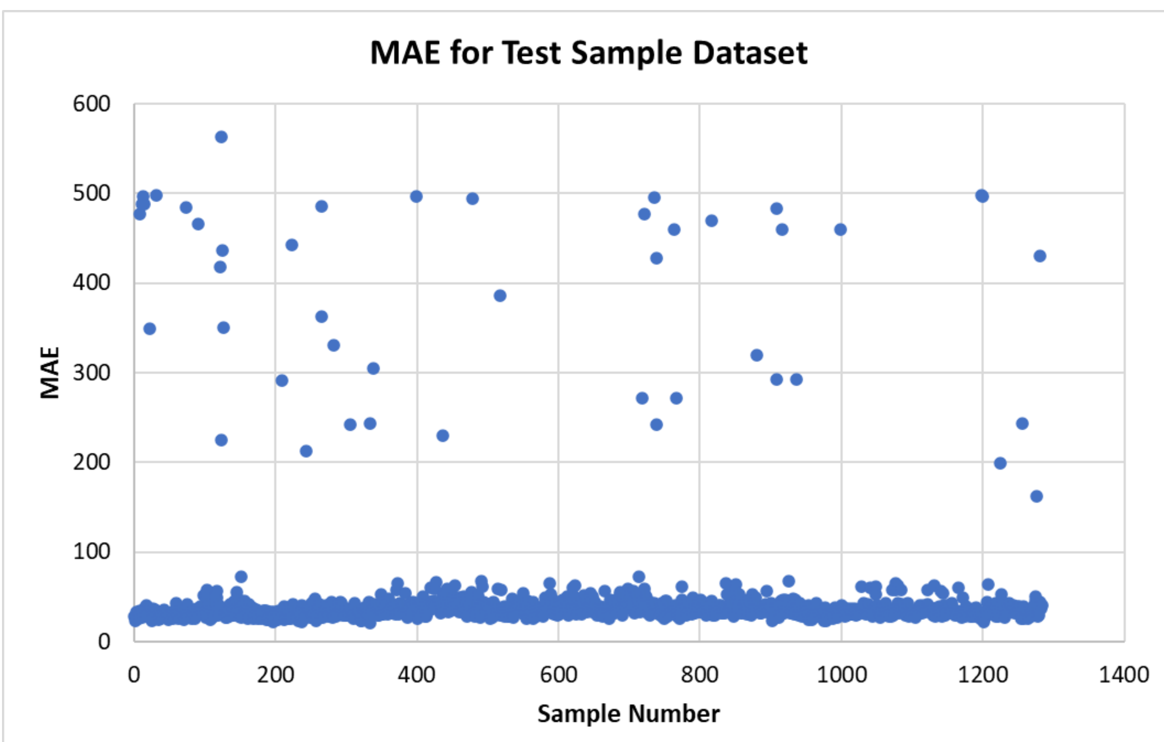


Figure 116: MAE of test samples

Based on the graph of the individual MAE values, it appears as if a there is a threshold at a score of approximately 100. To test this, the mean and standard deviation of the entire test data set was compared to the mean and standard deviation of just the samples under 100 MAE. Based on the statistics presented in Table 41, the anomalies identified by the autoencoder add significant amounts to the standard deviation and drive up the average value. As such, the thresholding should be based on statistics of the non-anomalous readings. Although the MAE results were non-normal, one general approach is using the mean plus three standard deviations to create an upper threshold. This works well for normally distributed data. This same approach can be used as a general guideline for distributions that are near-normal. In this example, the threshold would be 59.0. Anything above 59.0 would be flagged as an anomaly within the spray process.

Table 41: Test samples MAE Statistics

Test Data Set Statistics		
All Data	**# of Samples**	1283
	Mean MAE	48.4
	Standard Deviation of MAE	68.53
Under 100 MAE	**# of Samples**	1237
	Mean MAE	35.9
	Standard Deviation of MAE	7.7

To illustrate the anomalies, four samples were chosen to represent the range of MAE scores. One value was chosen at the upper edge of the under 100 mark. Then samples were selected between 200 to 300, 300 to 400, and 400 to 500. The error value for each along with the actual versus predicted graph is shown in Figure 117.

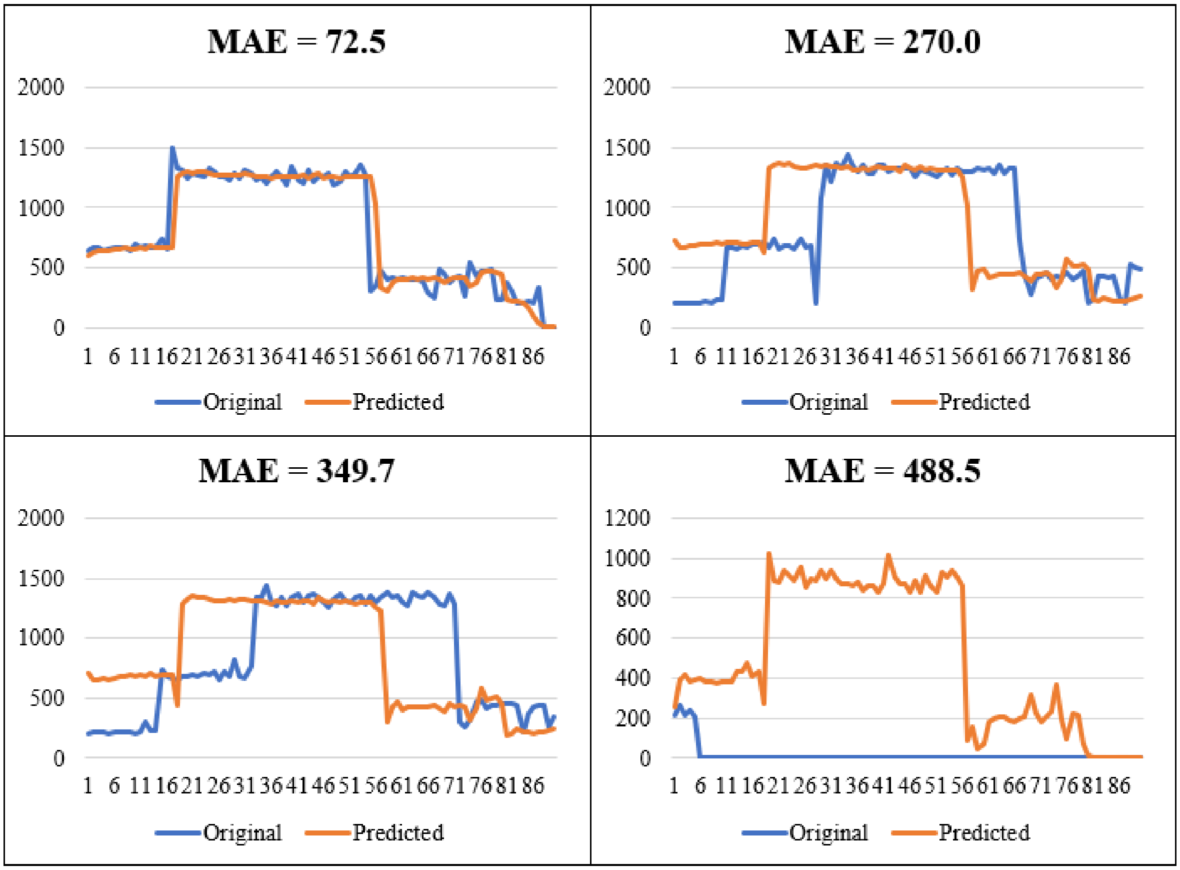


Figure 117: Illustrations of anomalies

Overall, the autoencoder performed well at identifying anomalies in the test data. Except for the 488.5 MAE profile which represents a spray cycle that was stopped, the rest of the anomalies detected represented shifts in the time-series data. This shift detection was typical of most of the anomalies detected by the algorithm. This was accounted for by a data collection trigger issue with the PLC that was longer than the corrected value that was done at the pre-processing. To further see how well the anomaly detection performs, additional theoretical testing will be performed in the next section.

Theoretical Anomaly Detection Testing

By their nature, anomalies do not often happen within the process. A set of theoretical failures was produced to test the performance of the autoencoder to detect anomalies. Two different sets of anomalies with three varying conditions were tested. These conditions include:

- 15%, 25%, and 35% flow rate reduction throughout the process – representing reduced pressure in the system
- 15%, 25%, and 35% flow rate reduction in the high flow cycle – representing a blocked or partially blocked nozzle

The original profile that was used as a baseline in this theoretical example, along with the modified profile that represents the anomaly and then the predicted value from the autoencoder were all plotted for comparison in Figure 118. Visually, the autoencoder did not perform well when the entire profile was scaled. It did perform well at identifying the anomaly when only the high portion of the flow rate was reduced.

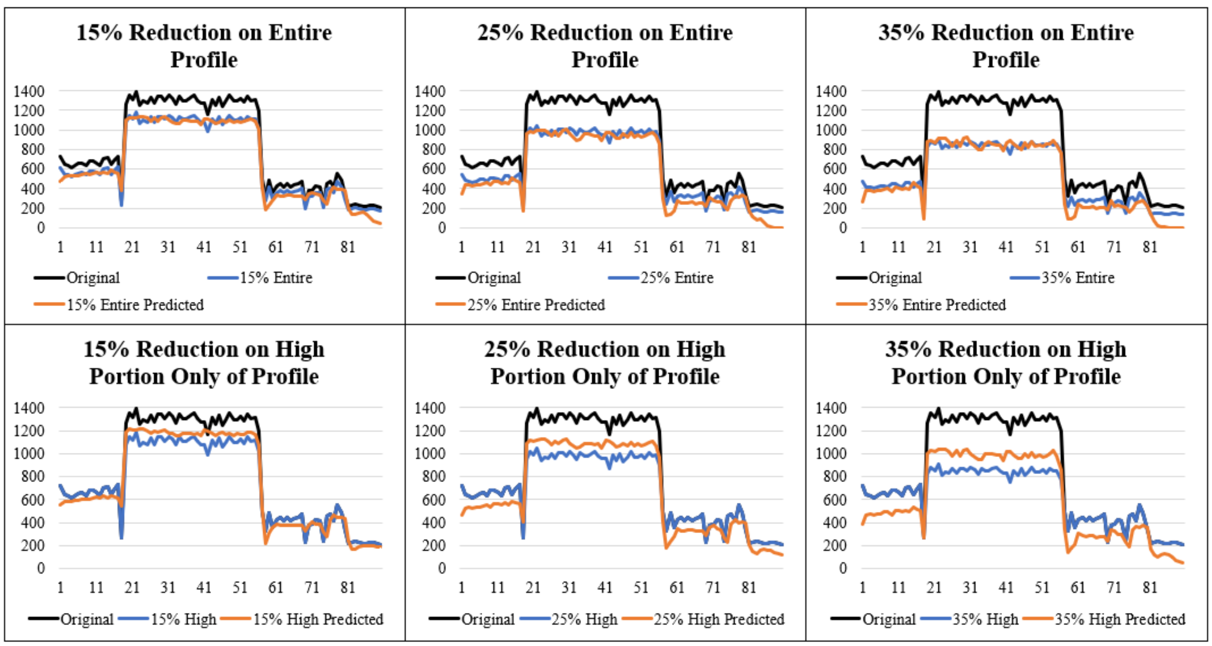


Figure 118: Theoretical anomaly detection profile graphs

The MAE was calculated for all six examples and the result can be seen in Table 42. The MAE values ranged from 45.5 to 144.9. Utilizing the statistical driven threshold of 59.0, all the profiles with a reduced high flow rate period would have been detected as an anomaly. However, only one out of the three shifts of the entire profile would have been identified as anomalous. Once the triggering issue on the PLC is corrected, further work will be done on this data to improve the training of the autoencoder. With more consistent training data, the anomaly detection for the autoencoder should improve beyond the results shown here.

Table 42: MAE Values for six different theoretical examples

Name	MAE Value
15% Entire Profile	45.5
25% Entire Profile	61.4
35% Entire Profile	58.4
15% Reduced High Profile Only	67.5
25% Reduced High Profile Only	104.0
35% Reduced High Profile Only	144.9

CONCLUSIONS

The HPDC process contains many different potential time-series data streams. From the typically reviewed velocity and pressures associated from the injection process to the flow rates of die lube during spray, the HPDC process can generate extraordinary amounts of data. This volume and velocity of data is too much for a human to analyze. The use of machine learning techniques provides opportunities to streamline the data analysis and help the operator know when important changes have occurred within time-series data.

The data being generated by the process must be analyzed to provide value to the user. Descriptive statistics involving a portion of the time-series data can be calculated and trended. This is an excellent first step and is currently utilized in many shot-monitoring systems throughout the HPDC industry. This descriptive statistics approach can help a foundry dial in its process and provide a high-level monitoring system for injection.

The shortcomings of descriptive statistical analytics need to be understood. Important information, such as slopes, noise, and other changes within the data will be lost with averages. Different profiles may have the same average, but not produce the same quality casting. Beyond shot monitoring, there is limited commercially available systems that allow a HPDC foundry to follow these steps for other time-series data. Approaches are needed to help analyze changes throughout the time-series data and flag irregularities when detected. Autoencoders provide a useful means for helping detect these types of anomalies.

Autoencoders utilize machine learning neural networks for time-series anomaly detection. As shown in the case study, these autoencoders can be quickly trained in a matter of minutes on typical computer hardware, utilize free open-source software, and be able to develop an effective anomaly detection algorithm on a limited size data set that can be created within a foundry in less than a week. The use of autoencoders can provide real benefit for advanced analytics within die casting. Autoencoders can also help remove the data processing from the operator or technician and provide warnings when the process is changing. This review of the data becomes more important as the industry sees increasing number of retirements and the skills gap that exists with newer employees. The use of machine learning to analyze the data, like autoencoders, could help foundries overcome the knowledge gap that exists between new employees entering the field and seasoned technicians heading to retirement.

It is also important to understand some of the shortcomings associated with autoencoders to set proper expectations with the technology. First, the input sample length must be fixed. In this example, changes in spray time where additional data is being collected, would error out any previously produced autoencoder because of the fixed input layer. In the case study, this input layer was 90 because of the 0.5 second time intervals and 45 seconds of spray. If a tool change drove a longer spray time, then the entire autoencoder would have to be rebuilt based on the new input data. Additionally, if known changes to the process occur, such as changes of when and how many nozzles spray, the autoencoder would need to be retrained. Otherwise, it may just detect everything as an anomaly based on the previous process when it was trained.

In conclusion, huge data sets and serialized product with detail part quality tracking are often the initial idea when considering machine learning implementation within manufacturing. There are considerable challenges to this type of analytics, which may not be the right first step for a foundry. Utilizing an autoencoder to detect anomalies can be viewed as a foundational building block to advanced foundry analytics. It provides an opportunity to collect additional time-series data while applying unsupervised machine learning to detect anomalies. Additionally, autoencoders can be developed on open-sourced software languages and packages, without the need for huge data sets. This use of autoencoders is just one of many opportunities for a foundry to engage in machine learning and Industry 4.0. Taking the first step in trying these technologies, is often the hardest. The goal of this work is to show the HPDC industry the challenge of implementing machine learning is not as insurmountable as one may imagine it to be.

Chapter 6: Conclusions and Future Work

Chapter 6 concludes this work and provides direction for future work for others to further explore this field. A summary of the chapters is provided before key conclusions resulting from this work are discussed. This chapter also contains a detailed list of the author's contributions that have been made during this PhD endeavor. Finally, a review of future areas of work are discussed in detail.

REVIEW

Two colliding factors are fundamentally and rapidly changing manufacturing: the manufacturing labor shortage and the digitization of manufacturing with Industry 4.0.

The COVID-19 pandemic has greatly accelerated the manufacturing labor shortage [240]. A recent study predicts that by 2030, the United States will have 2.1 million unfilled manufacturing jobs. Because manufacturing has the highest multiplier effect of any economic sector, these unfilled positions could negatively impact the US economy by more than $1 trillion USD by 2030 [241]. Businesses will need to become creative when trying to solve the labor shortage. Automation will assist with the shortage of manufacturing workers, but it will not fill the gap completely. Instead, companies must investigate means to increase productivity. Industry 4.0 provides tools that help improve productivity. However, this digitization of factories creates additional labor requirements. Creating and working within cyber-physical manufacturing worlds requires a skillset significantly different than the one possessed by the traditional manufacturing employee.

Data drives value in the cyber-physical manufacturing world of Industry 4.0. Data can be used to optimize processes, improve quality, reduce downtime, provide predictive maintenance, stop production equipment before it fails, automate scheduling, and provide insight on the entire manufacturing system. The improvement in productivity and quality could help offset labor shortages. However, as discussed in Chapter 2, Industry 4.0 is a complex system being added into an already complex manufacturing system. Many organizations have limited success, even after years of work.

The labor shortage and implementation of Industry 4.0 will change manufacturing. Factories must respond to the challenges and opportunities these two factors present. High pressure die casting is no exception to this change. This work presents a roadmap for the die casting industry to better understand the manufacturing systems that exists in its factories.

In Chapter 1, an introduction to die casting was provided. Additionally, the chapter highlighted a knowledge gap that exists. Although die casting is well studied and published, the research provides many shortcomings with optimization and application of machine learning. Much of this research is done in academic settings and even completely with simulation. The ranges used within the optimization algorithms often are not realistic and values would never be used within production industry. In other words, an experienced die casting engineer could look at the casting and the ranges being studied and select likely the best option without any advanced analytics. Applications of machine learning within production settings are limited and provided motivation for this work.

Chapter 2 provided the groundwork for understanding Industry 4.0 and complex systems to lead to the die casting data framework for future machine learning opportunities. Die casting contains the four characteristics of complex systems: network structure, adaptive, self-organizing, and nonlinear. The complexity of die casting requires a system engineering approach. A systems review has not been applied to production die casting to date. This decomposition of the die casting system provided the data framework, and the 10 different framework groups that were defined. The volume of data defined in the framework was orders of magnitude higher than ever discussed or considered within the industry previously. This volume of data requires machine learning to help humans process it in the short cycle times of die casting.

Before machine learning can be applied, additional research was completed into reasons why applications within production environments may failed. In Chapter 3, the stochastic nature of casting defects was reviewed. Defining macro versus micro porosity provided a foundation for then studying the macro porosity that mostly impacts part quality in production casting environments. Castings are judged to porosity specifications at a macro scale. Therefore, it is important to understand the formation of this macro porosity. The case study in Chapter 3 showed that although the porosity predicted zone could be established and was reliable with simulation, the precise prediction of size and location of macro porosity was randomly distributed within this area. Chapter 3 provided the foundation for a more thorough review of classification issues within die casting in Chapter 4. The random formation of porosity is one of four elements that impact defect classification. Combined with binary acceptance specifications, secondary process variation, and visual inspection, the level of misclassified castings in a production environment cause issues with machine learning applications. Misclassification creates data space overlap, which leads to poor machine learning accuracy.

In addition to describing the four misclassification elements for die casting, Chapter 4 also introduced an important topic in machine learning accuracy: the Critical Error Threshold. Because die casting and most other manufacturing processes produce high levels of quality (~95%), the accuracy of a machine learning algorithm applied must be near perfect to provide value for the investment. This accuracy level becomes a challenge. To provide perfect predictions, all critical input data must be available. Without a systems approach to the data and process (Chapter 2), algorithms would not have all the necessary information. The challenging part is even with all the critical input data, the accuracy of the machine learning algorithm still may not be at 100% due to the complex and stochastic nature of die casting.

Chapters 3 and 4 provide a frame of reference for applications of machine learning within a production environment. Knowing the shortcomings of machine learning applications is important because that provides insights into areas where machine learning can add significant value such as anomaly detection within the process. Chapter 5 highlights four different case studies of actual machine learning applications within a production manufacturing environment. Use of unsupervised machine learning, such as clustering and autoencoders, provided means to identify changes within the process. They also helped control and improve the part quality and productivity. Adding actual value with the use of machine learning in a production setting achieved the goal set at the onset of this work.

In summary, this work provides a roadmap for the die casting industry to assist with better understanding the process. Ideally, this will benefit not only the die cast industry, but manufacturing in general, confront the challenges of the future. Much was learned during this work with the key conclusions summarized in the next section.

CONCLUSIONS

Complexity and data are the two themes that run throughout this dissertation. Not surprisingly, these two items are related. Large volumes of data are needed to monitor and provide feedback because of the complexity of die casting. Control of the system can be improved with this feedback. In the end, these two themes create five key understandings:

1) Die casting is a complex system.
2) A systems engineering approach is needed for die casting.
3) The scale of data that can be generated is unlike anything discussed to date for die casting.
4) Because of complexity, 100% accuracy for machine learning models is improbable.
5) The focus of machine learning should be to detect anomalies and control the process.

By formal definition, die casting is a complex system. Die casting is a network structure of different systems interacting with each other to produce near net shape castings. The die castings system is a system-of-systems. Die casting adapts and evolves over time. The physical nature of die casting with wear and erosion of the tooling causes changes. Technology and humans within the process also make the system evolve. Die casting foundries will self-organize. Many improvements that are undertaken in a foundry are not top-down mandates. Instead, lessons learned by a technician in one area will be applied to other areas. Finally, die casting is nonlinear, and therefore, small changes within the system inputs can cause significant differences in the output. The physics of fluid flow as the metal fills the die and heat transfer as thermal balance is created within the process are examples of nonlinear systems in die casting. This nonlinearity can help explain the stochastic nature of porosity formation found in the macro porosity case study. Additionally, human input also provides a level of nonlinearity within die casting. In a technical sense, a casting with porosity should always be identified as such during quality inspections.

This input should match the output expected. However, fallible humans regularly misclassify castings. In real life, the classification input to output relationship does not follow a linear relationship. This research shows how technically complex die castings is.

The second takeaway is a logical approach once die casting is defined as a complex system. Systems engineering exists to create successful complex systems. Systems engineering tools are needed to improve the die casting process. Looking at a portion of the die casting process will not create the overall quality and production improvement that the industry could achieve. The entire system needs to be defined, analyzed, and controlled. The systems approach of fragmenting a system into smaller and smaller components provides the needed detail to capture the data that can be generated within the process. The data generated is really the feedback cycle of the system that is missing in today's industry. The industry knows the injection parameters well, since it is the only system that has had data collection associated with it for several decades. The same cannot be said for the thermal management of the tool, which is critical to part quality, nor understanding the die cast machine health in terms of maintenance, unplanned downtime, and production losses. By collecting the data within the die casting system, feedback can be monitored and applied to the system to improve control. This improved control will drive the business metrics the industry needs to make positive change.

The systems approach of defining this data creates a scale of data within die casting that has not been discussed before within the industry. Experience and publications have shown the industry may not be ready for this volume of data. The NADCA publication on *Die Casting Process Control* [23], stated the "revolution of process control" has ended and then describes only a handful of parameters to monitor. A multi-million dollar study by a European consortium defined a total of only 75 data parameters in die casting [86]. The industry has a short-sighted focus on output variables as they are tied to casting quality. The industry also fails to leverage the available data. For decades, large time-series data sets on the injection process have been collected and effectively thrown away, as only a few statistical averages are typically reviewed and monitored within the industry. The industry overlooked the benefits of the data. As described in this work, there are multiple types of data that need to be considered including design data, input setting data, output data (time series and discrete) and cycle time data. Without a complete view, the feedback and control for the entire die casting system will not be successful. This data generates multiple orders of magnitude larger data sets than the industry currently uses. Fundamentally, the industry must change to process and monitor this data.

Predicting casting quality is the holy grail of machine learning. Unfortunately, the focus provided by research on this topic is misleading for the die cast industry. Publications showing the applications of supervised machine learning to predict quality in small, academic exercises may paint a picture that it can be easily done. Having large scale software companies echoing these comments, but often lacking detailed examples, makes it challenging for those working on the front lines of die casting data. This work should make one question the feasibility of supervised machine learning for quality predictions in a traditional manufacturing setting. There are two items confounding this feasibility.

First, the die casting industry has a great quality track record. A high percentage of parts are produced at acceptable quality levels even with a complex system with limited feedback and control. Industry survey shows an 8% median scrap rate. Although there is room for improvement on this value, another interpretation is the high yield rate means the industry already has great knowledge of the process. For machine learning to create value for the organization, the Critical Error Threshold states the accuracy must be higher than the yield rate of the process. This drives the supervised machine learning accuracy to be near 100%. This level of accuracy is not possible without a complete understanding of the entire process.

The second item confounding this quality prediction is the complexity of die casting. Complex systems are non-linear and sensitive to initial conditions. As such, applications of modeling complex systems are driven to achieve a pattern that the complex system creates, not a precise prediction. Without a precise prediction, there likely is no benefit in predicting quality in die casting. The exception to this statement is if the process is producing terrible scrap rates, for example 40% scrap. This is where machine learning could potentially provide some benefit but is the wrong tool for the job. At scrap rates above 40%, either the casting design, tool design, or process settings selected are likely outside industry standards. Instead of trying to optimize with advanced analytics, the faster path to improvement is by addressing the issues with the part or tool design. In most cases with poor scrap rates, the technicians and engineers associated with the casting already know what needs to be done to fix it. The challenge is getting the customer to agree or the time to complete it. Machine learning will not likely fix these situations.

The question this work should pose is if quality prediction is possibly not feasible, then how can machine learning still be used to add value within the die casting industry. This becomes the final takeaway from this work. Today, instead of focusing on predicting part quality, machine learning should be used as a tool to process the large volume of data that is generated from die casting. Machine learning can be used to improve anomaly detection, and ultimately, improve process control. The case studies provided within Chapter 5 highlight some approaches that can be utilized. By controlling the process, the metrics of quality and productivity will improve, which is ultimately the goal. To get to this point, the industry must collect the data per the framework. With this data in place, then utilizing clustering, autoencoders, and other anomaly detection tools becomes the final step in creating value with machine learning. Fortunately, this anomaly detection approach can be piecemealed within a plant. A complete data framework approach could take years to achieve for a single machine. By utilizing the framework and starting the process on one system, anomaly detection algorithms can start to be applied as soon as the data is available. Unsupervised machine learning generates value immediately, which creates a culture of using these advanced tools within the industry.

As stated previously, quality prediction is the holy grail. There is potential to achieve value, but the complexity of die casting is likely to prevent success. However, the approach outlined with this work (collecting the system-wide data for die casting and utilizing anomaly detection to improve process control) creates the data foundation needed to better control the process and reduce the noise that exists in complex, human-dependent systems. The best opportunity to predict quality is to have a complete understanding of every system associated with die casting and make it as repeatable as possible. This work provides an outline for this, and shows how to use this data in a means that creates value immediately.

CONTRIBUTIONS

Research and learnings gained in a vacuum provide no benefit to society. The knowledge must be shared, discussed, and further developed. I have learned much from the work completed over the past four years and taken many different opportunities to share within the foundry industry through various formats including papers, presentations, and industrial organizational groups. My goal with this participation is to voice the tangible benefits of machine learning in complex, production focused manufacturing facilities. This is where real value can be added. Improving castings quality and reducing costs in processing is the goal of any manufacturing facility. I believe the work presented here provides a roadmap for others to follow on this journey.

Below is a list of the contributions made to date and future work planned based on the knowledge learned during this PhD endeavor:

CONFERENCE PAPERS AND PRESENTATIONS

- D. Blondheim, Jr., “Initial Development of Machine Learning Algorithms to Predict Casting Defects in High-Pressure Die Casting,” presented at the 2017 NADCA Congress and Tabletop, Atlanta, GA, Sep. 2017. [Online]. Available: http://www.diecasting.org/archive/transactions/T17-073.pdf

- D. Blondheim, Jr., “Unsupervised Machine Learning and Statistical Anomaly Detection Applied to Thermal Images,” presented at the 2018 NADCA Congress and Exposition, Indianapolis, IN, Oct. 2018, vol. T18-071. [Online]. Available: http://www.diecasting.org/archive/transactions/T18-071.pdf
    - Won NADCA 2018 Die Casting Congress & Exposition Paper of the Year

- D. Blondheim, Jr., “Utilizing Machine Learning Autoencoders to Detect Anomalies in Time-Series Data,” presented at the 2021 NADCA Congress and Exposition, Indianapolis, IN, Oct. 2021, vol. T21-082. [Online]. Available: https://www.diecasting.org/docs/congress/congress21/papers/T21-082.pdf

- D. Blondheim, Jr., “Understanding the Critical Error Threshold for Applying Machine Learning in Manufacturing,” IJMC paper presented at the 2021 NADCA Congress and Exposition, Indianapolis, IN, Oct. 2021, vol. T21-081. [Online]. Available: https://www.diecasting.org/docs/congress/congress21/presentations/T21-081ppt.pdf

JOURNAL ARTICLES AND OTHER PUBLICATIONS

- D. Blondheim, Jr. and S. Bhowmik, “Time-Series Analysis and Anomaly Detection of High-Pressure Die Casting Shot Profiles,” *Die Casting Engineer*, pp. 14–18, Nov. 2019.

- D. Blondheim, Jr. and A. Monroe, “Macro Porosity Formation - A Study in High Pressure Die Casting,” *International Journal of Metalcasting*, Apr. 2021, doi: 10.1007/s40962-021-00602-x.

- D. Blondheim, Jr., “Improving Manufacturing Applications of Machine Learning by Understanding Defect Classification and the Critical Error Threshold,” *International Journal of Metalcasting*, Jun. 2021, doi: 10.1007/s40962-021-00637-0.

PRESENTATIONS

- D. Blondheim, Jr., “Introduction to Machine Learning for Die Cast Manufacturing,” presented at 2018 NADCA Plant Management Conference, Minneapolis, MN, May 2018.

- D. Blondheim, Jr., “Smart Manufacturing: Data Analytics, Statistics, and Machine Learning in Die Casting,” presented at NADCA Chapter 12 Meeting, Slinger, WI, Nov. 2019.

- D. Blondheim, Jr., “Artificial Intelligence, Machine Learning, and Data Analytics: Understanding the Concepts to Find Value in Die Casting Data,” presented at 2020 NADCA Executive Conference, Clearwater, FL, Feb. 2020.

- D. Blondheim, Jr., “Challenges and Approaches for Machine Learning in Manufacturing,” webinar presented as part of TMS Artificial Intelligence in Materials: Research, Design, and Manufacturing Webinar Series, Feb. 2021.

ORGANIZATIONAL INVOLVEMENT

- Peer review for *International Journal of Metalcasting* from 2017 to present.

- Reviewer and moderator the Quality/Machine Learning section of the 2020 NADCA Virtual Congress and Tabletop.

- Initiating member of the American Foundry Society (AFS) Working Group: Data Collection and Storage, part of AFS Engineering & Smart Manufacturing Division, Jul. 2021 to present.

- Initiating member of the American Foundry Society (AFS) Working Group: Data Analytics, part of AFS Engineering & Smart Manufacturing Division, Jul. 2021 to present.

FUTURE CONTRIBUTIONS

- Instructor for TMS online course *Artificial Intelligence in Materials Science and Engineering*, " Module 5, AI/ML for Materials Manufacturing: Understanding the Applications, Building Predictive Modeling, and Uncertainty Quantification," Nov. 2021.

- Abstract submitted for TMS First World Congress on Artificial Intelligence in Materials and Manufacturing (AIM 2022), "Effects of Complex Die Cast Manufacturing Systems and the Critical Error Threshold on Applications of Machine Learning in Production" Pittsburg, PA, Apr. 2022, potential publication in journal *Integrating Materials and Manufacturing Innovation*

FUTURE WORK

For me, one of the most challenging aspects of writing a dissertation on systems engineering is knowing what to include, but possibly more important, what to omit. A systems approach looks at the big picture. As such, it was difficult to know when to stop research and exploring. The challenge is that each connection made to the die casting cell meant a new dataset was stored. This new data became another rabbit hole to get lost in. At some point, though, I had to shelve certain ideas. Below is a series of ideas that provide a possible framework for future research. Some of these challenges I know I will continue to explore within my role at Mercury Marine to help improve the industry.

The concept of complexity in die cast is a topic that can be further explored. The main point I had within this work is to show that die casting is a complex system. As a complex system, it needs a systems engineering approach to solve. It also will suffer from a model that only predicts the pattern and not the precision. This concept needs further exploration so a realistic expectation can be set within the industry. Today, executives are bombarded by software companies promising to solve all their problems with AI or machine learning. Fancy presentations and one-off case studies are reviewed to show the millions in savings they can help generate. In certain cases, this is possible. For most, however, it creates an unrealistic expectation. Because machine learning is so new to manufacturing, it's often overhyped and glamourized. This dialog within the industry could be improved through investigation and publication on complex system and model accuracy.

Weaved into this topic of complexity is the holy-grail conversation of quality predictions within die casting. Realistically, if the process is poor enough, machine learning could provide insight into the process for optimization. Could is the key word. To gain insight, a range of input parameters must be tested, and that data tracked to part quality. Experimenting with ranges of input parameters does not always happen in production settings. Focus and priority are often given to production output versus the non-production time needed to test and improve the process. This becomes a catch-22. Building on that issue, in cases with a significant quality issues, there likely is a list of actions the engineers and technicians want to accomplish to help fix the problem. However, they are not allowed the time to execute the improvements because of production needs. The question must be asked: What are the ranges of scrap where quality prediction becomes valuable and worth the effort, even if it is not 100% accurate? At 30% failures, I would argue fundamental die cast knowledge should be applied ahead of machine learning for improvements. What happens if the scrap rate is 20%? Or 10%? Less than 5%? Is there a cut-off that can be found that provides guidance? At what scrap level should the focus be on 'Die Casting 101' versus machine learning? The Critical Error Threshold provides insight into the economics of quality prediction, but there is much work to accomplish to help select applications that are best suited to attempt predicting quality with machine learning.

The macro porosity study from Chapter 2 also presents a future work opportunity in better understanding die casting porosity. The work that went into that paper was extremely useful in showing that even with identical process parameters, the results of the porosity can be significantly different. An interesting question exists with these parts regarding their density. How does the density of the Best-of-the-Best (BoB) and the Worst-of-the-Worst (WoW) parts compare? An initial theory based on limited density testing shows that there is no statistical difference in BoB and WoW castings even though visually the porosity is significantly different in X-Ray. A colleague at Mercury measured the specific gravity of the 9 BoB and 9 WoW parts, and they showed no statistical difference. Details on some of the parts can be seen in Figure 119. In this figure, a BoB and WoW casting are shown with identical specific gravities (2.635), even though visually the void space is different. The third part in the figure is another

WoW casting with a higher specific gravity than the two others, even though visually it appears worse. Further study in this area is needed to understand how porosity and casting density are related based on this work. This could provide insight into the porosity formation within castings and help explain the randomness associated with results seen in x-ray or visual inspection of machines surfaces. This knowledge would be helpful in looking to build better machine learning quality predictions.

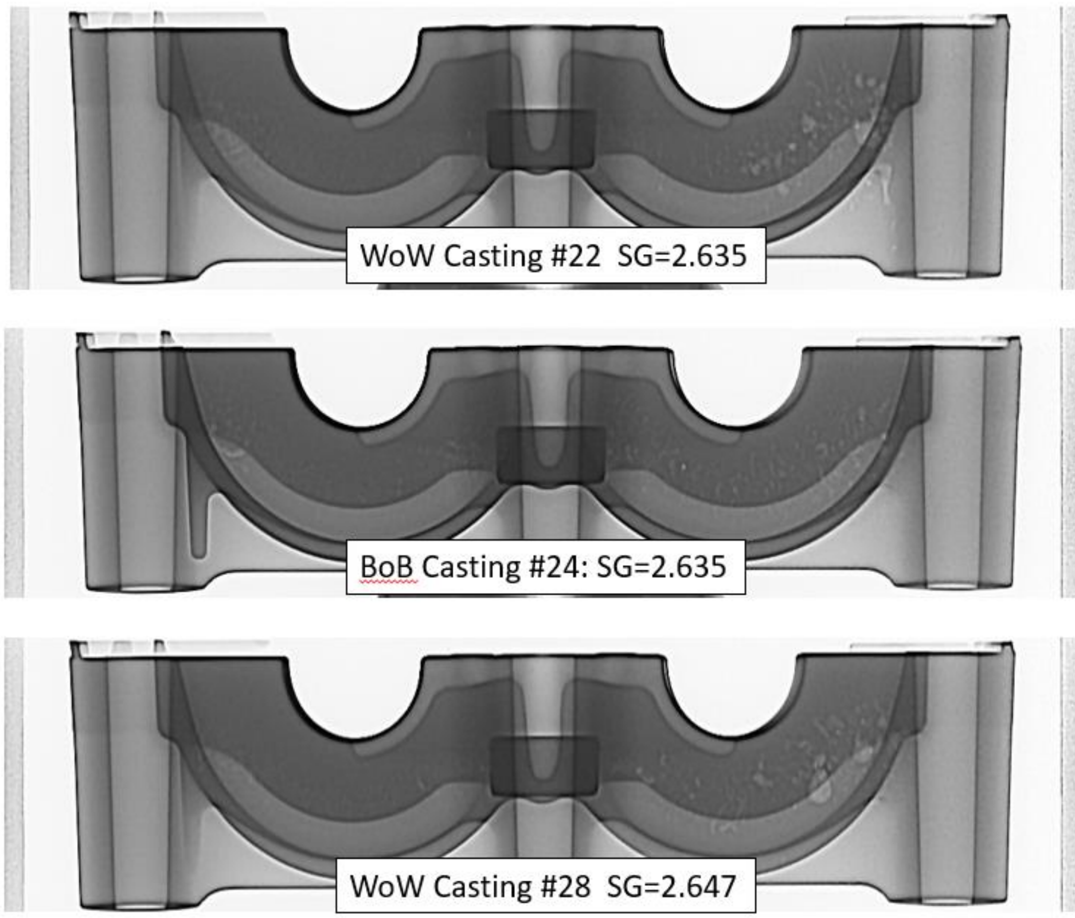


Figure 119: Specific gravity measurements and X-ray images on select BoB and WoW castings

Truly understanding process control is another phenomenon that builds off the predictive quality conversation. Experience has shown that many increases in scrap rates for a given casting can be attributed to special cause variation. Hundreds of water connections are made when setting up the die. If the inlet and outlet of a water fountain are connected in reverse, water stops flowing through that one cooling line. Is it enough to make a difference? The answer is usually not, until it does. At which point, the casting quality seems to fall off a cliff in terms of scrap percent. This is one example and hundreds or thousands more likely happen daily within the process. A study in production foundry environment researching the percentage of scrap generated by special cause variation versus defects generated from common cause variation would be beneficial to the industry. The industry tends to focus on the common cause variation defects, whereas I believe the special cause variation is a higher percentage and creates a much larger disruption to the production environment. Studies in this area would help guide the conversation to a better understanding of process control. This work provides the roadmap to machine learning applications for process control and anomaly detection through the data framework and the case studies.

In die casting, the thermal temperature profiles in the die steel drive the casting quality. Die steel temperatures can lead to changes in porosity locations, creation of blow-hole defects, and generate sticking or soldering of castings. The industry talks about steady state temperatures, with the assumption that after a limited number of cycles (typically 3 to 10), the die gets up to a set temperature and then remains at that temperature until the next downtime event. This thinking is flawed on both theoretical and practical levels. Miller argued in a 2016 NADCA paper that dies really enter a quasi-steady state equilibrium and can take hundreds of cycles to reach this steady state [44]. Practically, this assumption does not consider many of the different inputs into the die casting system. Metal temperatures delivered are not a flat temperature. The furnaces at the die cast machines go through heating cycles to try maintaining a set metal temperature. If left undisturbed, these cycles show a sinewave of readings. Its amplitude will be based on the design of the furnace controller to maintain a setpoint. In production, there is metal being delivered to that furnace as it is being removed by the casting process. The temperature of the metal add will vary based on how many stops the metal hauler had before getting to that furnace. This sinewave is reset each metal delivery. Figure 120 is data collected from Mercury Marine on a 2500T die casting machine during production. This is far from a flat line that goes into many of the assumptions on the die temperature reaching a steady state.

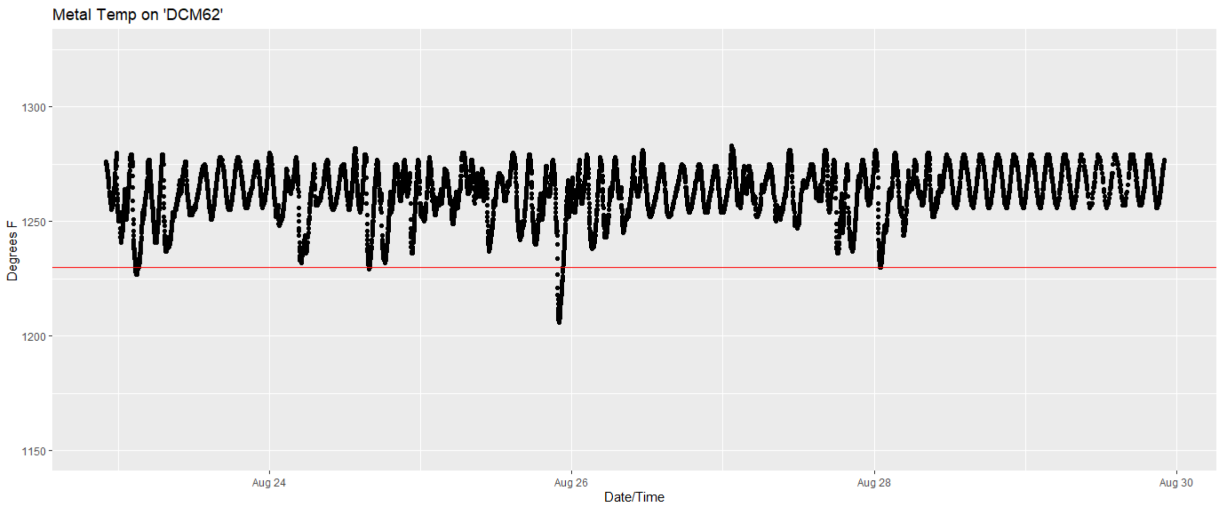


Figure 120: Example furnace temperature data

Furthering the study of die temperature understanding was an area of high interest to me after the 2018 NADCA paper [9]. Thermal imaging provides a significant amount of data and insight into the die surface temperature immediately before closing and the injection process. This is a powerful tool in better understanding thermals. In 2019, I presented at the November meeting of the Chapter 12 NADCA group on the topics of data and advanced analytics in die casting. In that presentation, I included the following two slides in Figure 121 that show some of the challenges with die temperature.

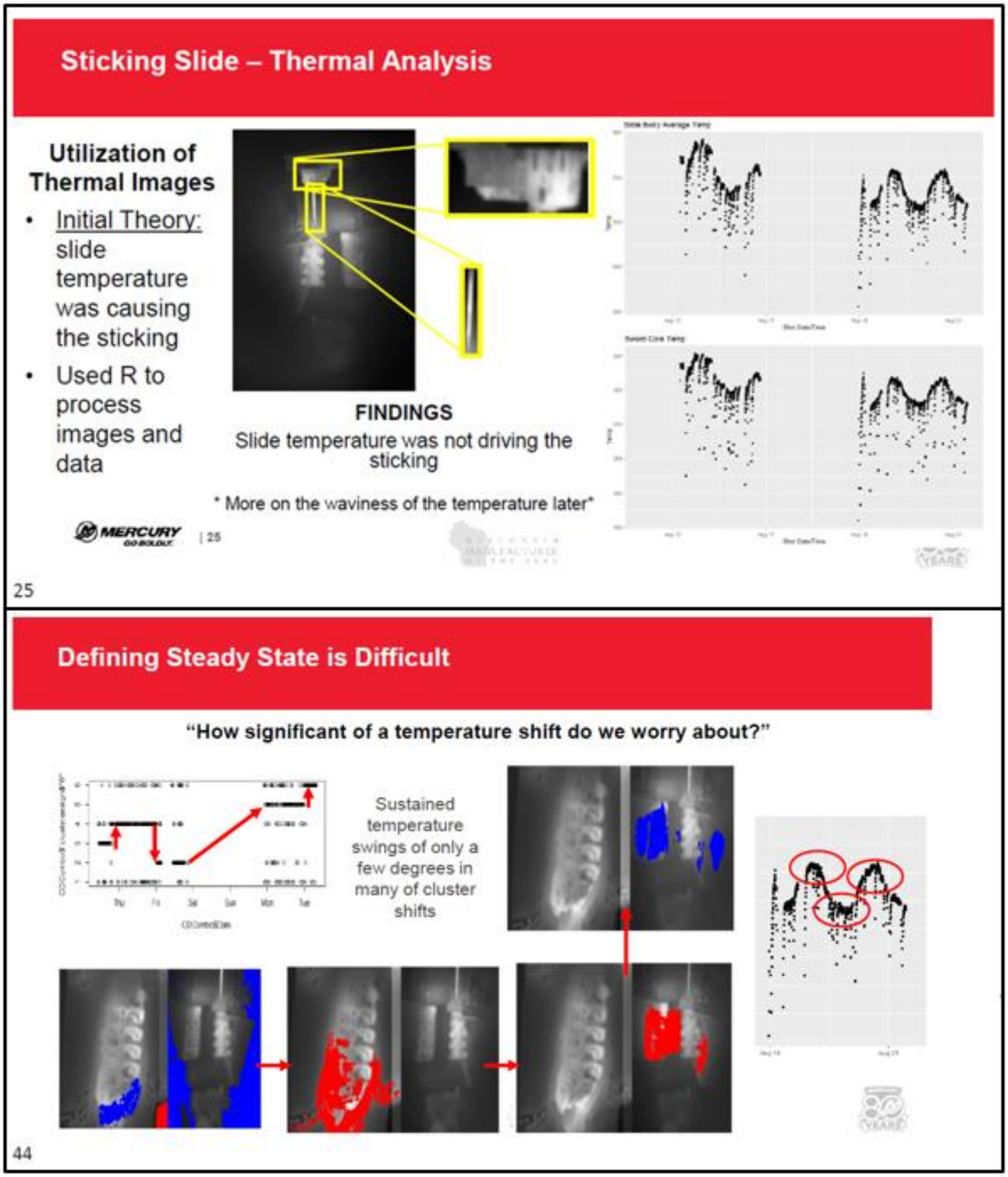


Figure 121: Slides author presented at November 2019 NADCA Chapter 12 meeting

The temperature of the die after spray seems to follow a sinewave but with a different period as the furnace temperature. This finding sparked my interest, and I pushed to have an advanced camera system installed on one of Mercury Marine's newest die cast machines at the beginning of 2020. Multiple vendor and IT delays have impacted this project. A solution is close to implementation and with the responsibility of schoolwork soon to be in the rearview mirror, this will be an area I plan to work on to help the industry gain a better understanding of die casting. Others, like Miller, have already discussed these types of topics, but with the data collected tangentially with this work, more effort should be given. The additional insights that can be gained from better data collection throughout the die casting process from the data framework provide a foundation to expand.

Finally, there are four areas of machine learning applications within die casting that I believe are due further investigation. The first of these is the concept of feature importance. In the conclusions of the *International Journal of Metalcasting* article that comprises Chapter 3 [133], an argument was made that feature importance could provide value in machine learning even if the prediction accuracy is not useful when compared to the CET. I believe the topic of feature importance could be expanded upon within die casting to help identify critical process parameters for a given casting. The data framework provided here again helps create that foundation. The second concept worth future endeavors is in predictive maintenance. The data from the equipment performance is by far the largest section of data from the framework. The vibration and amp draw sensors, at their high frequency rates, generate hundreds of thousands of data points. This data provides insights into the equipment health or performance. There has been no research published to date specifically focusing on predictive maintenance on die casting equipment. Since downtime is a large factor in die casting, studies in this area would yield meaningful benefits for the industry. Finding the ground truth in casting quality is the third area of research that is needed. As detailed in Chapter 4, misclassification of casting quality has significant impacts on supervised machine learning. One means to get to the ground truth of porosity in casting is with CT (computed tomography) scanning. An example of a CT scan on the casting used in Chapter 3 is seen in Figure 122. Industrializing this CT process for in-line inspection is key to gathering large volumes of data needed for supervised learning. Currently, the time it takes commercial equipment to create a CT scan of a casting significatnly longer than the cycle time of the die casting process. The CT equipment is also usually limited on the size of the castings it can scan. Research and improvements in this technology would drive a better ground truth result classification of castings for machine learning applications. The challenges to overcome are cost and processing time.

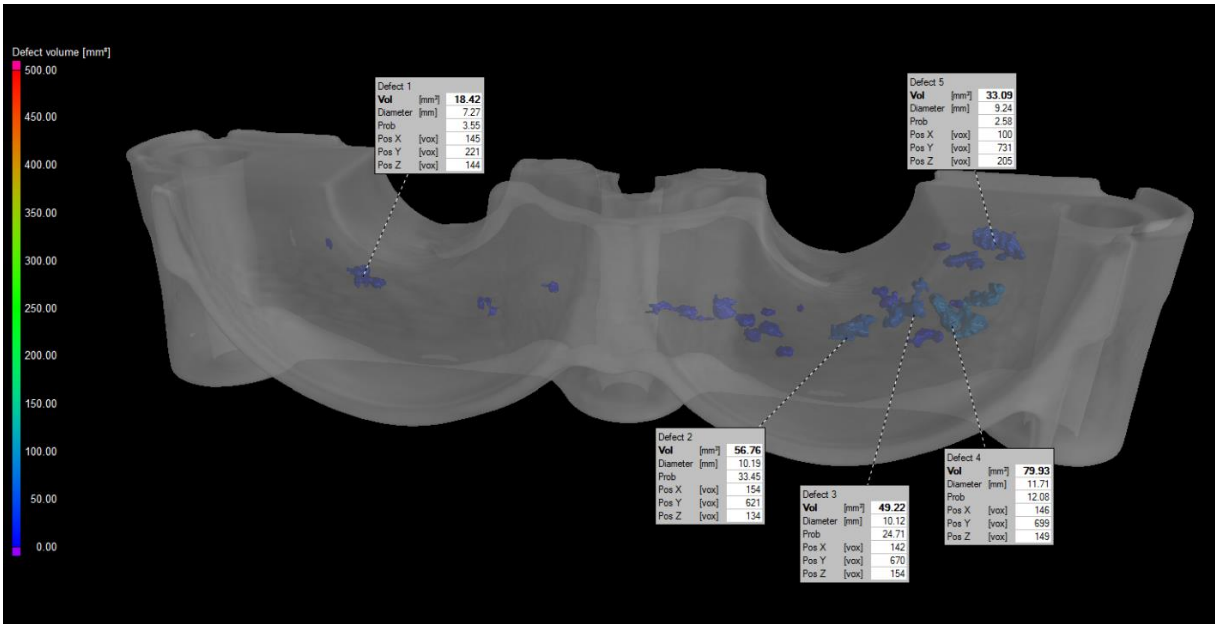


Figure 122: CT scan of porosity in casting from Chapter 3 study

The last opportunity is linked to the earlier conversation about steady state and die temperatures changing through time. The die casting cycle is not a unique, independent event. Previous cycles influence the die temperature. That is the entire reason the industry goes through a warm-up cycle and regularly scraps so many parts until the die is elevated to production temperatures. Therefore, as the understanding and data collection on temperatures improve within the industry, the use of machine learning tools like recurrent neural networks that carry information through time must be studied within die casting. This connection through time may provide the key insight into fully understanding the die casting process.

EPILOGUE

In my experience in the die casting industry, I have witnessed people always looking for the silver bullet. If I implement this piece of new technology, all my problems go away. If I use this new steel within my die, all my tool issues will end. If I collect this one more bit of data, I will prove this is the cause of all my defects. Prior to starting this work, I was convinced that it was not that simple. I used probability to reason to my position. If it is just one missing piece that solves all die castings problems, then someone, by chance, would have stumbled across it in the past 100 years. This person would have monetized the solution, and the entire industry would be performing at world-class levels.

The hard truth is there is no silver bullet in die casting. There is no “one more parameter” that one can start collecting today and use to solve all casting problems. There is no technology or material that will make the problems of die casting go away. Through this work I have learned die casting is, by definition, a complex system. Complex systems do not have simple solutions. If they did, they would not be complex.

The current die cast mentality must be replaced with a system thinking approach. Looking at the problems the industry faces with the entire system in mind changes how problems are approached. Issues would no longer be siloed and passed between departments. The system is reviewed, analyzed, prodded, tested, and changed to arrive at a solution. All areas are involved because everyone is an important piece in the overall system. Collecting the data from the system helps drive these types of system improvements, as the feedback from the system is known. This is easy to proclaim in a few statements within a paper. It is orders of magnitude harder to physically implement in a foundry.

This new approach requires different skillsets than exist today in most foundries. The analytics portion alone drives the need for data science skillsets. A recently graduated data scientist likely does not have a die cast foundry on the top of his or her future employer list. There are other issues to address as well like PLC programming, sensor connections, data storage, cyber security, considerations for technology within the manufacturing process, and systems engineering. All these require new skills not well known to foundries.

This work has also shown me that machine learning is as extremely useful as it is overhyped. I have found the challenge with machine learning does not come from the code writing and data preparation. Instead, success of machine learning comes from knowing where, when, and how to apply the different algorithms to different real-life problems. Success is also achieved by knowing where to say “no” for machine learning. The complexity of die casting, and manufacturing in general, provides a highly fascinating space to sensibly apply machine learning. I hope as an industry, die casting can achieve this.

Taking a systems approach with the addition of machine learning in die casting requires a culture change. Changing the die casting industry will take considerable time and effort. The punishment will be harsh for those that choose to stand on the sideline and wait for the silver bullet. The rewards will be significant for those foundries able and willing to make the investment.

There is no silver bullet in die casting.

References

[1] B. Andresen, *Die Casting Engineering: A Hydraulic, Thermal, and Mechanical Process*. New York: Marcel Dekker, 2005.

[2] D. Blondheim, Jr., “Artificial Intelligence, Machine Learning, and Data Analytics: Understanding the Concepts to Find Value in Die Casting Data,” presented at the NADCA Executive Conference, Clearwater, Florida, Feb. 25, 2020.

[3] J. Folk, “U.S. Aluminum Casting Industry - 2019,” *Die Casting Engineer*, vol. July 2019, pp. 16–19, 2019. [Online]. Available: https://www.diecasting.org/archive/dce/71916.pdf

[4] S. Midson, “Report on the 2014 Die Casting Benchmarking Survey Part 2 of 3: Operations,” in *Report on the 2014 Die Casting Benchmarking Survey*, North American Die Casting Association, 2014.

[5] T. Wuest, D. Weimer, C. Irgens, and K.-D. Thoben, “Machine learning in manufacturing: advantages, challenges, and applications,” *Production & Manufacturing Research*, vol. 4, no. 1, pp. 23–45, Jan. 2016, doi: 10.1080/21693277.2016.1192517.

[6] L. Baier, F. Johren, and S. Seebacher, “CHALLENGES IN THE DEPLOYMENT AND OPERATION OF MACHINE LEARNING IN PRACTICE,” in *In Proceedings of the 27th European Conference on Information Systems (ECIS)*, Stockholm & Uppsala, Sweden, Jun. 2019, p. 16. [Online]. Available: https://aisel.aisnet.org/ecis2019_rp/163

[7] N. Sun, A. Kopper, R. Karkare, R. C. Paffenroth, and D. Apelian, “Machine Learning Pathway for Harnessing Knowledge and Data in Material Processing,” *Inter Metalcast*, Sep. 2020, doi: 10.1007/s40962-020-00506-2.

[8] J. Landry, J. Maltais, J.-M. Deschênes, M. Petro, X. Godmaire, and A. Fraser, “Inline Integration of Shotblast Resistant Laser Marking in a Die Cast Cell,” Indianapolis, IN, 2018, vol. T18-123. [Online]. Available: https://www.diecasting.org/archive/transactions/T18-123.pdf

[9] D. Blondheim, Jr., “Unsupervised Machine Learning and Statistical Anomaly Detection Applied to Thermal Images,” Indianapolis, IN, Oct. 2018, vol. T18-071. [Online]. Available: http://www.diecasting.org/archive/transactions/T18-071.pdf

[10] “2018: Die Casting Award Winners,” *Die Casting Engineer*, p. 61, Sep. 2018. [Online]. Available: https://www.diecasting.org/docs/dce/issues/2018DCESep.pdf

[11] S. Alvarez, “Tesla Model Y single-piece rear casts spotted in Fremont factory,” Sep. 10, 2020. https://www.teslarati.com/tesla-model-y-unibody-casts-sighting-video/ (accessed Sep. 28, 2020).

[12] *Product Design for Die Casting E-606*, Sixth Edition., vol. E-606. North American Die Casting Association, 2009.

[13] H. H. Doehler, *Die Casting*. New York: McGraw-Hill Book Company, 1951.

[14] *NADCA Product Specification Standards for Die Casting*, 10th Edition. Arlington Heights, IL: North American Die Casting Association, 2018.

[15] E. A. Herman and J. S. Kirkman, *Designing Die Casting Dies Series - E-506*. Arlington Heights, IL: North American Die Casting Association, 2011.

[16] “Aluminum Alloys 101,” Oct. 13, 2020. https://www.mercalloy.com/aluminum-alloys-101/ (accessed Oct. 13, 2020).

[17] K. Ch. Apparao and A. K. Birru, “Optimization of Die casting process based on Taguchi approach,” *Materials Today: Proceedings*, vol. 4, no. 2, pp. 1852–1859, 2017, doi: 10.1016/j.matpr.2017.02.029.

[18] Q.-C. Hsu and A. T. Do, “Minimum Porosity Formation in Pressure Die Casting by Taguchi Method,” *Mathematical Problems in Engineering*, vol. 2013, pp. 1–9, 2013, doi: 10.1155/2013/920865.

[19] S. Balasubramaniam, S. Kannan, and R. Shivpuri, "Improving the Quality in Die Casting Production Using Statistical Analysis Procedures," presented at the 1999 NADCA World of Die Casting, Cleveland, OH, 1999. [Online]. Available: http://www.diecasting.org/archive/transactions/T99-071.pdf

[20] Q. Han, D. McClure, D. Wood, and D. Yang, "Statistical Analysis of the Effect of Operational Parameters on the Scrap Rates of Crossmember Casting," *Die Casting Engineer*, no. November 2017, pp. 38–43, 2017. [Online]. Available: http://www.diecasting.org/archive/dce/111738.pdf

[21] V. D. Tsoukalas, "Optimization of porosity formation in AlSi9Cu3 pressure die castings using genetic algorithm analysis," *Materials & Design*, vol. 29, no. 10, pp. 2027–2033, Dec. 2008, doi: 10.1016/j.matdes.2008.04.016.

[22] J. Zheng, Q. Wang, P. Zhao, and C. Wu, "Optimization of high-pressure die-casting process parameters using artificial neural network," *Int J Adv Manuf Technol*, vol. 44, no. 7–8, pp. 667–674, Oct. 2009, doi: 10.1007/s00170-008-1886-6.

[23] E. A. Herman, *Die Casting Process Control E-410*. North American Die Casting Association, 2012.

[24] I. Vuksanović Herceg, V. Kuč, V. M. Mijušković, and T. Herceg, "Challenges and Driving Forces for Industry 4.0 Implementation," *Sustainability*, vol. 12, no. 10, p. 4208, May 2020, doi: 10.3390/su12104208.

[25] A. Raj, G. Dwivedi, A. Sharma, A. B. Lopes de Sousa Jabbour, and S. Rajak, "Barriers to the adoption of industry 4.0 technologies in the manufacturing sector: An inter-country comparative perspective," *International Journal of Production Economics*, vol. 224, p. 107546, Jun. 2020, doi: 10.1016/j.ijpe.2019.107546.

[26] L. S. Dalenogare, G. B. Benitez, N. F. Ayala, and A. G. Frank, "The expected contribution of Industry 4.0 technologies for industrial performance," *International Journal of Production Economics*, vol. 204, pp. 383–394, Oct. 2018, doi: 10.1016/j.ijpe.2018.08.019.

[27] A. Mayr *et al.*, "Machine Learning in Production – Potentials, Challenges and Exemplary Applications," *Procedia CIRP*, vol. 86, pp. 49–54, 2019, doi: 10.1016/j.procir.2020.01.035.

[28] M. Bertolini, D. Mezzogori, M. Neroni, and F. Zammori, "Machine Learning for industrial applications: A comprehensive literature review," *Expert Systems with Applications*, vol. 175, p. 114820, Aug. 2021, doi: 10.1016/j.eswa.2021.114820.

[29] R. Kenett, I. Yahav, and A. Zonnenshain, "Chapter 13: Analytics as an Enabler of Advanced Manufacturing," in *Systems Engineering in the Fourth Industrial Revolution: Big Data, Novel Technologies, and Modern Systems Engineering*, First., John Wiley & Sons Inc, 2020, pp. 321–344.

[30] C. Hoyer, I. Gunawan, and C. H. Reaiche, "The Implementation of Industry 4.0 - A Systematic Literature Review of the Key Factors," *Syst Res Behav Sci*, vol. 37, no. 4, pp. 557–578, Jul. 2020, doi: 10.1002/sres.2701.

[31] S. Pfeiffer, "The Vision of 'Industrie 4.0' in the Making—a Case of Future Told, Tamed, and Traded," *Nanoethics*, vol. 11, no. 1, pp. 107–121, Apr. 2017, doi: 10.1007/s11569-016-0280-3.

[32] S. Wang, J. Wan, D. Zhang, D. Li, and C. Zhang, "Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination," *Computer Networks*, vol. 101, pp. 158–168, Jun. 2016, doi: 10.1016/j.comnet.2015.12.017.

[33] V. Roblek, M. Meško, and A. Krapež, "A Complex View of Industry 4.0," *SAGE Open*, vol. 6, no. 2, p. 215824401665398, Apr. 2016, doi: 10.1177/2158244016653987.

[34] "Industry 4.0 after the initial hype. Where manufacturers are finding value and how they can best capture it." McKinsey Digital, 2016. [Online]. Available: https://www.mckinsey.com/~/media/mckinsey/business%20functions/mckinsey%20digital/our%20insights/getting%20the%20most%20out%20of%20industry%204%200/mckinsey_industry_40_2016.ashx

[35] "German Industry 4.0 Index 2018." Staufen AG, 2018. Accessed: Aug. 21, 2021. [Online]. Available: https://www.staufen.ag/fileadmin/HQ/02-Company/05-Media/2-Studies/STAUFEN.-Study-Industry-4.0-Index-2018-Web-DE-en.pdf

[36] "German Industry 4.0 Index 2019." Staufen AG, 2019. Accessed: Aug. 21, 2021. [Online]. Available: https://www.staufen.ag/fileadmin/HQ/02-Company/05-Media/2-Studies/STAUFEN.-Study-Industry-4-0-index-2019-en_.pdf
[37] L. Bonekamp and M. Sure, "Consequences of Industry 4.0 on Human Labour and Work Organisation," *Journal of Business and Media Psychology*, vol. 6, no. 1, pp. 33–40, 2015.
[38] A. G. Frank, L. S. Dalenogare, and N. F. Ayala, "Industry 4.0 technologies: Implementation patterns in manufacturing companies," *International Journal of Production Economics*, vol. 210, pp. 15–26, Apr. 2019, doi: 10.1016/j.ijpe.2019.01.004.
[39] S. B. Gershwin, "The future of manufacturing systems engineering," *International Journal of Production Research*, vol. 56, no. 1–2, pp. 224–237, Jan. 2018, doi: 10.1080/00207543.2017.1395491.
[40] R. A. Miller, "Gate Speed, Fraction Solid, and the Effect on Mechanical Properties," Indianapolis, IN, Oct. 2018, vol. T18-042. [Online]. Available: https://www.diecasting.org/archive/transactions/T18-042.pdf
[41] H. Bramann, L. Leineweber, and J. C. Sturm, "Innovative Product Design and Robust Process Layout in Die Casting with Autonomous Engineering," Indianapolis, IN, Oct. 2018, vol. T18-093. [Online]. Available: https://www.diecasting.org/archive/transactions/T18-093.pdf
[42] M. Hartlieb and C. Raone, "Optimization of the Die Thermal Monitoring and Die Lubrication Process," Indianapolis, IN, Oct. 2018, vol. T18-023. [Online]. Available: https://www.diecasting.org/archive/transactions/T18-023.pdf
[43] M. Ognianov, "Influence of the Cooling Rate on the Microstructure and Mechanical Properties in 5wt.% Cr Martensitic Steels," Indianapolis, IN, Oct. 2018, vol. T18-013. [Online]. Available: https://www.diecasting.org/archive/transactions/T18-013.pdf
[44] R. A. Miller, "Multi-time Scale Systems and Quasi Equilibrium," 2016, vol. T16-082. [Online]. Available: https://www.diecasting.org/archive/transactions/T16-082.pdf
[45] D. Blondheim, Jr., "Initial Development of Machine Learning Algorithms to Predict Casting Defects in High-Pressure Die Casting," Atlanta, GA, Sep. 2017, vol. T17-073. [Online]. Available: http://www.diecasting.org/archive/transactions/T17-073.pdf
[46] E. Herman, "The Net Shape Paradox (Why Die Casting Sometimes Seems Mysterious)," *Die Casting Engineer*, vol. November 2018, pp. 20–22, Nov. 2018. [Online]. Available: https://www.diecasting.org/archive/dce/111820.pdf
[47] M. Mitchell, *Complexity: a guided tour*, 1. iss. as an paperback. New York, NY: Oxford University Press, 2011.
[48] N. F. Johnson, *Simply complexity: a clear guide to complexity theory*. Oneworld Publications, 2009. Accessed: Aug. 22, 2021. [Online]. Available: http://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=1792154
[49] H. G. Sillitto, "On Systems Architects and Systems Architecting: some thoughts on explaining and improving the art and science of systems architecting," *INCOSE International Symposium*, vol. 19, no. 1, pp. 970–985, Jul. 2009, doi: 10.1002/j.2334-5837.2009.tb00995.x.
[50] V. R. Vemuri, *Modeling of complex systems: an introduction*. New York: Academic Press, 1978.
[51] K. Warren, "Chaos Theory and Complexity Theory," in *Encyclopedia of Social Work*, NASW Press and Oxford University Press, 2013. doi: 10.1093/acrefore/9780199975839.013.45.
[52] F. Heylighen, "Complexity and Self-organization," in *Encyclopedia of Library and Informatoin Sciences*, Bates&Maack., Taylor & Francis, 2018.
[53] L. M. Jiji, *Heat Conduction*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. doi: 10.1007/978-3-642-01267-9.
[54] A. Kossiakoff, Ed., *Systems engineering: principles and practice*, 2nd ed. Hoboken, N.J: Wiley-Interscience, 2011.
[55] C. S. Wasson, *System engineering analysis, design, and development: concepts, principles, and practices*, Second edition. Hoboken, New Jersey: John Wiley & Sons Inc, 2016.

[56] D. D. Walden, G. J. Roedler, K. Forsberg, R. D. Hamelin, T. M. Shortell, and International Council on Systems Engineering, Eds., *Systems engineering handbook: a guide for system life cycle processes and activities*, 4th edition. Hoboken, New Jersey: Wiley, 2015.
[57] USA, Ed., *Systems engineering handbook*. Washington, D.C: National Aeronautics and Space Administration, 2007.
[58] A. Schneider, C. Wickert, and E. Marti, "Reducing Complexity by Creating Complexity: A Systems Theory Perspective on How Organizations Respond to Their Environments: Reducing Complexity by Creating Complexity," *Jour. of Manage. Stud.*, vol. 54, no. 2, pp. 182–208, Mar. 2017, doi: 10.1111/joms.12206.
[59] "Systems Engineering Definition," *Default*. https://www.incose.org/about-systems-engineering/system-and-se-definition/systems-engineering-definition (accessed Aug. 29, 2021).
[60] National ITS Architecture Team, "Systems Engineering for Intelligent Transportation Systems: An Introduction for Transportation Professionals," Department of Transportation, Office of Operations, FHWA-HOP-07-069, Jan. 2007. Accessed: Aug. 30, 2021. [Online]. Available: https://ops.fhwa.dot.gov/publications/seitsguide/
[61] M. Mennenga, F. Cerdas, S. Thiede, and C. Herrmann, "Exploring the Opportunities of System of Systems Engineering to Complement Sustainable Manufacturing and Life Cycle Engineering," *Procedia CIRP*, vol. 80, pp. 637–642, 2019, doi: 10.1016/j.procir.2019.01.026.
[62] R. Sturdivant, "Application of Systems Engineering to Complex Systems and System of Systems," Colorado State University, 2017.
[63] K. E. Boulding, "General Systems Theory - The Skeleton of Science," *Management Science*, vol. 2, no. 3, pp. 197–208, Apr. 1956.
[64] R. L. Ackoff, "Towards a System of Systems Concepts," *Management Science*, vol. 17, no. 11, pp. 661–671, Jul. 1971, doi: 10.1287/mnsc.17.11.661.
[65] *Control of complex systems: theory and applications*, 1st edition. Cambridge, MA: Elsevier, 2016.
[66] K. Åström, P. Albertos, M. Blanke, A. Isidori, W. Schaufelberger, and R. Sanz, Eds., *Control of Complex Systems*. London: Springer London, 2001. doi: 10.1007/978-1-4471-0349-3.
[67] "Mercury Marine's new high-pressure die casting machine is now the largest in North America," *Mercury Marine*. http://www.mercurymarine.com/en/us/news/mercury-marines-new-high-pressure-die-casting-machine-is-now-the-largest-in-north-america/ (accessed Aug. 31, 2021).
[68] "Tesla Giga Press Die Casting Machine - World's Largest Die Cast Machine," *Sunrise Metal - Aluminium Die Casting Expert*. https://www.sunrise-metal.com/tesla-giga-press-die-casting-machine/ (accessed Aug. 31, 2021).
[69] Idra Group, *IDRA GROUP | Flash News - Giga Press (Riccardo Ferrario)*, (Mar. 18, 2021). Accessed: Aug. 31, 2021. [Online Video]. Available: https://www.youtube.com/watch?v=o2Mi84wCRg8
[70] M. L. Cox, R. R. Fish, and J. P. Wronowicz, *PQ2 - Machine Power and Die Compatibility*. Arlington Heights, IL: North American Die Casting Association.
[71] J. Scott. Kirkman, *Thermal Design & Control of DIe Casting Dies*, Publication #415. Arlington Heights, IL: North American Die Casting Association, 2010.
[72] D. J. Armentrout, J. Brevick, and T. Altan, "Investigation and Reduction of Air Entrapment in the Shot Sleeve of a Cold Chamber Die Casting Machine (A Progress Report)," The Ohio State University, NSF Engineering Research Center for Net Shape Manufacturing, ERC/NSM-C-92-61, Oct. 1992. [Online]. Available: https://www.diecasting.org/archive/erc/ERC-92-61.pdf
[73] D. J. Armentrout, J. Brevick, and T. Altan, "Investigation of the Slow Shot Phase to Minimize Entrapped Air in Cold Chamber Die Casting," The Ohio State University, NSF Engineering Research Center for Net Shape Manufacturing, ERC/NSM-C-93-19, Jun. 1993. [Online]. Available: https://www.diecasting.org/archive/erc/ERC-93-19.pdf
[74] M. C. Thome and J. Brevick, "Optimal Slow Shot Profiles in Cold Chamber Die Casting," The Ohio State University, NSF Engineering Research Center for Net Shape Manufacturing, ERC/NSM-C-95-14, Dec. 1993. [Online]. Available: https://www.diecasting.org/archive/erc/ERC-95-14.pdf

[75] M. C. Thome, J. Brevick, and Y.-L. Chu, "Modeling the Effect of Shot Plunger Acceleration on Wave Formation and Air Entrapment in Cold Chamber DIe Casting (A Progress Report)," The Ohio State University, NSF Engineering Research Center for Net Shape Manufacturing, ERC/NSM-C-94-04, Dec. 1993. [Online]. Available: https://www.diecasting.org/archive/erc/ERC-94-04.pdf
[76] M. Thome and J. R. Brevick, "Optimal Slow Shot Velocity Profiles For Cold Chamber Die Casting," Indianapolis, IN, Oct. 1995, pp. 53–59. [Online]. Available: http://www.diecasting.org/archive/transactions/T95-024.pdf
[77] A. V. Deshpande, "Evaluation of Slow Shot Acceleration on Gas Porosity in Medium Sized Diecast Components," Master of Science, The Ohio State University, 1999. [Online]. Available: https://www.diecasting.org/archive/erc/ERC-99-01.pdf
[78] G. V. Sheffer, "Direct Recording In The Die Casting Industry," North American Die Casting Association Archive, 64–040d, 1964. [Online]. Available: http://www.diecasting.org/archive/reports/64-040d.pdf
[79] L. A. Bartling, "Recording instrumentation a valuable die casting tool," North American Die Casting Association Archive, 66–201, 1966. [Online]. Available: https://www.diecasting.org/archive/reports/66-201.pdf
[80] R. Moore, "Monitoring the Die-Casting Process," North American Die Casting Association Archive, 69–011, 1969. [Online]. Available: https://www.diecasting.org/archive/reports/69-011.pdf
[81] J. R. Vann, "A Building Block Approach to Shot End Monitoring and Control," *Die Casting Engineer*, no. September-October 1981, pp. 15–26, 1981. [Online]. Available: https://www.diecasting.org/archive/dce/98116.pdf
[82] J. R. Vann, "Automatic Shot Control with Computer Monitoring and Programming for Increased Productivity in Die Casting," *Die Casting Engineer*, no. March/April 1985, pp. 28–40, 1985. [Online]. Available: https://www.diecasting.org/archive/dce/38528.pdf
[83] L. X. Kong, F. H. She, W. M. Gao, S. Nahavandi, and P. D. Hodgson, "Integrated optimization system for high pressure die casting processes," *Journal of Materials Processing Technology*, vol. 201, no. 1–3, pp. 629–634, May 2008, doi: 10.1016/j.jmatprotec.2007.11.250.
[84] I. Pastor-Lopez, I. Santos, A. Santamaria-Ibirika, M. Salazar, J. de-la-Pena-Sordo, and P. G. Bringas, "Machine-learning-based surface defect detection and categorisation in high-precision foundry," in *2012 7th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, Singapore, Singapore, Jul. 2012, pp. 1359–1364. doi: 10.1109/ICIEA.2012.6360934.
[85] A. E. Kopper and D. Apelian, "Predicting Quality of Castings via Supervised Learning Method," *Inter Metalcast*, Apr. 2021, doi: 10.1007/s40962-021-00606-7.
[86] N. Gramegna and F. Bonollo, *Smart Control and Cognitive System applied to the HPDC Foundry 4.0*. MUSIC Consortium, 2016.
[87] K. J. Rogers, G. Savage, and B. Cooper, "Experiments With a New Generation Shot Monitoring System in High-Pressure Die Casting," *Die Casting Engineer*, no. January/February 2000, pp. 54–62, 2000. [Online]. Available: https://www.diecasting.org/archive/dce/10054.pdf
[88] G. Savage, M. Gershenzon, and K. J. Rogers, "The Role of Pressure in High Pressure Die Casting," in *NADCA Congress*, Cincinnati, OH, Oct. 2001, vol. T01-053, pp. 151–156. [Online]. Available: https://www.diecasting.org/archive/transactions/T01-053.pdf
[89] W. R. McAlister, "Continuous Process Monitoring Leads to Reduced Scrap and Shorter Cycle Times," *Die Casting Engineer*, vol. July/August 1998, pp. 14–20, 1998. [Online]. Available: https://www.diecasting.org/archive/dce/79814.pdf
[90] R. A. Miller and A. K. Monroe, "Benefits and Limitations of Time-varying Fast Shot Profiles," Atlanta, GA, Sep. 2017, vol. T17-092. [Online]. Available: https://www.diecasting.org/archive/transactions/T17-092.pdf
[91] M. Winkler, L. Kallien, and T. Feyertag, "Correlation between Process Parameters and Quality Characteristics in Aluminum High Pressure Die Casting," vol. T15-022. [Online]. Available: http://www.diecasting.org/archive/transactions/T15-022.pdf

[92] A. E. Kopper, “Knowledge Creation via Data Analytics in a High Pressure Die Casting Operation,” Worcester Polytechnic Institute, 2020. [Online]. Available: https://digital.wpi.edu/concern/etds/3484zk755
[93] D. Blondheim, Jr. and A. Monroe, “Macro Porosity Formation - A Study in High Pressure Die Casting,” *International Journal of Metalcasting*, 2021, doi: https://doi.org/10.1007/s40962-021-00602-x.
[94] J. Brevick, *Die Casting Porosity Guidebook*. North American Die Casting Association, 2008.
[95] W. G. Walkington, *Die Casting Defects: Causes and Solutions*. Rosemont, IL: North American Die Casting Association, 1997.
[96] D. Twarog, “State of the Industry 2012.” North American Die Casting Association, Feb. 2012. [Online]. Available: https://www.diecasting.org/archive/dce/212online2.pdf
[97] S. Viswanathan *et al.*, Eds., “Shrinkage Porosity and Gas Porosity,” in *Casting*, ASM International, 2008, pp. 370–374. doi: 10.31399/asm.hb.v15.a0005222.
[98] P. D. Lee, A. Chirazi, and D. See, “Modeling microporosity in aluminum–silicon alloys: a review,” *Journal of Light Metals*, vol. 1, no. 1, pp. 15–30, Feb. 2001, doi: 10.1016/S1471-5317(00)00003-1.
[99] J. Campbell, *Castings*, 2. ed. Oxford: Butterworth-Heinemann, 2003.
[100] E. Fiorese, F. Bonollo, G. Timelli, L. Arnberg, and E. Gariboldi, “New Classification of Defects and Imperfections for Aluminum Alloy Castings,” *Inter Metalcast*, vol. 9, no. 1, pp. 55–66, Jan. 2015, doi: 10.1007/BF03355602.
[101] R. Atwood, “A Combined Cellular Automata and Diffusion Model for the Prediction of Porosity Formation During Solidification,” University of London, 2001.
[102] I. Brill, B. Kappes, and S. Midson, “An Initial Evaluation of CT Scanning for Measuring and Characterizing Porosity in Aluminum Die Castings,” Indianapolis, IN, Oct. 2018, vol. T18-083. [Online]. Available: http://www.diecasting.org/archive/transactions/T18-083.pdf
[103] M. Weidt, R. A. Hardin, C. Garb, J. Rosc, R. Brunner, and C. Beckermann, “Prediction of porosity characteristics of aluminium castings based on X-ray CT measurements,” *International Journal of Cast Metals Research*, pp. 1–19, May 2018, doi: 10.1080/13640461.2018.1467105.
[104] C. Gu, Y. Lu, and A. A. Luo, “Three-dimensional visualization and quantification of microporosity in aluminum castings by X-ray micro-computed tomography,” *Journal of Materials Science & Technology*, vol. 65, pp. 99–107, Feb. 2021, doi: 10.1016/j.jmst.2020.03.088.
[105] H. Cao, M. Hao, C. Shen, and P. Liang, “The influence of different vacuum degree on the porosity and mechanical properties of aluminum die casting,” *Vacuum*, vol. 146, pp. 278–281, Dec. 2017, doi: 10.1016/j.vacuum.2017.09.048.
[106] X. P. Niu, B. H. Hu, I. Pinwill, and H. Li, “Vacuum assisted high pressure die casting of aluminium alloys,” *Journal of Materials Processing Technology*, vol. 105, no. 1–2, pp. 119–127, Sep. 2000, doi: 10.1016/S0924-0136(00)00545-8.
[107] Y. Zhang, E. Lordan, K. Dou, S. Wang, and Z. Fan, “Influence of porosity characteristics on the variability in mechanical properties of high pressure die casting (HPDC) AlSi7MgMn alloys,” *Journal of Manufacturing Processes*, vol. 56, pp. 500–509, Aug. 2020, doi: 10.1016/j.jmapro.2020.04.071.
[108] J. A. Dantzig and M. Rappaz, *Solidification*, 1 st. EPFL Press, 2009.
[109] J. Huang and J. G. Conley, “Modeling of Microporosity Evolution During Solidification Processes,” in *Review of Progress in Quantitative Nondestructive Evaluation*, D. O. Thompson and D. E. Chimenti, Eds. Boston, MA: Springer US, 1998, pp. 1839–1846. doi: 10.1007/978-1-4615-5339-7_238.
[110] T. Liang, C. Mobley, and N. Tsumagari, “The Effects of Shot Delay Time on the Microstructures and Mechanical Properties of a Die Cast Aluminum Alloy,” presented at the Die Casting Toward The Future, Rosemont, IL, Oct. 2002. [Online]. Available: https://www.diecasting.org/archive/transactions/T02-053.pdf

[111] B. Zhang, S. L. Cockcroft, D. M. Maijer, J. D. Zhu, and A. B. Phillion, “Casting defects in low-pressure die-cast aluminum alloy wheels,” *JOM*, vol. 57, no. 11, pp. 36–43, Nov. 2005, doi: 10.1007/s11837-005-0025-1.
[112] K. D. Carlson and C. Beckermann, “Prediction of Shrinkage Pore Volume Fraction Using a Dimensionless Niyama Criterion,” *Metall and Mat Trans A*, vol. 40, no. 1, pp. 163–175, Jan. 2009, doi: 10.1007/s11661-008-9715-y.
[113] G. K. Sigworth, “Shrinkage, Feeding and Riser Design,” *AFS Transactoins 2014*, vol. 14–002, pp. 25–36, 2014.
[114] M. Shabani and A. Mazahery, “Prediction of Mechanical Properties of Cast A356 Alloy as a Function of Microstructure and Cooling Rate,” *Archives of Metallurgy and Materials*, vol. 56, no. 3, Jan. 2011, doi: 10.2478/v10172-011-0073-1.
[115] M. Easton, C. Davidson, and D. St John, “Effect of Alloy Composition on the Dendrite Arm Spacing of Multicomponent Aluminum Alloys,” *Metall and Mat Trans A*, vol. 41, no. 6, pp. 1528–1538, Jun. 2010, doi: 10.1007/s11661-010-0183-9.
[116] J. Cho and C. Kim, “THE RELATIONSHIP BETWEEN DENDRITE ARM SPACING AND COOLING RATE OF Al-Si CASTING ALLOYS IN HIGH PRESSURE DIE CASTING,” *InternationalJournal ofMetalcasting*, vol. 8, no. 1, pp. 49–55, 2014, doi: https://doi.org/10.1007/BF03355571.
[117] “SRE MAX,” *Bosello High Technology , a ZEISS company*. https://bosello.eu/products/sre-max/ (accessed Dec. 29, 2020).
[118] “Phoenix Vtomex C | 3D CT Scanner (Mini Focus),” *Waygate Technologies*. https://www.bakerhughesds.com/industrial-x-ray-ct-scanners/phoenix-vtomex-c-ct (accessed Dec. 29, 2020).
[119] “Xradia 610 & 620 Versa.” https://www.zeiss.com/microscopy/us/products/x-ray-microscopy/zeiss-xradia-610-and-620-versa.html (accessed Dec. 29, 2020).
[120] T. J. Schorn, *Improving the Effectiveness of Visual Inspection*. Schaumburg, IL USA: American Foundry Society, 2018.
[121] J. F. Koretz and G. H. Handelman, “How the Human Eye Focuses,” *Sci Am*, vol. 259, no. 1, pp. 92–99, Jul. 1988, doi: 10.1038/scientificamerican0788-92.
[122] J. Schindelin *et al.*, “Fiji: an open-source platform for biological-image analysis,” *Nat Methods*, vol. 9, no. 7, pp. 676–682, Jul. 2012, doi: 10.1038/nmeth.2019.
[123] S. Preibisch, S. Saalfeld, J. Schindelin, and P. Tomancak, “Software for bead-based registration of selective plane illumination microscopy data,” *Nat Methods*, vol. 7, no. 6, pp. 418–419, Jun. 2010, doi: 10.1038/nmeth0610-418.
[124] “Standard Reference Radiographs for Inspection of Aluminum and Magnesium Die Castings E505-15.” ASTM International, 2015. doi: 10.1520/E0505-15.
[125] S. S. Shapiro and M. B. Wilk, “An Analysis of Variance Test for Normality (Complete Samples),” *Biometrika*, vol. Vol. 52, pp. 591–611, Dec. 1965, [Online]. Available: https://www.jstor.org/stable/2333709
[126] F. Wilcoxon, “Individual Comparisons by Ranking Methods,” *Biometrics Bulletin*, vol. 1, no. No. 6, pp. 80–83, Dec. 1945.
[127] *MAGMAsoft*. Kackerstrasse 11, 52072 Aachen, Germany: MAGMA Gmbh, 2019.
[128] M. R. Brand, “An examination of certain Bayesian methods used in reliability analysis,” *Reliability Engineering*, vol. 1, no. 2, pp. 115–125, Oct. 1980, doi: 10.1016/0143-8174(80)90003-7.
[129] S. G. Lee and A. M. Gokhale, “Formation of gas induced shrinkage porosity in Mg-alloy high-pressure die-castings,” *Scripta Materialia*, vol. 55, no. 4, pp. 387–390, Aug. 2006, doi: 10.1016/j.scriptamat.2006.04.040.
[130] F. Bonollo, N. Gramegna, and G. Timelli, “High-Pressure Die-Casting: Contradictions and Challenges,” *JOM*, vol. 67, no. 5, pp. 901–908, May 2015, doi: 10.1007/s11837-015-1333-8.

[131] C. H. Cáceres and B. I. Selling, "Casting defects and the tensile properties of an AlSiMg alloy," *Materials Science and Engineering: A*, vol. 220, no. 1–2, pp. 109–116, Dec. 1996, doi: 10.1016/S0921-5093(96)10433-0.
[132] R. Lumley, N. Deeva, and M. Gershenzon, "An Evaluation of Quality Parameters for High Pressure Die Castings," *Inter Metalcast*, vol. 5, no. 3, pp. 37–56, Jul. 2011, doi: 10.1007/BF03355517.
[133] D. Blondheim, Jr., "Improving Manufacturing Applications of Machine Learning by Understanding Defect Classification and the Critical Error Threshold," *Inter Metalcast*, Jun. 2021, doi: 10.1007/s40962-021-00637-0.
[134] Y. LeCun *et al.*, "Handwritten Digit Recognition with a Back-Propagation Network," 1990, pp. 396–404.
[135] S. Lalmuanawma, J. Hussain, and L. Chhakchhuak, "Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review," *Chaos, Solitons & Fractals*, vol. 139, p. 110059, Oct. 2020, doi: 10.1016/j.chaos.2020.110059.
[136] "The transformative potential of AI in the manufacturing industry - Microsoft Industry Blogs." https://cloudblogs.microsoft.com/industry-blog/manufacturing/2017/09/25/the-transformative-potential-of-ai-in-the-manufacturing-industry/ (accessed Nov. 27, 2020).
[137] "Bringing Intelligence to Industrial Manufacturing Through AWS IoT and Machine Learning | AWS Partner Network (APN) Blog." https://aws.amazon.com/blogs/apn/bringing-intelligence-to-industrial-manufacturing-through-aws-iot-and-machine-learning/ (accessed Nov. 27, 2020).
[138] P. K. Tseng, "Smart Manufacturing Becomes Key to Enterprise Survival in Post-Pandemic Era with Forecasted Revenue of US$400 Billion in 2024, Says TrendForce," Aug. 17, 2020. https://www.trendforce.com/presscenter/news/20200817-10437.html (accessed Nov. 27, 2020).
[139] "Thriving in an AI World." KPMG, Apr. 15, 2021. [Online]. Available: https://info.kpmg.us/content/dam/info/en/news-perspectives/pdf/2021/Updated%204.15.21%20-%20Thriving%20in%20an%20AI%20world.pdf
[140] G. Rebala, A. Ravi, and S. Churiwala, *An Introduction to Machine Learning*. Cham: Springer International Publishing, 2019. doi: 10.1007/978-3-030-15729-6.
[141] R. Fernandes de Mello and M. Antonelli Ponti, *Machine Learning: A Practical Approach on the Statistical Learning Theory*. Cham: Springer International Publishing, 2018. doi: 10.1007/978-3-319-94989-5.
[142] J. M. Juran, Ed., *Juran's quality handbook*, 5. ed. New York, NY: McGraw-Hill, 1999.
[143] J. J. Sheehan and C. G. Drury, "The analysis of industrial inspection," *Applied Ergonomics*, vol. 2, no. 2, pp. 74–78, Jun. 1971, doi: 10.1016/0003-6870(71)90073-1.
[144] C. G. Drury, "The effect of speed of working on industrial inspection accuracy," *Applied Ergonomics*, vol. 4, no. 1, pp. 2–7, Mar. 1973, doi: 10.1016/0003-6870(73)90002-1.
[145] C. G. Drury and M.-J. Wang, "Are Research Results in Inspection Task Specific?," *Proceedings of the Human Factors Society Annual Meeting*, vol. 30, no. 5, pp. 476–480, Sep. 1986, doi: 10.1177/154193128603000515.
[146] C.-F. Chi, "The human as an optimizer in inspection," State University of New York, Buffalo, NY, 1990.
[147] A. Fernández, S. García, and F. Herrera, "Addressing the Classification with Imbalanced Data: Open Problems and New Challenges on Class Distribution," in *Hybrid Artificial Intelligent Systems*, vol. 6678, E. Corchado, M. Kurzyński, and M. Woźniak, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 1–10. doi: 10.1007/978-3-642-21219-2_1.
[148] S. Viswanathan *et al.*, Eds., "High-Pressure Die Casting," in *Casting*, ASM International, 2008, pp. 715–718. doi: 10.31399/asm.hb.v15.a0005266.
[149] J. Folk, "The U.S. Aluminum Casting Industry," *Die Casting Engineer*, vol. May 2015, pp. 24–33, May 2015. [Online]. Available: https://www.diecasting.org/archive/dce/51524.pdf
[150] V. Khalajzadeh, K. D. Carlson, D. G. Backman, and C. Beckermann, "A Pore-Centric Model for Combined Shrinkage and Gas Porosity in Alloy Solidification," *Metall and Mat Trans A*, vol. 48, no. 4, pp. 1797–1816, Apr. 2017, doi: 10.1007/s11661-016-3940-6.

[151] P. F. Ostwald and J. Huang, "A Method for Optimal Tolerance Selection," *Journal of Engineering for Industry*, vol. 99, no. 3, pp. 558–565, Aug. 1977, doi: 10.1115/1.3439279.
[152] C. Zhang, H. P. Wang, and J. K. Li, "Simultaneous Optimization of Design and Manufacturing — Tolerances with Process (Machine) Selection," *CIRP Annals*, vol. 41, no. 1, pp. 569–572, 1992, doi: 10.1016/S0007-8506(07)61270-0.
[153] A. Jeang, "Tolerance chart optimization for quality and cost," *International Journal of Production Research*, vol. 36, no. 11, pp. 2969–2983, Nov. 1998, doi: 10.1080/002075498192238.
[154] M. D. Al-Ansary and I. M. Deiab, "Concurrent optimization of design and machining tolerances using the genetic algorithms method," *International Journal of Machine Tools and Manufacture*, vol. 37, no. 12, pp. 1721–1731, Dec. 1997, doi: 10.1016/S0890-6955(97)00033-3.
[155] S. A. Irani, R. O. Mittal, and E. A. Lehtihet, "Tolerance chart optimization," *International Journal of Production Research*, vol. 27, no. 9, pp. 1531–1552, Sep. 1989, doi: 10.1080/00207548908942638.
[156] A. N. Haq, K. Sivakumar, R. Saravanan, and V. Muthiah, "Tolerance design optimization of machine elements using genetic algorithm," *Int J Adv Manuf Technol*, vol. 25, no. 3–4, pp. 385–391, Feb. 2005, doi: 10.1007/s00170-003-1855-z.
[157] American Society of Mechanical Engineers and American National Standards Institute, *Dimensioning and tolerancing: engineering drawing and related documentation practices*. New York: American Society of Mechanical Engineers, 2019.
[158] J. D. Meadows, *Geometric Dimensioning and Tolerancing: Applications, Analysis & Measurement [per ASME Y14.5-2009]*. James D. Meadows & Associated, Inc., 2009.
[159] D. Steiner and R. Katz, "Measurement Techniques for the Inspection of Porosity Flaws on Machined Surfaces," *Journal of Computing and Information Science in Engineering*, vol. 7, no. 1, pp. 85–94, Mar. 2007, doi: 10.1115/1.2424244.
[160] S. J. Świłło and M. Perzyk, "Automatic inspection of surface defects in die castings after machining," *Archives of Foundry Engineering*, vol. Volume 11, no. Special Issue 3/2011, pp. 231–236, 2011.
[161] F. Peters, R. Stone, K. Watts, P. Zhon, and A. Clemons, "Visual Inspection of Casting Surfaces," *Transactions of the American Foundry Society*, pp. 45–52, 2013.
[162] D. H. Harris and F. B. Chaney, *Human Factors in Quality Assurance*. New York: John Wiley & Sons Inc, 1969.
[163] S. B. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques," *Informatica*, vol. 31, no. 3, pp. 249–269, Oct. 2007.
[164] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. New York, NY: Springer New York, 2009. doi: 10.1007/978-0-387-84858-7.
[165] Y. Wang, D. Zou, J. Yi, J. Bailey, X. Ma, and Q. Gu, "IMPROVING ADVERSARIAL ROBUSTNESS REQUIRES REVISITING MISCLASSIFIED EXAMPLES," Virtual Conference, 2020, p. 14. Accessed: Mar. 01, 2021. [Online]. Available: https://openreview.net/forum?id=rklOg6EFwS
[166] M. R. Smith and T. Martinez, "Improving classification accuracy by identifying and removing instances that should be misclassified," in *The 2011 International Joint Conference on Neural Networks*, San Jose, CA, USA, Jul. 2011, pp. 2690–2697. doi: 10.1109/IJCNN.2011.6033571.
[167] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, p. 6, Dec. 2020, doi: 10.1186/s12864-019-6413-7.
[168] D. Powers, "Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation," *Mach. Learn. Technol.*, vol. 2, p. 25, 2008, [Online]. Available: https://arxiv.org/abs/1504.00854
[169] D. Blondheim, Jr. and S. Bhowmik, "Time-Series Analysis and Anomaly Detection of High-Pressure Die Casting Shot Profiles," *Die Casting Engineer*, pp. 14–18, Nov. 2019. [Online]. Available: http://www.diecasting.org/archive/dce/111914.pdf

[170] P. Patel, M. I. Ali, and A. Sheth, "From Raw Data to Smart Manufacturing: AI and Semantic Web of Things for Industry 4.0," *IEEE Intell. Syst.*, vol. 33, no. 4, pp. 79–86, Jul. 2018, doi: 10.1109/MIS.2018.043741325.

[171] R. S. Peres, X. Jia, J. Lee, K. Sun, A. W. Colombo, and J. Barata, "Industrial Artificial Intelligence in Industry 4.0 - Systematic Review, Challenges and Outlook," *IEEE Access*, vol. 8, pp. 220121–220139, 2020, doi: 10.1109/ACCESS.2020.3042874.

[172] A. Bécue, I. Praça, and J. Gama, "Artificial intelligence, cyber-threats and Industry 4.0: challenges and opportunities," *Artif Intell Rev*, vol. 54, no. 5, pp. 3849–3886, Jun. 2021, doi: 10.1007/s10462-020-09942-2.

[173] A. B. Simmons and S. G. Chappell, "Artificial intelligence-definition and practice," *IEEE J. Oceanic Eng.*, vol. 13, no. 2, pp. 14–42, Apr. 1988, doi: 10.1109/48.551.

[174] S. J. Russell, P. Norvig, and E. Davis, *Artificial intelligence: a modern approach*, 3rd ed. Upper Saddle River: Prentice Hall, 2010.

[175] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. Cambridge, Massachusetts: The MIT Press, 2016.

[176] T. Sasakawa, J. Hu, and K. Hirasawa, "A brainlike learning system with supervised, unsupervised, and reinforcement learning," *Elect. Eng. Jpn.*, vol. 162, no. 1, pp. 32–39, Jan. 2008, doi: 10.1002/eej.20600.

[177] H. Oliff, Y. Liu, M. Kumar, M. Williams, and M. Ryan, "Reinforcement learning for facilitating human-robot-interaction in manufacturing," *Journal of Manufacturing Systems*, vol. 56, pp. 326–340, Jul. 2020, doi: 10.1016/j.jmsy.2020.06.018.

[178] L. Wang *et al.*, "Dynamic job-shop scheduling in smart manufacturing using deep reinforcement learning," *Computer Networks*, vol. 190, p. 107969, May 2021, doi: 10.1016/j.comnet.2021.107969.

[179] D. Blondheim, Jr., "Utilizing Machine Learning Autoencoders to Detect Anomalies in Time-Series Data," presented at the 2021 NADCA Congress and Exposition, Indianapolis, IN, Oct. 2021.

[180] S. Žapčević and P. Butala, "Adaptive process control based on a self-learning mechanism in autonomous manufacturing systems," *Int J Adv Manuf Technol*, vol. 66, no. 9–12, pp. 1725–1743, Jun. 2013, doi: 10.1007/s00170-012-4453-0.

[181] S. Profanter, A. Tekat, K. Dorofeev, M. Rickert, and A. Knoll, "OPC UA versus ROS, DDS, and MQTT: Performance Evaluation of Industry 4.0 Protocols," in *2019 IEEE International Conference on Industrial Technology (ICIT)*, Melbourne, Australia, Feb. 2019, pp. 955–962. doi: 10.1109/ICIT.2019.8755050.

[182] M. Silveira Rocha, G. Serpa Sestito, A. Luis Dias, A. Celso Turcato, and D. Brandao, "Performance Comparison Between OPC UA and MQTT for Data Exchange," in *2018 Workshop on Metrology for Industry 4.0 and IoT*, Brescia, Apr. 2018, pp. 175–179. doi: 10.1109/METROI4.2018.8428342.

[183] G. Büchi, M. Cugno, and R. Castagnoli, "Smart factory performance and Industry 4.0," *Technological Forecasting and Social Change*, vol. 150, p. 119790, Jan. 2020, doi: 10.1016/j.techfore.2019.119790.

[184] Y.-H. Choi and S.-H. Choi, "A Study of Crossing the Chasm in applying Smart Factory System for SMEs," *International Journal of Pure and Applied Mathematics*, vol. 118, no. No 19, pp. 469–487, Jan. 2018, [Online]. Available: http://www.ijpam.eu

[185] M. Khan, X. Wu, X. Xu, and W. Dou, "Big data challenges and opportunities in the hype of Industry 4.0," in *2017 IEEE International Conference on Communications (ICC)*, Paris, France, May 2017, pp. 1–6. doi: 10.1109/ICC.2017.7996801.

[186] A. Khan and K. Turowski, "A Survey of Current Challenges in Manufacturing Industry and Preparation for Industry 4.0," in *Proceedings of the First International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'16)*, vol. 450, A. Abraham, S. Kovalev, V. Tarassov, and V. Snášel, Eds. Cham: Springer International Publishing, 2016, pp. 15–26. doi: 10.1007/978-3-319-33609-1_2.

[187] T. Wuest, C. Irgens, and K.-D. Thoben, “An approach to monitoring quality in manufacturing using supervised machine learning on product state data,” *J Intell Manuf*, vol. 25, pp. 1167–1180, 2014, doi: 10.1007/s10845-013-0761-y.
[188] S. Sun, X. Zheng, J. Villalba-Díez, and J. Ordieres-Meré, “Data Handling in Industry 4.0: Interoperability Based on Distributed Ledger Technology,” *Sensors*, vol. 20, no. 11, p. 3046, May 2020, doi: 10.3390/s20113046.
[189] R. Sahal, J. G. Breslin, and M. I. Ali, “Big data and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case,” *Journal of Manufacturing Systems*, vol. 54, pp. 138–151, Jan. 2020, doi: 10.1016/j.jmsy.2019.11.004.
[190] K. Villalobos, V. J. Ramírez-Durán, B. Diez, J. M. Blanco, A. Goñi, and A. Illarramendi, “A three level hierarchical architecture for an efficient storage of industry 4.0 data,” *Computers in Industry*, vol. 121, p. 103257, Oct. 2020, doi: 10.1016/j.compind.2020.103257.
[191] F. Y. Kuo and I. H. Sloan, “Lifting the Curse of Dimensionality,” *Notices of the AMS*, vol. 52, no. 11, pp. 1320–1329, 2005.
[192] G. Hughes, “On the mean accuracy of statistical pattern recognizers,” *IEEE Trans. Inform. Theory*, vol. 14, no. 1, pp. 55–63, Jan. 1968, doi: 10.1109/TIT.1968.1054102.
[193] P. Domingos, “A few useful things to know about machine learning,” *Commun. ACM*, vol. 55, no. 10, pp. 78–87, Oct. 2012, doi: 10.1145/2347736.2347755.
[194] F. Provost, “Machine Learning from Imbalanced Data Sets 101,” in *AAAI Technical Report WS-00-05*, 2000, p. 3. [Online]. Available: https://www.aaai.org/Papers/Workshops/2000/WS-00-05/WS00-05-001.pdf
[195] M. Kuhn and K. Johnson, *Applied Predictive Modeling*. New York, NY: Springer New York, 2013. doi: 10.1007/978-1-4614-6849-3.
[196] A. Paleyes, R.-G. Urma, and N. D. Lawrence, “Challenges in Deploying Machine Learning: a Survey of Case Studies,” *arXiv:2011.09926 [cs]*, Jan. 2021, Accessed: Feb. 07, 2021. [Online]. Available: http://arxiv.org/abs/2011.09926
[197] L. Breiman, “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
[198] S. Kuutti, R. Bowden, Y. Jin, P. Barber, and S. Fallah, “A Survey of Deep Learning Applications to Autonomous Vehicle Control,” *IEEE Trans. Intell. Transport. Syst.*, vol. 22, no. 2, pp. 712–733, Feb. 2021, doi: 10.1109/TITS.2019.2962338.
[199] A. Kuusk and J. Gao, “Factors for successfully integrating operational and information technologies,” in *2015 Portland International Conference on Management of Engineering and Technology (PICMET)*, Portland, OR, USA, Aug. 2015, pp. 1513–1523. doi: 10.1109/PICMET.2015.7273136.
[200] E. J. M. Colbert and A. Kott, Eds., *Cyber-security of SCADA and Other Industrial Control Systems*, vol. 66. Cham: Springer International Publishing, 2016. doi: 10.1007/978-3-319-32125-7.
[201] L. Thames and D. Schaefer, Eds., *Cybersecurity for Industry 4.0: Analysis for Design and Manufacturing*. Cham: Springer International Publishing, 2017. doi: 10.1007/978-3-319-50660-9.
[202] G. Culot, F. Fattori, M. Podrecca, and M. Sartor, “Addressing Industry 4.0 Cybersecurity Challenges,” *IEEE Eng. Manag. Rev.*, vol. 47, no. 3, pp. 79–86, Sep. 2019, doi: 10.1109/EMR.2019.2927559.
[203] M. Lezzi, M. Lazoi, and A. Corallo, “Cybersecurity for Industry 4.0 in the current literature: A reference framework,” *Computers in Industry*, vol. 103, pp. 97–110, Dec. 2018, doi: 10.1016/j.compind.2018.09.004.
[204] P. O’Donovan, K. Leahy, K. Bruton, and D. T. J. O’Sullivan, “Big data in manufacturing: a systematic mapping study,” *Journal of Big Data*, vol. 2, no. 1, p. 20, Dec. 2015, doi: 10.1186/s40537-015-0028-x.
[205] M. Fahmideh and G. Beydoun, “Big data analytics architecture design—An application in manufacturing systems,” *Computers & Industrial Engineering*, vol. 128, pp. 948–963, Feb. 2019, doi: 10.1016/j.cie.2018.08.004.

[206] I. Nunes and D. Jannach, “A systematic review and taxonomy of explanations in decision support and recommender systems,” *User Model User-Adap Inter*, vol. 27, no. 3–5, pp. 393–444, Dec. 2017, doi: 10.1007/s11257-017-9195-0.
[207] C. Rudin and K. L. Wagstaff, “Machine learning for science and society,” *Mach Learn*, vol. 95, no. 1, pp. 1–9, Apr. 2014, doi: 10.1007/s10994-013-5425-9.
[208] N. Werts and M. Adya, “Data Mining in Healthcare: Issues and a Research Agenda,” in *AMCIS 2000 Proceedings*, 2000, pp. 94–97.
[209] R. Boutaba *et al.*, “A comprehensive survey on machine learning for networking: evolution, applications and research opportunities,” *J Internet Serv Appl*, vol. 9, no. 1, p. 16, Dec. 2018, doi: 10.1186/s13174-018-0087-2.
[210] “NADCA - Schedule of Events.” https://www.diecasting.org/wcm/Meetings/Congress_and_Exposition/Schedule_of_Events/wcm/Meetings/Schedule_of_Events.aspx?hkey=03f18550-584d-4f4f-9e69-df3668429226 (accessed Sep. 22, 2021).
[211] “AIM 2022: Home.” https://www.tms.org/portal/MEETINGS___EVENTS/TMS_Meetings___Events/Upcoming_TMS_Meetings/AIM2022/portal/Meetings___Events/2022/AIM2022/default.aspx?hkey=7050d00c-58bb-4f5c-854b-84486f96c37b (accessed Sep. 22, 2021).
[212] “AI Course: Home.” https://www.tms.org/portal/Meetings___Events/2021/AIcourse/default.aspx (accessed Sep. 22, 2021).
[213] F. von Hayek, “Friedrich von Hayek Prize Lecture,” *NobelPrize.org*, Dec. 11, 1974. https://www.nobelprize.org/prizes/economic-sciences/1974/hayek/lecture/ (accessed Aug. 22, 2021).
[214] B. Marr, “A Short History of Machine Learning -- Every Manager Should Read,” *Forbes*, Feb. 19, 2016. https://www.forbes.com/sites/bernardmarr/2016/02/19/a-short-history-of-machine-learning-every-manager-should-read/
[215] R Core Team, *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2016. [Online]. Available: https://www.R-project.org
[216] B. Lantz, *Machine Learning with R*, 2nd edition. Packt Publishing Ltd., 2015.
[217] E. Siegel, *Predictive Analytic: the power to predict who will click, buy, lie or die*. John Wiley & Sons Inc, 2013.
[218] M. Kuhn, *caret: Classification and Regression Training*. 2017. [Online]. Available: https://CRAN.R-project.org/package=caret
[219] W. N. Venables and B. D. Ripley, *Modern Applied Statistics with S*. New York, NY: Springer New York, 2002. doi: 10.1007/978-0-387-21706-2.
[220] S. Fritsch and F. Guenther, *neuralnet: Training of Neural Networks*. 2017. [Online]. Available: https://CRAN.R-project.org/package=neuralnet
[221] D. Meyer, E. Dimitriadou, K. Hornik, A. Weingessel, and F. Leisch, *e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien*. 2017. [Online]. Available: https://CRAN.R-project.org/package=e1071
[222] J. Macqueen, “SOME METHODS FOR CLASSIFICATION AND ANALYSIS OF MULTIVARIATE OBSERVATIONS,” in *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1967, pp. 281–297.
[223] P. Catterina and C. Raone, “Towards the ‘Total’ certification in diecasting process Thermographic monitoring and traceability with TOTAL THERMAL VISION (TTV),” 2013, vol. T13-041. [Online]. Available: https://www.diecasting.org/archive/transactions/T13-041.pdf
[224] L. X. Kong, F. H. She, S. Nahavandi, and L. H. Wang, “Thermal control with image processing and fuzzy controllers,” in *Proceedings Joint 9th IFSA World Congress and 20th NAFIPS International Conference (Cat. No. 01TH8569)*, Vancouver, BC, Canada, 2001, vol. 3, pp. 1443–1449. doi: 10.1109/NAFIPS.2001.943761.

[225] S. Simske, *Meta-Algorithmics: Patterns for Robust, Low Cost, High Quality Systems*. Wiley-IEEE Press, 2013.
[226] W. A. Shewhart, *Economic Control of Quality of Manufactured Products*. New York, NY: D. Van Nostrand Company, Inc., 1931.
[227] D. L. Simon and J. C. Lafeber, "Optimization of Die Casting Processes Using Infrared Thermography," 2011, vol. T11-071. [Online]. Available: https://www.diecasting.org/archive/transactions/T11-071.pdf
[228] A. Reikher, H. Gerber, and A. Starobin, "Multi-Stage Plunger Deceleration System," Atlanta, GA, May 2008, vol. T08-011. [Online]. Available: https://www.diecasting.org/archive/transactions/T08-011.pdf
[229] J. Liang and S. Parthasarathy, "Robust Contextual Outlier Detection: Where Context Meets Sparsity," in *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, Indianapolis Indiana USA, Oct. 2016, pp. 2167–2172. doi: 10.1145/2983323.2983660.
[230] S. Tartaglia and J. Vandertouw, "Benchmarking Study Comparison of North American and Japanese Companies," *Die Casting Engineer*, no. March/April 2000, pp. 104–109, 2000. [Online]. Available: https://www.diecasting.org/archive/dce/200104.pdf
[231] K.-D. Thoben, S. Wiesner, and T. Wuest, "'Industrie 4.0' and Smart Manufacturing – A Review of Research Issues and Application Examples," *Int. J. Automation Technol.*, vol. 11, no. 1, pp. 4–16, Jan. 2017, doi: 10.20965/ijat.2017.p0004.
[232] S. P. Udvardy, "2019 State of the Die Casting Industry," *Die Casting Engineer*, no. January 2020, pp. 10–15, 2020. [Online]. Available: https://www.diecasting.org/archive/dce/12010.pdf
[233] S. P. Udvardy, "2020 State of the Die Casting Industry," *Die Casting Engineer*, vol. January 2021, pp. 8–13, 2021. [Online]. Available: https://www.diecasting.org/docs/information/stateoftheindustry-dcejan2021.pdf
[234] A. Heidarian and M. J. Dinneen, "A Hybrid Geometric Approach for Measuring Similarity Level Among Documents and Document Clustering," in *2016 IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService)*, Oxford, United Kingdom, Mar. 2016, pp. 142–151. doi: 10.1109/BigDataService.2016.14.
[235] S. Samarasinghe, *Neural networks for applied sciences and engineering: from fundamentals to complex pattern recognition*. Auerbach Publications, 2007.
[236] D. Forsyth, *Applied Machine Learning*. Cham: Springer International Publishing, 2019. doi: 10.1007/978-3-030-18114-7.
[237] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, pp. 11–26, Apr. 2017, doi: 10.1016/j.neucom.2016.12.038.
[238] G. E. Hinton, "Reducing the Dimensionality of Data with Neural Networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006, doi: 10.1126/science.1127647.
[239] Y. LeCun, C. Cortes, and C. J. C. Burges, "The MNIST Database of handwritten digits." http://yann.lecun.com/exdb/mnist/ (accessed Jun. 25, 2021).
[240] "2.1 Million Manufacturing Jobs Could Go Unfilled by 2030," *NAM*, May 04, 2021. https://www.nam.org/2-1-million-manufacturing-jobs-could-go-unfilled-by-2030-13743/ (accessed Sep. 04, 2021).
[241] P. Wellener, V. Reyes, H. Ashton, and C. Moutray, "Creating pathways for tomorrow's workforce today," *Deloitte Insights*. https://www2.deloitte.com/us/en/insights/industry/manufacturing/manufacturing-industry-diversity.html (accessed Sep. 04, 2021).

List of Abbreviations

AFS – American Foundry Society

AI – Artificial Intelligence

BoB – Best of Best

CET – Critical Error Threshold

CT Scanning – Computed Tomography Scanning

HPDC – High Pressure Die Casting

IIoT – Industrial Internet of Things

IT – Information Technology

INCOSE – International Council on Systems Engineering

ML – Machine Learning

MQTT – Message Queuing Telemetry Transport

NADCA – North American Die Casting Association

NN – Neural Network

OPC UA – Open Platform Communications Unified Architecture

OT – Operational Technology

PLC – Programmable Logic Controller

PSI – Pounds per Square Inch

RPM – Revolutions per Minute

SE – Systems Engineering

SoS – Systems-of-Systems

SVM – Support Vector Machine

WoW – Worst of worst