7-26-2018

A probabilistic model to estimate visual inspection error for metalcastings given different training and judgment types, environmental and human factors, and percent of defects

Michelle M. Stallard-Voelker
Iowa State University

Cameron A. MacKenzie
Iowa State University, camacken@iastate.edu

Frank E. Peters
Iowa State University, fpeters@iastate.edu

A probabilistic model to estimate visual inspection error for metalcastings given different training and judgment types, environmental and human factors, and percent of defects

Abstract

Current methods for visual inspection of cast metal surfaces are variable in both terms of repeatability and reproducibility. Because of this variation in the inspection methods, extra finishing operations are often prescribed; much of this is over processing in attempt to avoid rework or customer rejection. Additionally, defective castings may pass inspection and be delivered to the customer. Given the importance of ensuring that customers receive high-quality castings, this article analyzes and quantifies the probability of Type I and II errors, where a Type I error is a false alarm, and a Type II error misses a present defect. A probabilistic model frequently used in risk analysis, called an influence diagram, is developed to incorporate different factors impacting the chances of Type I and II errors. These factors include: training for inspectors, the type of judgment used during the inspection process, the percentage of defective castings, environmental conditions, and the inspectors' capabilities. The model is populated with inputs based on prior experimentation and the authors' expertise. The influence diagram calculates the probability of a Type I error at 0.35 and the probability of a Type II error at 0.40. These results are compared to a naïve Bayes model. A manufacturer can use this analysis to identify factors in its foundry that could reduce the probability of errors. Even under the best-case scenario, the probability of Type I error is 0.18 and the probability of Type II error is 0.30 for visual inspection. This indicates improvements to the inspection process for cast metal surfaces is required.

Keywords

Risk assessment, Cast surfaces, Visual inspection, Influence diagrams, Surface inspection

Disciplines

Industrial Engineering | Systems Engineering

Comments

This article is published as Michelle M.Stallard (Voelker)Cameron A.MacKenzieFrank E.Peters. A probabilistic model to estimate visual inspection error for metalcastings given different training and judgment types, environmental and human factors, and percent of defects Author links open overlay panel. *Journal of Manufacturing Systems*, 2018 48(Part A); 97-106. DOI: 10.1016/j.jmsy.2018.07.002. Posted with permission.

A probabilistic model to estimate visual inspection error for metalcastings given different training and judgment types, environmental and human factors, and percent of defects

Michelle M. Stallard (Voelker)[1], Cameron A. MacKenzie[2], and Frank E. Peters[3]

Industrial and Manufacturing Systems Engineering at Iowa State University

Ames, Iowa, 50011, United States

[1] 0000-0002-3521-694X
[2] 0000-0002-4851-8509
[3] 0000-0002-8998-0062

ABSTRACT

Current methods for visual inspection of cast metal surfaces are variable in both terms of repeatability and reproducibility. Because of this variation in the inspection methods, extra finishing operations are often prescribed; much of this is over processing in attempt to avoid rework or customer rejection. Additionally, defective castings may pass inspection and be delivered to the customer. Given the importance of ensuring that customers receive high-quality castings, this article analyzes and quantifies the probability of Type I and II errors, where a Type I error is a false alarm, and a Type II error misses a present defect. A probabilistic model frequently used in risk analysis, called an influence diagram, is developed to incorporate different factors impacting the chances of Type I and II errors. These factors include: training for inspectors, the type of judgment used during the inspection process, the percentage of defective castings, environmental conditions, and the inspectors' capabilities. The model is populated with inputs based on prior experimentation and the authors' expertise. The influence diagram calculates the probability of a Type I error at 0.35 and the probability of a Type II error at 0.40. These results are

compared to a naïve Bayes model. A manufacturer can use this analysis to identify factors in its foundry that could reduce the probability of errors. Even under the best-case scenario, the probability of Type I error is 0.18 and the probability of Type II error is 0.30 for visual inspection. This indicates improvements to the inspection process for cast metal surfaces is required.

KEYWORDS: risk assessment; cast surfaces; visual inspection; influence diagrams; surface inspection

I. INTRODUCTION

Inspecting parts to meet quality standards is important for meeting customer needs. In metal casting, current standards use qualitative methods to determine acceptability of surface quality. The inspection process involves one or more trained operators to visually examine the surface to determine if the part is acceptable. Variation exists among interpretation of the standard not only in relation to the repeatability and reproducibility of the inspection process, but also in regards to interpretations between the manufacturer and the customer. The variability in the casting process itself is often less than that of the visual inspection process [1]. This stack-up in variation results in inconsistencies in acceptance criteria and increases the occurrence of Type I and II errors. A Type I error, also known as a false alarm, occurs when a defect is identified on the casting although no defect is present. Type II errors, or misses, occur when a casting passes inspection with a defect present. Although the determination of Type I and II errors is in itself subjective, these errors could be detrimental to the performance of the parts and could lead to disagreements between the manufacturer and customer if not interpreted as intended.

As a labor-intensive process, visual inspection requires the utmost attention to detail by the operator to minimize Type I and II errors. If at any time operators are not focused on their jobs or not physically and mentally alert, the risk of scrap or nonconformance increases. For instance, foundry environments where inspection takes place may be noisy and have poor lighting or extreme temperatures, which may be a distraction and impede the inspector's judgment. Assuring environmental and human factors are optimal will allow operators to perform at their best. Additionally, training operators on best practices to identify defects, such as rastering or using a visual aid, will improve consistency in identifying defects between operators resulting in a more stable process. These factors influencing Type I and II errors are not exhaustive; however, they do play a major role on casting inspection. Megaw [2] provides an extensive list of sources that can affect the accuracy of visual inspection.

The unique contribution of this article is the combination of various sources that impact the accuracy of visual inspection, as measured by Type I and II errors, to model the effectiveness of cast metal surface visual inspection. This article develops an influence diagram to calculate the probability of a Type I or Type II error. Although influence diagrams have frequently been used to assess risks and identify the optimal alternatives in business and public policy decisions, they have only rarely been applied to manufacturing decisions. Additionally, previous work exploring Type I and Type II errors in the casting industry only examines a single factor's impact.

This article incorporates and predicts the impact of several factors that contribute to Type I and II errors. Management at a manufacturing company can use this type of model to identify factors to focus improvement efforts on to decrease the number of Type I and II errors. The article presents a methodology for using influence diagrams to probabilistically assess the effect of different factors on the visual inspection process. An illustrative example for foundries in general, using results from previous research, is provided to demonstrate how this methodology can be applied. Foundries are encouraged to use their own data and expertise to reassess the probabilities given in this paper and determine likelihood of Type I and II errors for their own inspection processes. Although this article describes how the probabilities have been assessed for this illustrative example, the purpose of the article is not to describe the specific methodology for assessing probabilities either from data or from experts. Readers interested in learning more about how to assess the influence among factors and the likelihood of events are referred to [3-9].

II. BACKGROUND INFORMATION

Since this article draws from two distinct fields (manufacturing inspection and probabilistic risk analysis), it is necessary to provide background and cite the relevant literature for both fields. The first part of this section introduces the visual inspection standards and reviews the relevant literature on the inspection process. The second part of this section presents the influence diagram model, which will be used to assess the uncertainty in Type I and II errors. This brief review of both fields will provide the foundation to understand the model in Section III.

A. Current Visual Inspection Standards

Visual inspection of castings often occurs several times during their production and often is the final processing step before they are shipped. The workstation varies widely depending on many factors including the shop layout and size of castings. In almost all cases, the castings are delivered to the inspection station via a fork truck, overhead crane with a magnet, or via a roller crane. Depending on the size of the castings, they could be delivered individually or as a group of castings. For those that can be safely handled, they are often inspected as the inspector manipulates the part on a steel workbench. Medium sized castings are picked up via a jib crane operated by the inspector to safely access all sides of the castings. Very large castings are inspected on the floor, and then moved by the overhead crane to access the other sides. The environmental conditions of the inspection workstation will vary in these scenarios, but they are essentially always in a shop environment in the midst of the other processing steps. As with the casting size, the production volumes vary greatly where an inspector could be inspecting a few dozen or maybe a couple thousand castings in a day, which often consists of a variety of geometries. Any problem areas that need additional attention are highlighted with chalk or a special marking pen directly on the casting surface.

Many qualitative standards exist for the surface inspection of cast metal including company and industry specific standards. The Manufacturer Standardization Society (MSS) SP-55 Visual Method, American Society for Testing and Materials (ASTM) A802 which references the use of comparator from the Steel Castings Research and Trade Association (SCRATA), Alloy Casting Institute (ACI) Surface Indicator Scale, and GAR Electroforming Cast Comparator C9 are the most commonly used metal casting standards in industry. Inspectors use comparators and images in these methods to visually classify the surface roughness and abnormalities on an actual casting. The methods are primarily qualitative and based on a discretized scale, as opposed to a continuous scale, of classification.

In the MSS SP-55 method, images are used for comparison to cast surfaces. Twelve abnormality types, ranging from porosity to weld repair areas, are identified and images of acceptable and non-acceptable surfaces are provided for each [10]. Plastic replications of actual metal castings are used for comparison in the SCRATA method and adopted by ASTM [11]. Lettered plates representing one of nine abnormalities are used, each with various severity levels.

The abnormalities represented are similar to the MSS method. This standard is the most widely used standard in the U.S. steel casting industry. For the surface inspection process, inspectors compare the image or comparator associated with the surface specification to surface characteristics (abnormalities and roughness) of the casting. They then judge whether the surface characteristics fall below the threshold established by the plates. If the surface characteristics exceed the threshold, the part is rejected.

The ACI Surface Indicator evaluates “general smoothness, height and depth of irregularities extending beyond the range of general variations, and frequency and distribution of such irregularities” [12]. Designations SIS-1 through SIS-4 correspond to the root mean square (RMS) average deviation in micro-inches. The standard also specifies criteria for the height and frequency of surface abnormalities. Inspection is executed similarly to the two standards mentioned previously.

Less widely used than the other methods is the GAR C9 Comparator. Comparator swatches (each 12 x 36 mm) quantify the surface roughness based on root mean square (RMS) values in micro-inches. No abnormalities are defined in this standard. In addition to a visual examination, inspectors are instructed to “draw the tip of the fingernail across each surface at right angles” to match the texture of the inspected part [13].

Inspectors compare the surface of the casting to the appropriate standard in order to make the determination of whether or not the surface is acceptable. Regardless of the standard, inspectors should be trained in the applicable standard and have access to documentation to determine the acceptability of a part. Training should be ongoing to ensure inspectors remain calibrated [14]. Additionally, any errors identified downstream should be fed back to the inspector as soon as possible to reduce the likelihood of future occurrences [15]. Although these measures are in place to combat errors, the current standards lack robustness as they can be interpreted differently between people, rely on inspectors' sensory capabilities, and lack definition regarding rarely occurring abnormalities and their distribution over the surface. As long as there is a human element involved in the inspection process, various factors can affect their performance, which risk inaccurately determining whether or not a surface is acceptable. A digital standard is under development, which can be used to verify inspectors' judgments per customer requirements [16]. This will also lay the groundwork for more quantitative specifications for cast metal surfaces in the future, which would be an ideal method by reducing the human element and subjectivity of inspection.

While machine vision is readily applied for some casting surface inspection tasks, it is limited to a range of defects in certain areas. For example, online vision systems are used to detect defects on flat surfaces [17] and to match morphological features on a part surface to a database of similar geometrical defects [18]. However, this is not feasible for many castings as their geometries are complex and their defects are inconsistently shaped or located. A vision system would require that the orientation of the component is known, which would be time consuming and costly for the large variety of shapes produced in small quantities. Additionally, cleaning and maintenance of vision systems in a steel foundry would be a further disadvantage. Other methods compare the casting to a CAD model to identify defects [19], but these geometries may be in tolerance but differ from the perfect nominal due to inherent process variation. Thus, visual inspection methods are preferred for the several in-process inspection steps of a wide variety of castings within the production facility.

B. Influence Diagrams

An influence diagram—also called a Bayesian belief net or a decision diagram—models factors that contribute to a final outcome or uncertainty [3-4]. The influence diagram calculates the probability of the final outcome conditioned on all the factors in the model. The factors relate to each other and to the final outcome via conditional probabilities. Decisions can also be included in the influence diagram where a decision maker can understand how the probability of an outcome is influenced by each alternative [5]. For example, Fig. 1 depicts an influence diagram using Netica software where B is an uncertain outcome with two possible outcomes b1 and b2 (with probabilities 17.5% and 82.5%, respectively), A is an uncertain factor with two possible states a1 and a2 (with probabilities 25% and 75%, respectively), and D is a decision with two alternatives d1 and d2. The arrows in the model show that the uncertainty in B is conditionally dependent on the uncertainty in A and the decision D. In the decision node D, the graphical representation indicates that alternative d1 is selected.

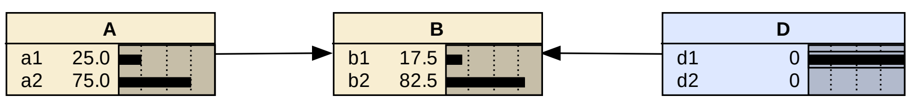

Fig. 1. An influence diagram with one factor A, one outcome B, and one decision D.

Computing the probability of b1 and b2 requires several assessments. First, it is necessary to assess the probability of a1 and a2 for factor A. Fig. 1 displays the probabilities: P(A = a1) = 0.25 and P(A = a2) = 0.75. Second, the probability of b1 and b2 should be assessed conditionally on factor A and decision D. For example, the probability of b1 given A = a1 and D = d1 is assessed as 0.1 and the probability of b2 given A = a1 and D = d1 equals 0.9. The example in Fig. 1 requires four such conditional assessments because A has two states and D has two alternatives. After these probabilities are assessed, typically through a combination of data and expert elicitation [6], the influence diagram calculates the probability of the outcome given each alternative. In Fig. 1,

$$P(B = b1 \mid D = d1) \tag{1}$$

$$= P(B = b1 \mid D = d1, A = a1) * P(A = a1) + P(B = b1 \mid D = d1, A = a2) * P(A = a2)$$

$$= 0.175$$

Software such as Netica enables the calculation of probabilities after the assessed probabilities are entered into the model.

Influence diagrams have been popular modeling tools for analyzing the risks of engineered systems [7, 20], decision making in business and public policy [21-23], and diagnosing disease [24]. Their role in assessing manufacturing problems and uncertainties has been much more limited, however. Some exceptions include diagnosing faults in manufacturing systems [25-27], optimizing a maintenance policy [28], modeling manufacturing processes with several control variables [29-30], and determining the optimal site for a manufacturing facility [31-32]. Influence diagrams typically are constructed from collected data as well as subject matter expertise to assess uncertainties for which data is not available [22, 33-34]. By combining data and personal expertise, influence diagrams represent a different modeling approach than most machine-learning algorithms, which require a large data set to estimate model parameters. This paper constructs an influence diagram in which some of the uncertainties and model parameters are derived from prior experimental data and some of the probabilities are assessed based on the authors' own expertise and research.

Influence diagrams can also be used to optimize a decision under uncertainty to maximize a decision maker's expected value or expected utility. Examples of using an influence diagram to optimize a decision include: choosing the most cost-effective strategy for managing river water quality [35], managing groundwater contamination [36], land management [37], mitigating the risk of an unmanned aerial vehicle crash [38], and managing highway maintenance projects [39]. Since there is no value function in this article, the influence diagram does not determine an optimal alternative, although discussion will be included on how the influence diagram could be extended to mitigate the risk in the visual inspection process.

III. INFLUENCE DIAGRAM FOR VISUAL INSPECTION

The visual inspection methods discussed in Section II are used to help determine if a part is defective; however, errors are frequent with these methods. This section builds an influence diagram to assess the likelihood of Type I and II errors in the visual inspection of cast metal surfaces and the effects of different interacting factors on them.

A. Overarching Model

Fig. 2 depicts an influence diagram to calculate the probability of a Type I error and a Type II error. The diagram is constructed in Netica to analyze various scenarios causing errors. A Type I error (false alarm) occurs when a defect is identified on the casting although no defect is present. A Type II error (miss) occurs when a casting passes inspection with a defect present. The nodes Type I error and Type II error represent uncertain nodes, and each node has two outcomes: the error occurs or does not occur. Two decisions influencing the probabilities of Type I and II errors are included: the training for the inspector and the judgment type used in the inspection process (on the right-hand side of Fig. 2). The manufacturer can determine the judgment to use in the inspection process (relative or absolute) and training type (basic or raster). The arrows from the decision nodes training and judgment type to Type I error and Type II error indicate the probability of each error is conditioned on the manufacturer's decision.

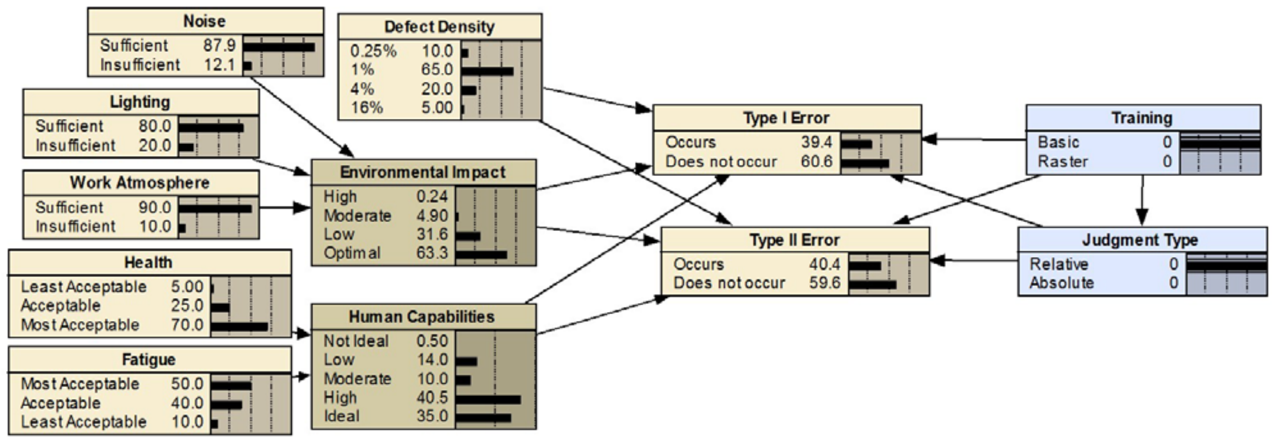

Fig. 2. Influence diagram for Type I and II errors for cast metal surface inspection. The non-zero numbers represent the probabilities of different states for each of the uncertainties. The two decision nodes on the right-hand side depict that basic training and relative judgment type are selected.

The left-hand side of Fig. 2 displays other uncertain factors influencing the probabilities of Type I and II errors. Three uncertain factors directly influence the likelihood of errors: defect density, environmental impact, and human capabilities. The arrows indicate conditional probability. For example, the probability of Type I error is conditional on defect density, environmental impact, and human capabilities, as well as on the training and judgment decisions. Since the node defect density does not have any arrows going into it, defect density is not influenced by any other factor in this model. The environmental impact depends on the noise, lighting, and work atmosphere, each of which has its own node. Human capabilities depend on the health and fatigue of the inspectors. The node health and the node fatigue each has arrows into human capabilities, which means the probability of an outcome under human capabilities is probabilistically dependent on health and fatigue.

The words inside each of the nodes in Fig. 2 represents the possible outcomes for each factor, and the number indicates the chance for that outcome. For example, the node noise has two outcomes, sufficient and insufficient. The probability of sufficient noise is 87.9%, and the probability of insufficient noise is 12.1%. As will be explained in the following sections, probabilities need to be assessed for each uncertain node. If an uncertain node has an arrow pointing to it, then conditional probabilities must be assessed.

After probabilities are assessed for all uncertainties in the influence diagram, the Netica software calculates the probability of a Type I and II error for each alternative in the training and judgment type decision. The output of the influence diagram is the probability of a Type I error and the probability of a Type II error for each combination of decisions: (i) basic training and relative judgment, (ii) raster training and relative judgment, (iii) basic training and absolute judgment, and (iv) raster training and absolute judgment. These probabilities will enable a manufacturer to quantify the impact of training and judgment on Type I and II errors while considering all environmental and human factors also contributing to those errors. Fig. 2 depicts the probabilities conditioned on the first combination of decision: basic training and relative judgment type.

The remainder of this section describes each factor in the influence diagram (training and judgment type, environmental factors, human capabilities, and defect density), describes how probabilities are assessed for each of the uncertain nodes, and explains each factor’s impact on Type I and II errors. The probabilities are based on previously conducted experiments, research, and the authors’ own expertise and knowledge about manufacturing conditions. Each of these sources are assumed to be conducted in ideal conditions; therefore, the results of the comprehensive model can be compared to the original source to better understand how these factors interact and how each individual factor impacts the overall outcome of a Type I or II error.

B. Training and Judgment Type

Methodologies used to calibrate inspectors affect the likelihood of Type I and II errors and consistency of identifying defects. This can be attributed to the enforcement of inspection procedures and effectiveness of training. Enforcing methodologies for inspection is a major factor in the consistency of identifying defects. This consistency helps analyze the reliability of the estimates for our Type I and II errors since the judgment of these errors are, in fact, as subjective as the inspection process. The type of judgment as well as the inspection sampling method impacts how defects are identified.

A manufacturer can choose to enforce a relative or absolute judgment in visual inspection. This explains why the node in Fig. 2 for judgment has two possibilities: relative or absolute. The model in this article assumes if the manufacturer chooses one of the two judgments, then all inspectors follow that judgment. Future research can study how well the manufacturer can enforce the type of judgment. Relative judgment occurs when the inspector has a comparator or image of the inspection criteria in hand for direct comparison to the cast part, while absolute judgment occurs when the inspector recalls the criteria from memory. Weber and Brewer [40] conducted a study to determine the differences in relative versus absolute judgment in relation to eye-witness accounts. In the relative judgment experiment, participants were asked to compare two individuals and pick which was previously shown in an image. For the absolute judgment experiment, the same participants were shown a single individual and asked if he or she had appeared in the previous image. Accuracy of absolute judgment in the study was found to be 69%, whereas for relative judgment it was found to be 80% as seen in Table 1. Although this study did not directly relate to the casting inspection process, these values can be used as insight into the impact of judgment type on Type I and II error. In the context of this study, an incorrect identification leads to a Type I or II error.

Table 1. Judgment type's effects on identification of defects from [40]

Judgment Type	Correct ID	Incorrect ID
Absolute	0.69	0.31
Relative	0.80	0.20

Peters et al. [41] evaluated the inspection of castings with and without comparators; data was collected in relation to Type I and II errors. Participants in the study were asked to categorize 25 casting surfaces as acceptable or not. For some surfaces, participants were given the comparator to use for references (relative), while others were to recall the criteria from memory (absolute). Table 2 shows the results of this study.

Table 2. Judgment type's effects on Type I and II errors from [41]

Judgment Type	Type I Error	Type II Error
Absolute	0.33	0.26
Relative	0.22	0.30

Training techniques also impact error in visual inspection, and the training node in Fig. 2 has two alternatives: basic and raster. In one case study, basic training and raster training were evaluated in casting inspection using absolute judgment [41]. Basic training involves giving the subject a general overview of which defects to look for on a casting; raster training also includes teaching subjects to systematically scan the part in a zig-zag pattern. This study also used eye tracking software to determine the percentage of the casting viewed under these conditions. Overall, the specific technique used to locate defects not only allowed the individual to view a greater percentage of the part, but it decreased the effects of Type I and II errors in the inspection process. The results of this study are found in Table 3; however, it is noted Type II error in raster training was about 16% more variable than for basic training. The subjects in this study had no prior experience with inspecting castings, which allowed for an unbiased result in the analyzing the overall effectiveness in training [40-42].

Table 3. Training effects on Type I and II errors and percent of part viewed [42]

Training	Type I Error	Type II Error	% Part Viewed
Basic	0.41	0.45	68
Raster	0.26	0.55	75

The decisions of training and judgment type, as seen on the right in Fig. 2, impact both Type I and Type II errors. The influence diagram depicts the judgment and training as decisions, which means that the manufacturer can choose absolute or relative judgment and basic or raster training. As will be described in more detail in Section IV, the probability of Type I / Type II error given judgment type (Table 2) is combined with the Type I / Type II error given training type (Table 3) in order to derive a probability conditioned on each combination of judgment and training. It is also necessary to factor in the environment factors, human factors, and defect density, which are now explained more fully.

C. Environmental Factors

Inspectors can be influenced by various environmental factors including the physical environment and work atmosphere. These aspects can reduce the inspector's effectiveness in the visual inspection process. The physical work environment includes auditory noise, light level, temperature, and humidity [1]. These can all distract the inspector and even reduce his or her capability to locate defects. For example, the just noticeable difference between the defect and surrounding area will reduce significantly if the lighting is poor, making the defect more difficult to locate. In general, both Type I and II errors increase in suboptimal conditions [43]. Additionally, the temperature and humidity can affect the inspector's cognitive ability. In fact, the ideal humidity of 65% and temperature of 70 degrees Fahrenheit in the presence of a fan can stimulate brain activity and increase alertness of the inspector [41].

The work atmosphere can also affect the inspector's likelihood to locate defects. In some workplaces, workers are rewarded for doing their job well while others are disciplined if quality is subpar. In some corporations, inspectors are required to re-inspect parts, either from a previous inspection or from another inspector. These are referred to as motivational losses. If inspectors receive a part that has already passed inspection once or know a part will be inspected later, they may not look as closely for defects because they feel it is a poor use of time. Both instances will increase the likelihood of Type II errors [43].

As depicted in Fig. 2, the factors of noise, lighting, and work atmosphere are assigned binary states of sufficient or insufficient in the influence diagram. It is necessary to assess the probability each one of the three factors is insufficient and assess how these three factors influence the overall environmental impact. These probabilities are subjectively estimated based on previous reports and the authors' expertise. Each of the main factors (noise, lighting, and atmosphere) are examined to determine the likelihood that each is in an acceptable or unacceptable state.

The noise element is a major environmental factor in steel foundries. Based on data collected in foundries, the noise level of the processes can range from 70 decibels in areas further from equipment to well above 85 decibels with some as high as 110 decibels. This not only affects the environment in which they currently work, but it can also affect long term health of the individual [44]. As is common with subjective probability assessments, an assumption is made that the noise level in a foundry follows a triangle probability distribution with the minimum, mode, and maximum of the triangle equal to 70, 85, and 110 decibels, respectively. Most foundries require their employees to wear at minimum noise reduction rated (NRR) 25dB hearing protection; therefore, the distribution was shifted to the left nine units to account for this practice (i.e., the minimum, mode, and maximum equal 61, 76, and 101 decibels, respectively). According to the Occupational Safety and Health Administration, exposure to sound levels above 90 decibels for an eight-hour work day can cause hearing damage, so any decibel above this level is classified at an unacceptable state [45]. Therefore, the probability the noise level is insufficient is 12.1% for this model, which is depicted in the noise node in Fig. 2.

Additional lighting at inspection stations is typically installed to increase visibility of the inspector; however, if the light levels become too bright, individuals may experience glare on the surface of the part reducing the ability to effectively inspect the surface. Placement of the casting in the lighting can also play a major role in successfully detecting defects due to shadows that may appear on the surface [2]. Based on a study on casting inspection, the range of lighting seen in inspection stations was from 150 to 15,000 lux with a mean of approximately 675 lux [46]. A beta probability distribution was fit to these parameters to model lighting. Ideally, the acceptable range to avoid glare-out and excessive shadows on the part is from 500 to 900 lux. Light levels outside of this range are considered insufficient. According to the beta distribution, there is 20% probability lighting will be insufficient.

Most foundries typically have more than one inspector for each casting process, whether it be on the same or different shifts. The larger foundries with more inspectors are likely to be more at risk for providing rewards to high performing inspectors or creating unintentional competition among the inspectors increasing the likelihood for error. According to a study in the United States, 20% of foundries were considered large businesses, which consisted of 100 or more employees [47]. Since the influence of incentives or competition among inspectors has not been studied in detail, a conservative assumption is made that 50% of the large businesses create an insufficient work environment. Thus, 10% of all foundries have an insufficient work environment as depicted in Fig. 2.

These three factors were chosen based on the estimated impact of each on the inspector. The environmental impact can either be high, moderate, low, or optimal based on the noise, lighting, and work atmosphere. The environmental state is assessed based on the number of insufficient factors as depicted in Table 4. If none of the factors (noise, lighting, and work atmosphere) are insufficient, the environmental state is optimal, and the probability of Type I and II error remains at the base level. If one factor is insufficient, the environmental impact is low, and the probabilities of Type I and II errors increase by 0.05. If two of three factors are insufficient, the environmental impact is moderate, and the probabilities increase by 0.1. If the all three factors are insufficient, the environmental impact is high, and the probabilities increase by 0.2. Since previous studies of Type I and II errors assumed ideal conditions for all nodes, if all factors are at a sufficient state, there is no change in the probability of Type I and II errors. The increase in probabilities based on the environmental state is incorporated into the influence diagram in Fig. 2.

Table 4. Environmental states and their impact on Type I and II errors

Number of Insufficient States (noise, lighting, work atmosphere)	Environmental State	Impact on Error
3	High	+0.20
2	Moderate	+0.10
1	Low	+0.05
0	Optimal	0.00

D. Human Capabilities

The capabilities of the individual performing the inspection also play a role in his or her ability to detect defects. These capabilities can be either physical, such as vision ability, or perceptual, such as memory ability.

As shown in Fig. 2, two uncertainties impact an inspector's capabilities: health and fatigue. An individual's health and fatigue can be impacted by several factors in a foundry environment, such as air quality, heat exposure, and overtime [47-48]. Visual, mental, and physical fatigue in inspectors can affect the judgment of whether or not a defect is present. When inspectors are tired, they can lose focus in the task at hand and become easily distracted [43]. Although fatigued inspectors may take additional time to view each part, errors generally increase [41]. Since no studies exist to our knowledge on the health of inspectors, an assumption is made that 50% of the time fatigue is most acceptable, 40% of the time fatigue is acceptable, and 10% of the time fatigue is least acceptable.

The age and health of the inspector can also be a limiting physical capability. This includes vision impairment, such as near or far sightedness, which could reduce the individual's ability to identify defects. This model assumes most inspectors have good health, and 70% of the time health is most acceptable, 25% of the time health is acceptable, and 5% of the time health is least acceptable, as show in Fig. 2.

The factors of fatigue and health were assigned states in the influence diagram. The states of fatigue and health are least acceptable, acceptable, and most acceptable. These factors were chosen based on the estimated impact of each on the inspector. The human capabilities node has five possible states: not ideal, low, moderate, high, and ideal. The impact on human capabilities is based on the states of each factor: least acceptable (LA), acceptable (A), and most acceptable (MA). Since previous studies of Type I and II errors assumed ideal conditions and human capabilities, if both fatigue and health are MA, there is no change in the probability of Type I and II errors. Table 5 depicts how the fatigue and health states combine to determine human capabilities and their impact on Type I and II errors.

Table 5. Deterministic values of human capabilities on Type I and II errors

Fatigue and Health States	Human Capabilities States	Impact on Error
2 LA	Not Ideal	+0.20
LA + MA/A	Low	+0.15
2 A	Moderate	+0.05
A + MA	High	+0.02
2 MA	Ideal	0.00

E. Defect Density

An inspector's perception of a task can greatly influence the likelihood of Type I and II errors. This includes developing a memory of past inspections and expectations over time. Inspectors who inspect the same part constantly develop a memory of where defects are most common. This may cause them to overlook other areas of the part to be inspected where defects are less common. In general, Type I errors become less common, and Type II errors increase [43].

Defect density, or the overall number of defects on a part, can affect Type I and II errors. Generally, as the defect density decreases, Type I and II errors increase. For example, if an inspector recalls from previous experience the number of overall unacceptable parts was approximately one out every five, he or she may begin to second guess previously inspected parts if ten or more in a row are found without any defects causing a Type I error. Similarly, if many parts with a lower number of defects are observed, parts with even fewer defects may be overlooked causing a Type II error. A study [49] using test samples with 0.25, 1, 4, and 16% defect densities was administered to 80 inspectors with no prior inspection experience. These inspectors were asked to identify all defects on each sample without being told how many defects to expect. If the inspector could not decide whether a specific feature was considered a defect, the test monitor acted as an inspection supervisor and advised them on how to classify the area in question. Results from this study can be found in the Table 6. The probability for the percentage of defects was determined by sampling actual castings produced in a foundry.

Table 6. Defect density's effect on Type I and II errors from [49]

Total Defects	**Type I Error**	**Type II Errors**
0.25%	0.85	0.42
1%	0.41	0.29
4%	0.15	0.25
16%	0.05	0.18

IV. DISCUSSION

A. Results

Populating the influence diagram in Fig. 2 requires combining data from different sources in order to assess the probabilities of Type I and II errors. Since each dataset that relates judgment type (Table 2), training (Table 3), or defect density (Table 6) to Type I and II errors does not consider the other two elements, an average of the three probabilities are used to determine the probability of an error conditioned on the judgment, training, and defect density. For example, if judgment is relative, training is basic, and the defect density is 0.25%, the probability of a Type I error is:

P(Type I Error):

$$= \frac{\text{P(Type I }|\mathit{relative\ judgement}) + \text{P(Type I }|\mathit{basic\ training}) + \text{P(Type I }|0.25\%\ \mathit{defect\ density})}{3}$$

$$= \frac{Table\ 2 + Table\ 3 + Table\ 6}{3}$$

$$= \frac{0.22 + 0.408 + 0.85}{3}$$

$$= 0.493 \qquad (2)$$

However, since it is assumed these studies were conducted under optimal conditions for environmental conditions and the ideal state for human capabilities, it is necessary to account for the possibility of less-than-ideal conditions in assessment of Type I and II probabilities. The influence diagram calculates the final probabilities for Type I and II errors based on the probabilities the factors are in given states and based on the adjustment for Type I and II errors as given in Tables 4 and 5.

Fig. 2 displays the influence diagram if training is basic and judgment is relative. If a manufacturer chooses these alternatives for its training and judgment, the probability of a Type I error is 0.39 and the probability of a Type II error is 0.40. Fig. 3 depicts the probability of Type I and II errors given each alternative for judgment and training type where each of these probabilities are computed via the influence diagram and the conditional probabilities. As seen in Fig. 3, relative judgment and raster training results in the smallest probability of a Type I error at 0.35, but it increases the probability of a Type II error to 0.44. Absolute judgment and basic training result in the smallest probability of a Type II error at 0.39, but leads to a 0.43 probability of a Type I error. The training has opposite effects on Type I and Type II errors. More robust training and judgment types (raster and relative) decrease the probability of false alarms (Type I error) and increase the probability of misses (Type II error). This result is from the studies [41-42] as depicted in Tables 2 and 3 in which raster training results in more Type II errors than basic training and relative judgment results in more Type II errors than absolute judgment.

The probabilities of Type I and II errors are fairly high, and it may be worrisome that the probabilities of these errors are between 0.35 and 0.45 regardless of the training and judgment type chosen by the manufacturer. However, these probabilities align closely with the experiments previously cited in which the probability of a Type I error ranges between 0.22 and 0.41 and a probability of a Type II error ranges between 0.26 and 0.55 without considering defect density. When defect density is included (Table 6), the probability of Type I error can be as large as 0.85. Although the prior studies provide estimates of the probabilities, actual values will vary among individual foundries. An individual manufacturer can substitute probabilities of the different factors more accurate for its foundry.

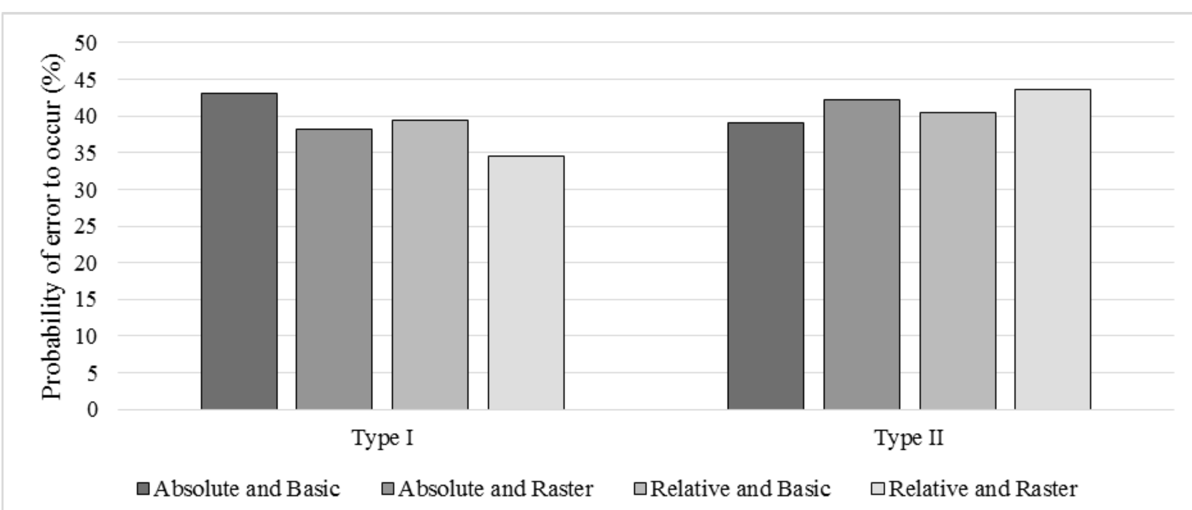

Fig. 3. Base values of error comparing judgment and training type decision without certainty of other factors for influence diagram model

B. Sensitivity Analysis

Since many of the probabilities in the model are based on assumptions of how the different factors interact with each other, sensitivity analysis can help determine to what extent the probabilities for Type I and II errors depend on these assumptions. An influence diagram can easily demonstrate how changing a factor from one state to another state impacts the final outcome. Fig. 4 shows the probabilities of Type I and II errors when each factor is moved from its best state to its worst state and the other factors remain constant. The Type I error probabilities are based on raster training and relative judgment, and the Type II error probabilities are based on basic training and absolute training. For example, if defect density is 16%, the probability of a Type I error is 0.23. If defect density is 0.25%, the probability of a Type I error is 0.50, as depicted in Fig. 4. Defect density has the largest impact on the probability of a Type I error. If fewer defects are present, inspectors have fewer defects to identify, which increases their tendency to over inspect parts and cause false alarms. Each of the other five factors only reduce the probability of a Type I error by approximately 0.03 if one of them is at the best state. If fatigue or health is in the worst state, however, the probability of a Type I error increases significantly to more than 0.46 in each case.

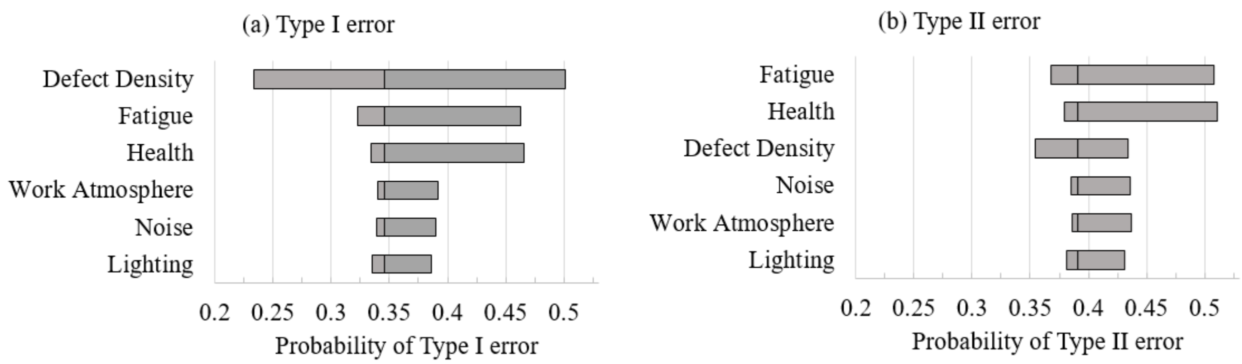

Fig. 4. Sensitivity analysis for (a) Type I error with raster training and relative judgment and (b) Type II error with basic training and absolute judgment

Fatigue and health have the largest effect on the probability of a Type II error. If fatigue or health is in the worst state, the probability of a Type II error increases to more than 0.5. When inspectors are fatigued or in bad health, their attention is less focused, resulting in a tendency to miss defects. If fatigue is in its best state, the probability of a Type II error decreases to 0.37. Defect density also has large impact on the probability of a Type II error. If defect density is 0.16%, the probability of a Type II error is 0.35.

The influence diagram can also be used to ascertain how good and how bad the outcomes can be. If the environmental impact is optimal, the human capabilities is ideal, and defect density is 16%, the probability of a Type I error is 0.18 with relative judgment and raster training. The probability of a Type II error under these same uncertainty conditions is 0.30 with absolute judgment and basic training. By ensuring ideal conditions exist for manual inspection (e.g. sufficient lighting, less noise, healthy and well-rested inspectors), a manufacturer can significantly decrease the probability of a Type I error from the base-case probability of 0.35. The probability of a Type II error can only be decreased from 0.39 to 0.30. However, a key contributing factor to the lower probability of a Type I error is a high defect density, which does not seem desirable for a manufacturer. If the defect density is 1%, the environmental impact is optimal, and the human capabilities ideal, the probability of a Type I error is 0.30 with raster training and relative judgment, which is only slightly less than the base-case probability.

However, if the manufacturer ignores the environmental conditions and human capabilities and lets these conditions deteriorate to their worst cases, the probabilities of Type I and II errors increase dramatically. If environmental impact is high, human capabilities is not ideal, and defect density is 0.25%, the probability of a Type I error is the probability of a Type I error is 0.85 with relative judgment and raster training, and the probability of a Type II error is 0.78 with absolute judgment and basic training. Although such an extreme case is very unlikely, it demonstrates how much error would result from visual inspection if conditions are extremely poor.

This sensitivity analysis demonstrates what a manufacturer could do to improve its inspection process in addition to choosing the training and judgment type. Each factor that contributes to the environmental impact (noise, lighting, and work atmosphere) has little individual effect on the probabilities of Type I and II errors. The two factors for human capabilities (health and fatigue) have a larger effect on Type I and II errors than the environmental factors. For example, ensuring inspectors are not fatigued decreases the probability of a Type II error. Targeting areas like fatigue and defect density would be ideal if a manufacturer wants to reduce one type of effect; this could include requiring visual exercises to reduce eye strain or increasing awareness of defect density among inspectors.

C. Other Factors

Another factor that can influence the validity of the inspection process is how specifications are interpreted. Factor interpretation was not included in the influence diagram because it is unclear how the interpretation directly impacts the probability of a Type I or II error. Interpretation is an important factor and deserves some discussion. Since various standards can be used to inspect cast metal surfaces and there is no easy way to calibrate inspectors, the results from visual inspection are subjective [1]. As discussed in Section II, inspection standards may consist of methods using images while others use physical comparators. Some standards identify specific types of abnormalities to look during inspection. If a standard does not define an abnormality, there is no way for the customer to specify what is desired. On the other hand, if the customer only specifies criteria for porosity and the part has inclusions, the inspector must determine whether to only inspect for the porosity or consider other abnormalities. This causes confusion for both parties. The interpretation of the standard can contribute to uncertainty and variability in the inspection process.

Issues with repeatability (variation for a single inspector) and reproducibility (variation between inspectors) may also arise within a company's inspecting team, which affects the consistency of identifying defects. Visual inspection methods show large variation in measurement error for both repeatability and reproducibility due to inconsistencies for a single inspector between parts and between inspectors on the same part [50]. The average repeatability across six operators from three foundries was 66.83%, while the average reproducibility for operators at the same facilities was 63.33% [51]. Since the consistency of an inspection requires that the inspection is both repeatable and reproducible, consistency can be calculated as the product of the probability of repeatability and reproducibility.

The variation in identifying defects will impact Type I and II errors, but it is not known whether it would increase the chances of missing a defect and false alarms. The lack of consistent standards and the lack of repeatability and reproducibility signify that the probabilities of Type I and II errors will vary among inspectors and from one inspection to another inspection. Even if a foundry has optimal environmental impact and ideal human capabilities, if it does not have consistent standards, some inspectors may find many more defects and other inspectors may find far fewer defects. Without a consistent standard, it is difficult to know if the inspectors who are finding more defects are making a lot of Type I errors or if the inspectors who are finding few defects are making a lot of Type II errors. Judgment type and the inspection method will also impact the consistency of evaluation.

Qualitative standards for cast metal surfaces rely on an individual's capability to judge if a part is acceptable. Individuals must differentiate between the types of abnormalities present. It can be unclear if a part is acceptable when an unexpected abnormality appears on the final part if the abnormality was not taken into consideration by the customer when specifying the surface. The interpretation of the standard or specification varies greatly among inspectors and between the customer and manufacturer, and these factors increase the risk of Type I and II errors resulting from the inspection process.

D. Comparison with Naïve Bayes

The data depicted in Tables 2-4 are used to construct a naïve Bayes model to estimate the likelihood of Type I and II errors given judgment and training type. The results from the naïve Bayes model can be used to validate the influence diagram approach. The naïve Bayes model uses Bayes' theorem but assumes that the probability of judgment type and probability of training type are conditionally independent of each other. The probability of Type I error or Type II error given judgment and training type is calculated:

$$P(error|judgment, training) = \frac{P(error)P(judgment|error)P(training|error)}{P(judgment,training)} \quad (3)$$

where

$$P(judgment, training) = P(error) * P(judgment|error) * P(training|error) + P(no\ error) * P(judgment|no\ error) * P(training|no\ error),$$

$error$ is either Type I error or Type II error, $judgment$ is either absolute or relative, and $training$ is either basic or raster. The variable $no\ error$ is the complement of $error$, so $P(no\ error) = 1 - P(\text{Type I error})$ if $error$ is Type I error.

This model requires $P(error)$, which is the marginal probability of a Type I error or Type II error. Tables 2-4 are used to estimate this probability by averaging the likelihood of a type of error from each table and then averaging the three averages. This method assumes that the two judgment types are equally likely, the two training types are equally likely, and the four defect densities are equally likely. The probability of a Type I error is 0.33 and the probability of a Type II error is 0.35. The conditional probability of judgment given the error type $P(judgment|error)$ equals the probability of the error type given judgment divided by the sum of the probabilities of error type given each judgment as depicted in Table 2. The conditional probability of training given the error type $P(training|error)$ equals the probability of the error type given training divided by the sum of the probabilities of error type given each training as depicted in Table 3. For example, the probability of absolute judgment given Type I error is calculated as $0.33/(0.33 + 0.22) = 0.6$.

Fig. 5 depicts the results of the naïve Bayes model where the probabilities of Type I and Type II errors are conditioned on judgment and training. The naïve Bayes does not consider the environmental conditions and the human capabilities modeled in the influence diagram. The probabilities in the naïve Bayes model (Fig. 5) have a greater spread than the probabilities in the influence diagram (Fig. 3). The probability of Type I error ranges from 0.2 to 0.46 and the probability of Type II error ranges from 0.29 to 0.43 in the naïve Bayes model, whereas the probability of Type I error ranges from 0.35 to 0.43 and the probability of Type II error ranges from 0.39 to 0.44 in the influence diagram. The naïve Bayes model has greater ranges because the naïve Bayes model assumes that the four defect densities in Table 5 are equally likely, but the influence diagram assumes that a 1% defect density is much more likely than the other defect densities. The naïve Bayes model and influence diagram exhibit very similar trends because the absolute judgment and basic training result in the largest probability of a Type I error and the smallest probability of a Type II error in both models. Relative judgment and raster training generate the smallest probability of a Type I error and the largest probability of a Type II error in both models.

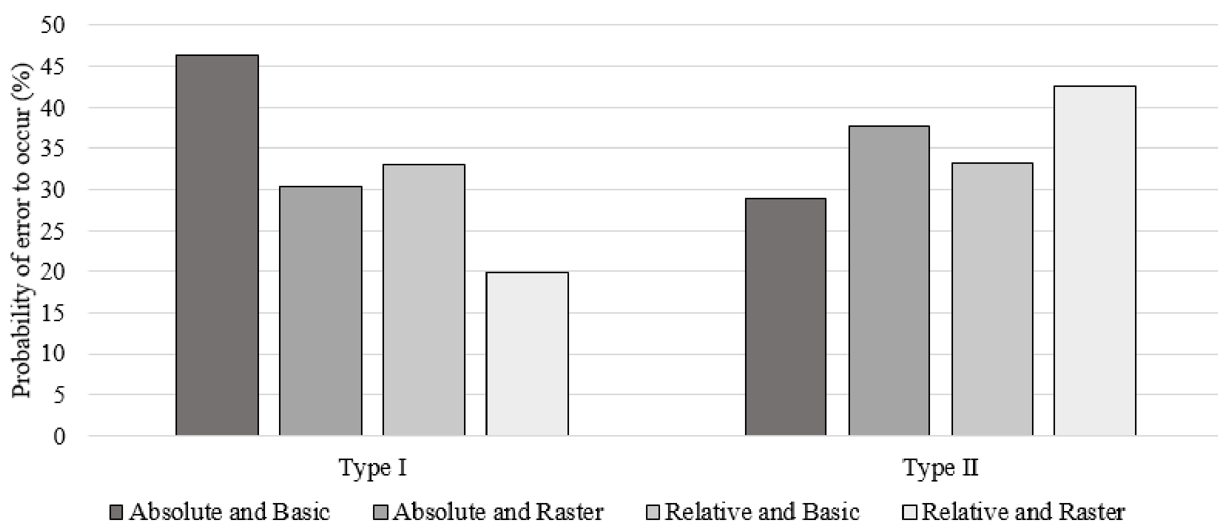

Fig.5. Base values of error comparing judgment and training type decision without certainty of other factors for naïve Bayes model

The influence diagram incorporates the impact of environmental factors and health capabilities, which is left out of the naïve Bayes model. The influence diagram is flexible enough that it can incorporate the subjective assessments on how detrimental environmental factors and health capabilities affect the likelihood of the Type I and II errors. Since the naïve Bayes model requires data, which is not available for environmental factors and human capabilities. The influence diagram approach enables an analysis such as Fig. 4, which quantifies the impact of moving the worst level and the best level for each factor. Whether an individual favors the naïve Bayes' forecast or the influence diagram's forecast depends to a large extent on the individual's comfort with including subjective assessments.

V. CONCLUSIONS

Surface standards for metal cast surfaces help to determine the acceptability of surface quality; however, with current standards and capabilities, a large amount of variability exists in the visual inspection process. This article represents the first use of an influence diagram to model the inspection process of surface capabilities. The influence diagram models and demonstrates how the different factors interact to impact Type I and II errors. The probabilities in the influence diagram are derived from previous studies and the authors' expertise. According to the model, Type I errors appear slightly less frequently than Type II errors. However, each type of error must be examined independently of one another to understand the impact. In the case of a Type I error, acceptable parts are being held at the manufacturer unnecessarily causing an increase in work-in-process inventory and adding additional labor for rework and re-inspection. If multiple inspectors arrive at this same conclusion, the parts may even be scrapped. In the case of a Type II error, parts are leaving the manufacturer and arriving at the customer in an unacceptable condition. If the customer does not do an in-house inspection before using the parts, they could be assembled into final products and could damage the customer's reputation to the consumer.

The influence diagram developed in this assessment provides additional insight into the visual inspection process. The model of individual factors and their interactions with one another present a broader picture of the problem. Using Netica allowed for a simple means of comparing scenarios when uncertainty nodes changed state or decision nodes were declared. This provides a better understanding of how a variety of factors play a role in affecting Type I and II errors. Individual foundries can use this model input with the current probability of occurrence of these factors in their facilities to compare with actual results from their visual inspection process.

A limitation of this research is the subjective assessment both in terms of how factors relate to each other and in the estimation of parameters. Due to the lack of solid data collected by a foundry or a careful design of experiments that measures how factors combine to affect Type I and II errors, it is recognized that these estimates and functional relationships may have substantial uncertainty. Others may disagree with the assessments provided in this article, and this model is flexible enough to incorporate their own estimates. Having more precise data on how the environmental factors, human capabilities, and defect density interact to affect the likelihood of Type I and II errors would enable a more reliable means of estimating parameters.

However, the limitations of this article also point to a strength and usefulness of using influence diagrams to model risk in the visual inspection process. The influence diagram is well suited to integrate subjective assessments with data, which fits well with the knowledge basis of the visual inspection process. The influence diagram is constructed based on the authors' expertise into the inspection process, a handful of prior experiments testing Type I and II errors, and the authors' ongoing conversations with foundry operators. The manner in which the factors relate to each other and are modeled within the influence diagram is based on this expertise, and many assessments of the probabilities are derived from prior experiments. Without a model that can integrate data with subjective assessments, the analysis would either rely on the prior experiments that only measure the influence of a single factor (as in the case of the naïve Bayes model) or be completely qualitative and subjective. The influence diagram developed in this article can combine subjective assessments and data (which are derived from experiments) into a probabilistic model that provides additional insight into misclassification errors in a manual inspection process.

Future research could compare the results of the influence diagram modeling approach with more data-intensive approaches, such as naïve Bayes which was used as a comparison in this paper. Although the goal of this article is not to optimize the inspection process, the influence diagram can be used to measure the benefits of improving conditions, instituting a different training regimen, and enforcing a judgment methodology. If the costs of these actions are known, the manufacturer can use the influence diagram to optimize its action based on maximizing the benefit-cost ratio.

The consistency of identifying defects, however, is extremely variable, which means the estimates for Type I and II errors contain a significant amount of variability. The judgment of these errors are as subjective as the inspection process. Clearer communication of expectations of cast surface specifications is needed between the manufacturer and customer.

To improve communication in visual inspection, the manufacturer and customer should convene to discuss their expectations of surface quality in regards to the comparator methods available. Additionally, training procedures should be developed so inspectors are calibrated with one another. A yearly refresher course, at minimum, would be ideal to verify the inspectors remain calibrated throughout the duration of their inspection duties.

To reduce the subjectivity and variability of visual inspection, quantitative criteria should be implemented. A digital surface standard can be developed to provide a quantitative method of inspecting cast metal surfaces. This standard should reduce the variation and improve the accuracy in the surface inspection process. The influence diagram could be expanded to assess how the probabilities of errors change with such a standard.

VI. ACKNOWLEDGEMENTS

This research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-12-2-033. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

VII. REFERENCES

[1] Monroe, R. "Design "As Is" and "To Be"" Steel Founder's Society of America Meeting. July 2015.

[2] Megaw, E. "Factors Affecting Visual Inspection Accuracy." *Applied Ergonomics* 10.1 (1979): 27-32. Web. 11 Nov. 2015.

[3] Howard, R. and J. Matheson. “Influence Diagrams.” *The Principles and Applications of Decision Analysis*. 1984, Strategic Decisions Group: Menlo Park, CA.

[4] Pearl, J. “Probabilistic Reasoning in Intelligent Systems.” *Networks of Plausible Reasoning*. 1988, Morgan Kaufmann Publishers, Los Altos.

[5] Shachter, R. "Model Building with Belief Networks and Influence Diagrams." *Advances in Decision Analysis: From Foundations to Applications*. (2007): 177-207.

[6] Woodberry, O., A. Nicholson, K. Korb, and C. Pollino. “Parameterising Bayesian networks.” *Australasian Joint Conference on Artificial Intelligence*. 2004, Springer Berlin Heidelberg: 1101-1107.

[7] Bedford, T. and R. Cooke, *Probabilistic Risk Analysis: Foundations and Methods*. 2001, Cambridge: Cambridge University Press.

[8] Morgan, G., and M. Henrion. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge: Cambridge University Press, 1990.

[9] Spetzler, C., and Stael von Holstein, C.-A. “Probability Encoding in Decision Analysis.” *Management Science* 22.3 (1975): 340-358.

[10] MSS. Quality Standard for Steel Castings for Valves, Flanges and Fittings and Other Piping Components (Visual Method). Tech. no. SP-55. New York: Manufacturers Standardization Society of the Valve and Fittings Industry, 1961.

[11] ASTM International. *Standard Practice for Steel Castings, Surface Acceptance Standards, Visual Examination*. Tech. no. A802/A802M-95. West Conshohocken: ASTM Int'l, 2006.

[12] Eubanks, P. *Standard for the Visual Inspection of Casting Surfaces*. Manual. Des Plaines: Steel Founders' Society of America Alloy Casting Institute, 1969.

[13] ASME. *Surface Texture: Surface Roughness, Waviness and Lay*. New York: American Society of Mechanical Engineers, 2009. Print. ASME B46.1-2009.

[14] A997-08, “Standard Practice for Investment Castings, Surface Acceptance Standards, Visual Examination,” *ASTM International*, West Conshohocken, PA, 2012.

[15] Thapa, V., A. Gramopadhye, B. Melloy, and L. Grimes. “Evaluation of different training strategies to improve decision-making performance in inspection,” *International Journal of Human Factors in Manufacturing.* 6.1 (1996): 243-261.

[16] Voelker, M. M. and Peters, F. E., "Development of a Digital Standard to Specify Surface Requirements of Cast Metal Surfaces." *Materials Performance and Characterization* 6.2 (2017): 130-143.

[17] Fernandez, C., C. Platero, P. Campoy, and R. Aracil. “Vision system of on-line surface inspection in aluminum casting process.” *IEEE Proceedings from the Conference on Industrial Electronics, Control, and Instrumentation.* (1993).

[18] Frayman, Y., H. Zheng, and S. Nahavandi. “Machine vision system for automatic inspection of surface defects in aluminum die casting.” *Journal of Advanced Computational Intelligence and Intelligent Informatics.* 10.3 (2006): 281-286.

[19] Newman, T., and A. Jain. “A system for 3D CAD-based inspection using range images.” *Pattern Recognition*. 28.10 (1995): 1555-1574.

[20] Trucco, P., E. Cagno, F. Ruggeri and O. Grande, “A Bayesian Belief Network Modelling of Organisational Factors in Risk Analysis: A Case Study in Maritime Transportation.” *Reliability Engineering and System Safety*, 93 (2008): 823-834.

[21] Varis, O. "Bayesian Decision Analysis for Environmental and Resource Management." *Environmental Modelling & Software* 12.2-3 (1997): 177-185.

[22] Neil, M., N. Fenton, S. Forey, and R. Harris. “Using Bayesian belief networks to predict the reliability of military vehicles.” *Computing & Control Engineering Journal,* 12(1) (2001): 11-20.

[23] Paté-Cornell, E. and S. Guikema. "Probabilistic Modeling of Terrorist Threats: A Systems Analysis Approach to Setting Priorities among Countermeasures." *Military Operations Research* 7.4 (2002): 5-23.

[24] Owens, D., R. Schachter and R. Nease. "Representation and Analysis of Medical Decision Problems with Influence Diagrams." *Medical Decision Making* 17.3 (1997): 241-262.

[25] Lewis, R. and R Ransing. "A Semantically Constrained Bayesian Network for Manufacturing Diagnosis." *International Journal of Production Research* 35.8 (1997): 2171-2188.

[26] Mehranbod, N., M. Soroush, M. Piovoso and B. Ogunnaike. "Probabilistic Model for Sensor Fault Detection and Identification." *AIChE Journal* 49.7 (2003): 1787-1802.

[27] Jeong, I., V. Leon and J. Villalobos. "Integrated Decision-Support System for Diagnosis, Maintenance Planning, and Scheduling of Manufacturing Systems." *International Journal of Production Research* 45.2 (2007): 267-285.

[28] Vatn, J., P. Hokstad and L. Bodsberg. "An Overall Model for Maintenance Optimization." *Reliability Engineering & System Safety* 51.3 (1996): 241-57.

[29] Agogino, A., S. Srinivas and K. Schneider. "Multiple Sensor Expert System for Diagnostic Reasoning, Monitoring and Control of Mechanical Systems." *Mechanical Systems and Signal Processing* 2.2 (1988): 165-185.

[30] Nadi, F., A. Agogino and D. Hodges. "Use of Influence Diagrams and Neural Networks in Modeling Semiconductor Manufacturing Processes." *IEEE Transactions on Semiconductor Manufacturing* 4.1 (1991): 52-58.

[31] Canbolat, Y, K. Chelst and N. Garg. "Combining Decision Tree and MAUT for Selecting a Country for a Global Manufacturing Facility." *Omega* 35.3 (2007): 312-325.

[32] Dogan, I. "Analysis of facility location model using Bayesian Networks." *Expert systems with applications* 39.1 (2012): 1092-1104.

[33] Agogino, A., and A. Rege "IDES: Influence Diagram Based Expert System." *Mathematical Modelling* 8 (1987): 227-233.

[34] Neil, M., N. Fenton, and M. Tailor. "Using Bayesian Networks to Model Expected and Unexpected Operational Losses." *Risk Analysis* 25.4 (2005): 963-72.

[35] Vans, O. "A Belief Network Approach to Optimization and Parameter Estimation: Application to Resource and Environmental Management." *Artificial Intelligence* 101.1 (1998): 135-63.

[36] Farmani, R., H. Henriksen, and D. Savic. "An Evolutionary Bayesian Belief Network Methodology for Optimum Management of Groundwater Contamination." *Environmental Modelling & Software* 24.3 (2009): 303-10.

[37] Marcot, B., R. Holthausen, M. Raphael, Mary M. Rowland, and Michael J. Wisdom. "Using Bayesian Belief Networks to Evaluate Fish and Wildlife Population Viability under Land Management Alternatives from an Environmental Impact Statement." Forest ecology and management 153.1 (2001): 29-42

[38] Luxhøj, J. "A Conceptual Object-Oriented Bayesian Network (OOBN) for Modeling Aircraft Carrier-Based UAS Safety Risk." *Journal of Risk Research* 18.10 (2015): 1230-58.

[39] Bayraktar, M. and M. Hastak. "Bayesian Belief Network Model for Decision Making in Highway Maintenance: Case Studies." *Journal of Construction Engineering and Management* 135.12 (2009): 1357-69.

[40] Weber, N., and N. Brewer. "Confidence-Accuracy Calibration in Absolute and Relative Face Recognition Judgments." *Journal of Experimental Psychology: Applied* 10.3 (2004): 156-72.

[41] Peters, F., R. Stone. "Visual Inspection of Casting Surfaces." *AFS Transactions* (2013): 45-52.

[42] Watts, K. *The effect of visual search strategy and overlays on visual inspection of castings*. M.S. Thesis, Iowa State University, 2011.

[43] Schorn, T. "Management Risks in Multiple 100% Inspections." American Foundry Society (2013).

[44] National Occupational Health and Safety Commission. Foundry Health Hazards. Canberra: AGPS, 1989. Safe Work Austrailia. Web. 10 Nov. 2015.

[45] Occupational Safety & Health Administration. "Regulations and Standards." U.S. Department of Labor. Web. 10 Nov. 2015.

[46] Stone, R. "Environmental and Social Factors." IE 577- Human Factors Lecture. Howe Hall, Iowa State University, Ames. 16 Nov. 2015. Lecture.

[47] Spada, A. "Revitalization of American Metalcasting." American Foundry Society. Minnesota, Twin Cities. 17 Feb. 2015. Lecture.

[48] Occupational Safety & Health Administration, United Food and Commercial Workers International Union. *Hazards of Extreme Temperatures: Handling Heat and Cold*. 2015. Web. 11 Nov. 2015.

[49] Harris, D. "Effect of Defect Rate on Inspection Accuracy." *Journal of Applied Psychology* 52.5 (1968): 377-79. Web. 15 Oct. 2015.

[50] Daricilar, G. *Measurement Error of Visual Casting Surface Inspection*. M.S. Thesis, Iowa State University, 2005.

[51] Daricilar, G., and F. Peters. "Methodology for Assessing Measurement Error for Casting Surface Inspection." *International Journal of Metalcasting* (2011): 7-15.