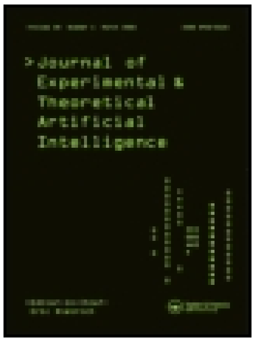

Taylor & Francis
Taylor & Francis Group

Journal of Experimental & Theoretical Artificial Intelligence

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/teta20

A comparison between the Bayesian network model and the logistic regression model in prevention of the defects on ceramic tiles

Volkan Sevinç & Meryem Merve Kırca

To cite this article: Volkan Sevinç & Meryem Merve Kırca (2022): A comparison between the Bayesian network model and the logistic regression model in prevention of the defects on ceramic tiles, Journal of Experimental & Theoretical Artificial Intelligence, DOI: 10.1080/0952813X.2022.2153273

To link to this article: https://doi.org/10.1080/0952813X.2022.2153273

Published online: 08 Dec 2022.

JOURNAL OF EXPERIMENTAL & THEORETICAL ARTIFICIAL INTELLIGENCE
https://doi.org/10.1080/0952813X.2022.2153273

Taylor & Francis
Taylor & Francis Group

ARTICLE

Check for updates

A comparison between the Bayesian network model and the logistic regression model in prevention of the defects on ceramic tiles

Volkan Sevinç [a] and Meryem Merve Kırca [b]

[a]Faculty of Science, Department of Statistics, Muğla Sıtkı Koçman University, Muğla, Turkey; [b]Graduate School of Natural and Applied Sciences, Statistics Program, Muğla Sıtkı Koçman University, Muğla, Turkey

ABSTRACT

One of the most important problems encountered in ceramic tile industry is defective product problem. Defective ceramics lead to loss of income and waste of resources in enterprises. However, it is generally unknown that which factors and production stages cause what kinds of defects. On the other hand, in the literature, the existing modelling studies usually consider the defects seen on industrial ceramics. The defect types of industrial ceramics and those of ceramic tiles are different. This article investigates the reasons behind the defect occurrences on ceramic tiles, along with a comparison between a logistic regression model and a Bayesian network model. The study shows that the Bayesian network model is more successful in estimating the defect types. The constructed Bayesian network model indicates that, in general, the high levels of the production band speed significantly increase the probabilities of all kinds of defects except the deformation defect. Additionally, the high densities of the glaze also increase the occurrence levels of the defects except the deformation defect. Similarly, the high levels of the engobe weight and the engobe density are also among the factors increasing the defect occurrences.

ARTICLE HISTORY
Received 6 December 2021
Accepted 25 November 2022

KEYWORDS
Ceramic tiles; defective tile production; Bayesian networks; machine learning; logistic regression

Introduction

One of the most important problems encountered in the production of ceramic tiles is defective product problem. The defects, which negatively affect the quality of the ceramics, also cause loss of raw materials, time, and income in enterprises. However, it is not possible to prevent the product defects completely, because it is not definitely known that at what stage and for what reason the defects occur. Thus, defective ceramics lead to loss of income and waste of the resources in enterprises. An important step in prevention of the ceramic defects is to determine the production stages and conditions causing the defects.

There are various studies in the literature concerning the quality and the defects of industrial ceramics. Agarwal and Rao (2005) conducted a study on the estimation of the surface quality in ceramic sharpening. They focused on the surface quality parameters and the surface roughness. They aimed to create a model for the estimation of the surface roughness. It is emphasised that the existing models are inadequate and not used in practice. In addition, it is stated that many variables that are affected by the process are nonlinear, interdependent, or difficult to measure, making the modelling process difficult. The ceramic sharpening model was designed in accordance with the

CONTACT Volkan Sevinç vsevinc@mu.edu.tr Faculty of Science, Department of Statistics, Muğla Sıtkı Koçman University, Kötekli Kampüsü, Muğla 48000, Turkey

stochastic nature of the sharpening process, the random geometries, and the distributions of the cutting edges. The theoretical results of the model have been shown to be compatible with the experimental results. Chen et al. (2011) created a simplified thermochemical model, and issued environmental stress distribution rules at the boundary of contact between the ceramic shell and patterns. They also introduced the environmental stress variation rule. Everhart et al. (2012) worked on ageing-based injury and developed a three-dimensional nonlinear model to try to predict cracking in the ceramic during wax removal. According to the results of the study, it is stated that ageing should be done over the glass transition temperature for a minimum of 24 hours. Everhart et al. (2013) later proposed a static displacement-controlled modelling method for estimating the defeat stress in the corner areas of the ceramic shell. The model analyzes the effect of shell porosity on stress on sharp edges. Various experimental studies have also been conducted to examine the deformation in the ceramic shell. Among these studies, Jiang and Liu (2004) examined the dimensional changes in the shell samples for different conditions. Wereszczak et al. (2002) compared the creep decay in two ceramic shells and characterised the decay. In another study, regarding deformation in ceramic shell, Jin et al. (2017) adopted a Bayesian network approach. According to the results given by the Bayesian model, the most important factor causing deformation on industrial ceramic shells appeared to be dimensional expansion, while the most important factor leading to dimensional expansion was found to be firing temperature.

In another study on estimating the wear of ceramics in industrial use, Marshall et al. (1982) developed a model using the progression of lateral cracks. Kim et al. (1986) proposed the contact severity index and stated that by defining the wear modes under different contact severity, the surface strength of the material and the tendency to fracture based on the applied stress can be estimated. The stress crack model, developed by Wang and Hsu (1996), focuses on the formation of cracking mechanisms, and explains the ceramic wear. A concept, which is based on energy balance, is used in the model and the equation of wear is obtained. Ting and Winer (1989) have proposed a thermochemical model and used the thermal stresses from friction to calculate the total stress. Kong and Ashby (1992) proposed to evaluate both bulk temperature and flash temperature in defining ceramic wear. Hsu and Shen (2004) described the ceramic wear as the microstructure, the grain shape, the size of the ceramic, the grain boundary strength, and a complex function of operating conditions. In a study evaluating various deterioration parameters in industrial ceramic production, Liu et al. (2019) proposed the Permafrost Analysis Method (PAM) for the analysis of deformation in the ceramic shells during the firing process. By comparing the simulation-based and the experimental results, they concluded that the PAM method predicted the deformations occurring in the ceramic shells with high accuracy.

In the literature, there are also studies concerning the detection of the defects on ceramic tiles Younas et al. (2022) suggested a novel approach for detecting the ceramic tile cracks based on image processing and morphological operation. They report an accuracy level of 92%. In another study by Wang et al. (2022), however, they developed an algorithm called the Concatenated Network Fusion Algorithm (CNFA) to improve the detection accuracy of surface defects on ceramic tiles. They report that their algorithm has an accuracy rate of 98.02%, which is 13% more than other algorithms. Hanzaei et al. (2017) used a support vector machine classifier to develop an automatic defect-type detection algorithm on ceramic tiles. They compared the results of the support vector machine classifier with the results of an artificial neural network model that they built. They suggest that the support vector machine classifier that they proposed performed better than the artificial neural network. In another study by Li et al. (2021), they used YOLOv3 algorithm to detect six kinds of tile defects and they analysed the experimental results they obtained. Wan et al. (2022), however, suggested a deep learning method for ceramic tile surface defect detection, based on the improved YOLOv5s network architecture. They point out that their improved model can localise and identify the defects with an accuracy rate of 94.27%. Additionally, Stephen et al. (2021), suggested a machine learning model, which is based on a convolutional neural network architecture, to detect the surface defects on ceramic tiles. They remark that their model achieved an approximate classification accuracy of 0.9943. Also, Ighe et al. (2021) proposed an optimised quality control approach for ceramic tiles produced in smaller businesses. They used digital image processing to inspect the dimensions, flatness, and other design defects on ceramics. Similarly, Dong et al. (2022) developed a rapid detection method especially for the surface defects on small mosaic type of ceramic tiles.

Research gap and motivation of the study

When the studies in the literature concerning ceramics are examined, it is seen that a significant part of the modelling studies focus mostly on deformation defect occurring on industrial ceramics and defect detection through different methods. Moreover, there is no study concerning the reasons behind the defect occurrences for different defect types on ceramic tiles and using a machine learning method. However, in addition to the deformation defect, there are other types of defects that are considerably important in the production of ceramic tiles. For example, while discoloration defect may not affect the quality of industrial ceramics very drastically, it is a very serious problem for ceramic tiles, which reduces the quality. In another example, industrial ceramics may need to be produced to withstand extreme temperatures for purposes such as coating the interior of industrial kilns; however, ceramic tiles are not expected to be resistant to such extreme temperatures.

Prices of ceramic tiles are determined by considering the types of the defects on the products. The products having defects are sold in the market as low-quality tiles for lower prices, which means loss of income for tile producers. From another point of view, the defects also mean waste of the raw materials, the energy, and time in enterprises.

Thus, the aim of this study is to provide an analysis and suggestions to prevent the defects on the ceramic tiles by examining the production conditions and the stages that cause various types of defects. The analysis also includes a comparison and a selection process between a conventionally used model, logistic regression; and a comparatively new modelling approach, Bayesian networks.

Materials and methods

Logistic regression

Logistic regression and linear regression are similar. The main goal is to fit a straight line in both. In linear regression, the fitted straight line is used to construct a functional relationship between the explanatory and the dependent variables. In logistic regression, however, the fitted line is used to separate the variables in the data using a logistic function. Thus, while the value of the continuous dependent variable is predicted in linear regression, the values of categorical variables are predicted in logistic regression. Logistic regression can also be defined as a classification algorithm used to assign observations to discrete set of classes. The simple form of the logistic regression is the case where the dependent variable is dichotomous. This type of logistic regression is called binary logistic regression. When there are multiple explanatory variables in a data set, the probability of the outcome occurrences is estimated with the following equation.

$$P(Y) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon_i)}} \tag{1}$$

The ratio $P(Y)$ given in Equation (1) models the probability of an outcome based on individual characteristics. However, what is really modelled is the logarithm of the probability presented in the following equation.

$$log\left(\frac{P(Y)}{1 - P(Y)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon_i \tag{2}$$

For detailed information about logistic regression, Hosmer et al. (2013) can be referred.

Logistic regression is a model, which has been used conventionally for classification of multiple explanatory variables. However, because it is an additive model, it has some restrictions such as, only the relations between the dependent and explanatory variables can be considered and any possible relations among the explanatory variables cannot be examined. This problem can be exceeded by using a non-additive model like a Bayesian network, in which it is possible to observe the relations between the independent variable and the explanatory variables, as well as the multilateral relations among the explanatory variables.

Bayesian networks

Bayesian network is a type of machine learning, which is a branch of artificial intelligence. The main idea of Bayesian networks is to reflect the conditional probabilistic relations of the variables to a graphical model. Bayesian networks, which were first introduced by Pearl (1985), have two main parts as the graphical and the probabilistic structures. The graphical structure of the Bayesian networks is based on a visual architecture called directed acyclic graph (DAG). The structure of a DAG consists of nodes and directed arrows (edges), which connect the nodes. The nodes represent the variables, and the arrows indicate the conditional probabilistic relations among the nodes. Estimations of the conditional probabilistic relations in Bayesian networks are obtained by creating tables called conditional probability tables (CPT), which form the probabilistic structure of the Bayesian networks.

The term acyclic in the name of the DAG structure means that the path of an arrow directed from any node, must not allow returning to the same node. A node from which an arrow is directed to another node is called the parent node; and the node the arrow arrives is called the child node. Moreover, nodes having no parent nodes are called root nodes, and nodes having no child nodes are called leaf nodes. An example of a Bayesian network structure is given in Figure 1.

The Bayesian network in Figure 1 consists of six nodes with the titles A, B, C, D, E, and F. For example, node A is the parent node of nodes E, B, and C, whereas nodes B, F, and D are the child nodes of node C. Node A is a root node because it does not have any parent node and node D is a leaf node because it does not have any child node.

The composite probability function of a Bayesian network can be given as follows.

$$P_B(X_1, \ldots, X_n) = \prod_{i=1}^{n} P_B(X_i \backslash \Pi_{X_i}) \tag{3}$$

where B denotes the Bayesian network, X_i the nodes in the network, and n the number of nodes in the network. The first step in the construction of Bayesian networks is to determine the approach to be used to form its graphical structure. There are two approaches to determining the network structure of the Bayesian networks. The first approach is based on expert opinion. In this kind of approach, the dependency relations among the variables are considered and the directions of the arrows are manually determined by experts. The parent–child relations that are established using this approach are within the scope of cause–effect relationship. Therefore, these kinds of networks

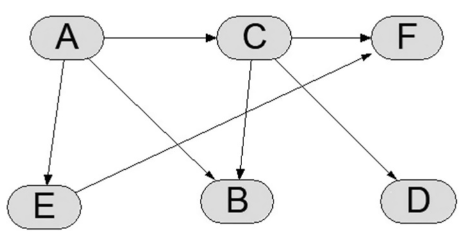

Figure 1. Bayesian network.

are called causal Bayesian networks. According to the second approach, the creation of the network structure is carried out with various algorithms applied to the data set. In this way, the structure of the network is created based on structure learning. In Bayesian networks, which are created with algorithms based on structure learning, directions of arrows indicate probabilistic dependences. However, they may not indicate cause–effect relationships. Such networks are called Bayesian networks based on structure learning.

Causal Bayesian networks

Causal Bayesian networks are constructed within the framework of expert opinions. The causal Bayesian network structure is advantageous in that, it enables visual expressions of cause–effect relationships among the variables. In this way, the analyses performed using the network become more explanatory by giving results within the framework of causality. However, as the number of the variables increases in some applications, the number of the relationships that need to be established manually increases too. Then, construction of causal Bayesian networks becomes difficult. This is the most important disadvantage of the causal Bayesian networks.

Bayesian networks based on structure learning

The second approach, which is called structure learning, constructs Bayesian networks by forming the network structure using various algorithms with the help of a data set. There are various structure learning methods, which can be generally grouped as score-based, constraint-based, and hybrid methods.

Score-based methods. Score-based methods form the structure of a Bayesian network by forming many candidate networks at the beginning. Then, using some scoring functions, they choose the network structure with the highest score. There are various suggested scoring functions in the literature. Among these functions, Cooper and Herskovits (1992) suggested the K2 score. The K2 algorithm is one of the first algorithms used to create Bayesian networks. Another one is the Bayesian Dirichlet (BD) score, which was suggested by Heckerman et al. (1995). BD score is obtained by additional assumptions on likelihood, equivalence, and probability regarding the structure. The most frequently used scoring functions are Log-Likelihood (LL) by Heckerman et al. (1995), Akaike Information Criterion (AIC) by Akaike (1973), and Bayes Information Criterion (BIC) by Schwarz (1978).

Let B denote a Bayesian network, let $|B|$ denote the number of parameters in the network, let T denote the data set, and let $f(N)$ be a nonnegative penalty function. Then, LL scoring function can be given as follows.

$$LL = (B\backslash \mathrm{T}) = \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{q_i} \sum\nolimits_{k=1}^{r_i} N_{ijk} \log\left(\frac{N_{ijk}}{N_{ij}}\right) \tag{4}$$

where n denotes the number of nodes in the network; q_i, the number of possible configurations in the network; r_i, the number of levels in the node X_i; N_{ijk}, the total number of observations, and N_{ij}, the number of instances in the data T.

Let φ be a scoring function, which is given as follows.

$$\varphi(B\backslash T) = LL(B\backslash T) - f(N)|B| \tag{5}$$

When $f(N) = 1$ in Equation (5), AIC scoring function is obtained as follows.

$$AIC(B\backslash T) = LL(B\backslash T) - |B| \tag{6}$$

If $f(N) = 1/2\log(N)$ in Equation (5), BIC scoring function can be calculated as follows.

$$BIC(B\backslash T) = LL(B\backslash T) - 1/2\log(N)|B| \tag{7}$$

Constraint-based methods. Verma and Pearl (1990) proposed the fundamentals of the constraint-based methods. When estimating a Bayesian network structure, constraint-based methods begin with a complete graphical form. After performing some conditional independence tests among the nodes, the arrows among the nodes that do not have any conditional probability relations, are eliminated According to Triantafillou et al. (2014), the constraint-based methods have two disadvantages. The first is that the application time increases exponentially with the number of variables and the other is that their reliability is low. This is because small changes in input can create large changes in the structure of the Bayesian networks and cause errors in independence tests. Some examples of constraint-based methods can be given as Max-Min Parents, and Children (MMPC) algorithm developed by Tsamardinos et al. (2006), Grow Shrink (GS) algorithm by Margaritis (2003), Fast Incremental Association Markov Blanket (Fast-IAMB) algorithm by Yaramakala and Margaritis (2005), Inter-Leaved IAMBD (Inter-IAMBD) algorithm by Tsamardinos et al. (2003), and Peter-Clark (PC) algorithm by Spirtes and Glymour (1991).

Hybrid methods. Hybrid methods combine the properties of the constraint-based methods and the score-based methods in learning Bayesian network structure. The first step of the hybrid methods is the same as the constraint-based methods, and in this step, the structures of the Bayesian networks are created using independence tests. In the second step, the scoring functions are applied to the network structures created in the first step. In the literature, there are many studies suggesting various hybrid methods to build Bayesian networks, such as Singh and Valtorta (1995), who developed a hybrid learning method that integrates the conditional independence tests and expert-opinion-based network structures. Dash and Druzdzel (1999), on the other hand, proposed a method based on scoring with a Bayesian metric, after searching the area consisting of equivalent classes of basic graphs and using conventional heuristic constraint-based algorithms. Acid and De Campos (2001) developed a scoring-metric based on contradictions and an intuitive search strategy that emphasises model complexity and accuracy. Tsamardinos et al. (2006) introduced the Max-Min Hill Climbing (MMHC) algorithm. This algorithm combines the score-based and the constraint-based methods. Another hybrid method. The Separation and Reunion (SAR) algorithm, was developed by Liu et al. (2017). The algorithm decomposes the undirected independent plots for a full set of nodes in the separation phase. Then, it applies a score-based method to build a DAG structure.

Results and discussion

Data and variables

The data were taken from a private ceramic factory. The data were derived from the quality control records kept in the factory by the quality control department. A data set including 1000 cases, which belong to the defective ceramics, were used. Among these defective ceramics, 70 burr, 114 crack, 64 drip, 74 deformation, 134 pinhole, 67 pitting, 79 glaze blistering, 122 colour tone, 64 black spot, 72 plucking, 58 tear, and 82 surface defects were observed. The research includes 14 variables determined by the expert opinion as the factors of production that are likely to affect the defect occurrences on ceramic tiles. These variables are size, glaze density, band speed, press surface moisture, glaze weight, engobe weight, engobe density, clay type, printing colour, drying temperature, firing time, firing temperature, post-drying moisture, and shift variables. In addition to these variables, there is another variable named defect type, which represents the 12 different defect types observed on ceramic tiles. These defect types are burr, crack, drip, deformation, pinhole, pitting, glaze blistering, colour tone, black spot, plucking, tear, and surface defect. In this study, the effects of the 14 variables on the defect type variable were examined.

With the help of the expert opinion, the levels of the numerical variables, which are glaze density, band speed, press surface moisture, glaze weight, engobe weight, engobe density, drying temperature, firing time, firing temperature, and post-drying moisture, were formed by dividing their observations into two or three categories. The levels of the categorical variables size, clay type, printing colour, shift, and defect type, however, were formed by their own categories.

Glaze is a substance that is applied on ceramic tiles to add shine; and engobe, which can be defined as a layer between clay and glaze, plays a role in hiding the colour of the underlying clay. The variables used in this study are listed in Table 1 with their descriptions, levels, and units of measurement.

The constructed logistic regression model

The logistic regression model was built by using the 15 variables presented in Table 1. Weka software (Frank et al., 2021) was employed to construct the logistic regression model of the defect type variable on the independent variables glaze density, band speed, press surface moisture, glaze weight, engobe weight, engobe density, clay type, printing colour, drying temperature, firing time, firing temperature, post-drying moisture, and shift. Receiver operating characteristic (ROC) curves can be used to measure the estimation performances of the constructed models (Hand, 1997). Area under the curve (AUC) value, which can be defined as the area under the ROC curve, is a score ranging from zero to one. If the AUC value is one, the estimation is error-free. The estimation performance of the logistic regression model, for each type of defect, is presented in Table 2 with the corresponding AUC scores sorted in a descending order.

When the AUC scores in Table 2 are examined, it is seen that the logistic regression model estimated the probability of the deformation defect occurrence with the highest rate of 67%. The probabilities of the glaze blistering and the drip defects were estimated with the second highest success level of 63% for each. The colour tone defect and the surface defect, however, are the least successfully estimated defect types with the rates of 48% and 46%, respectively.

The constructed Bayesian network model

To examine the relationships among the factors, the materials, the ceramic types, and the ceramic defect types, three different Bayesian network models were built in GeNIe (Druzdzel & Sowinski, 2019) software applying Bayesian Search, Peter-Clark, and Greedy Thick Thinning algorithms. The parameter settings of the experimented algorithms are presented in Table 3.

Model selection can be made with the LL values given by Equation (4). The algorithms experimented and the corresponding LL values are given in Table 4.

In Table 4, it is seen that −7340.11 is the largest LL value, which belongs to Bayesian Search algorithm. Thus, the Bayesian network estimated by the Bayesian Search algorithm was selected as the model to be used in the analysis. The Bayesian network model, which was constructed with the help of GeNIe (Druzdzel & Sowinski, 2019) software, was rearranged in Netica (Norsys Software Corp, 2019) software and given in Figure 2. The scores behind the bars are percentages, which show the probabilities of the corresponding levels.

To evaluate the estimation performance of the constructed Bayesian network for each type of defect, the calculated AUC scores are presented in Table 5 in a descending order.

Given the AUC scores in Table 5, the constructed Bayesian network model estimated the crack, the pinhole, and the colour tone defects with the highest probability of 0.86 for each. The next most accurate estimate belongs to the glaze blistering defect with a success rate of 0.83. The third most successful estimate, however, was made for both the plucking defect and the surface defect with a rate of 0.82 for each. The pitting defect, however, was the least accurately estimated type of defect with the lowest probability of 0.77. In general, it is possible to conclude that the performance of the Bayesian network model in estimating the defect types is at a satisfactory level for each of the defect types.

To make a comparison between the estimation performances of the constructed Bayesian network and the logistic regression models, the AUC scores of both models are presented side by side in Table 6.

Table 1. Variables used in the study, their descriptions, and levels.

Variable	Description	Levels	
Size	Indicates the two types of the ceramic tiles	Wall tile	
		Floor tile	
Glaze Density	Indicates the density of the glaze substance on the ceramic tiles (g/l)	0-1780	Low
		1780–1800	Moderate
		1800 and over	High
Band Speed	Indicates the number of ceramic tiles passing on the production band per minute (number/min)	0-50	Low
		50 and over	High
Press Surface Moisture	Indicates the amount of moisture at the surface of the press during the pressing process (%)	0–0.05	Low
		0.05–0.06	Moderate
		0.06 and over	High
Glaze Weight	Indicates the weight of the glaze substance on the ceramic tiles (g)	0-64	Low
		64-80	Moderate
		80 and over	High
Engobe Weight	Indicates the weight of the engobe substance on the ceramic tiles (g)	0-65	Low
		65-70	Moderate
		70 and over	High
Engobe Density	Indicates the density of the engobe substance on the ceramic tiles (g/l)	0-1760	Low
		1760–1810	Moderate
		1810 and over	High
Clay Type	Indicates the two different clay types used in the production	Dry grinding	
		Spray	
Printing Color	Indicates the two main printing color types used to color the ceramic tiles	Light coloured	
		Dark coloured	
Drying Temperature	Indicates the temperature that the ceramic tiles are subjected to during the drying process (°C)	0-255	Low
		255-267	Moderate
		267 and over	High
Firing Time	Indicates the duration of time when the ceramic tiles are kept in the kiln during the firing process (min)	0-40	Low
		40 and over	High
Firing Temperature	Indicates the temperature that the ceramic tiles are subjected to in the kiln during the firing process (°C)	0-1118	Low
		1118–1124	Moderate
		1124 and over	High
Post-Drying Moisture	Indicates the percentage of the moisture in the ceramic tiles after leaving the drying unit (%)	0–0.40	Low
		0.40–0.55	Moderate
		0.55–0.80	High
Shift	Indicates the shift hours in the factory	16:00-24:00	
		24:00-08:00	
		08:00-16:00	
Defect Type	Indicates the 12 defect types observed on the ceramic tiles	Burr	
		Crack	
		Drip	
		Deformation	
		Pinhole	
		Pitting	
		Glaze blistering	
		Color tone	
		Black spot	
		Plucking	
		Tear	
		Surface defect	

When the AUC scores presented in Table 6 are examined, it is possible to see that the AUC scores of the Bayesian network model are considerably higher than the ones of the logistic regression model for all defect types. Thus, it is possible to conclude that the Bayesian network model performed quite better than the logistic regression model in estimating the probabilities of the defect types.

The most important thing to consider when comparing Bayesian network and logistic regression models is that the Bayesian network is a graphical model, while the logistic regression model is an additive model. Therefore, both models work with different estimation mechanisms compared to each other. Due to its additive mechanism, logistic regression generates an estimate for the

Table 2. The AUC scores of the constructed logistic regression model.

Defect type	AUC Scores
Deformation	0.67
Glaze blistering	0.63
Drip	0.63
Pinhole	0.62
Plucking	0.62
Tear	0.60
Black spot	0.58
Crack	0.53
Burr	0.52
Pitting	0.51
Color tone	0.48
Surface defect	0.46

Table 3. The parameter settings of the Bayesian network algorithms.

	Bayesian Search	Peter-Clark	Greedy Thick Thinning
Discrete threshold	20	20	20
Maximum parent count	8	-	8
Maximum adjacency size	-	8	-
Iterations	20	-	-
Significance threshold	-	0.05	-
Sample size	50	-	-
Seed	0	-	-
Link probability	0.1	-	-
Prior link probability	0.001	-	-
Maximum time (seconds)	0	0	-

Table 4. Log-likelihood (LL) values.

Algorithm	Log-Likelihood (LL)
Bayesian Search	−7340.11
Peter-Clark	−7705.80
Greedy Thick Thinning	−8246.99

dependent variable by calculating the sum of the corresponding explanatory variables, which are reduced to the minimum number until they become independent from each other, and the coefficients of these variables determined by various methods. Thus, when estimating with the logistic regression model, it is assumed that the independent explanatory variables do not affect each other, but only the dependent variable. Because of this requirement, it is obvious that this estimation mechanism cannot produce very good results for situations where explanatory variables also interact with each other. In Bayesian networks, there are no assumptions on explanatory variables such as being independent of each other or being as few as possible. For instance, in our article concerning the defects on the ceramic tiles, the size and engobe weight variables are clearly dependent because as the ceramic size increases, the engobe weight that must be used will increase. Still, it is not a problem to include these two variables in the Bayesian network model and to estimate the probability of failure on ceramic tiles with the help of these variables. This feature is one of the most important advantages of Bayesian networks over the logistic regression model. In addition, unlike the logistic regression, distributional assumptions are not required in Bayesian network models. Moreover, once a logistic regression model has been constructed, the addition or extraction of an explanatory variable may endanger the significance of the model. On the other hand, this process usually does not lead to drastic changes in a Bayesian network model. Furthermore, unlike the logistic regression model, the Bayesian network model also has the superiority of examining the

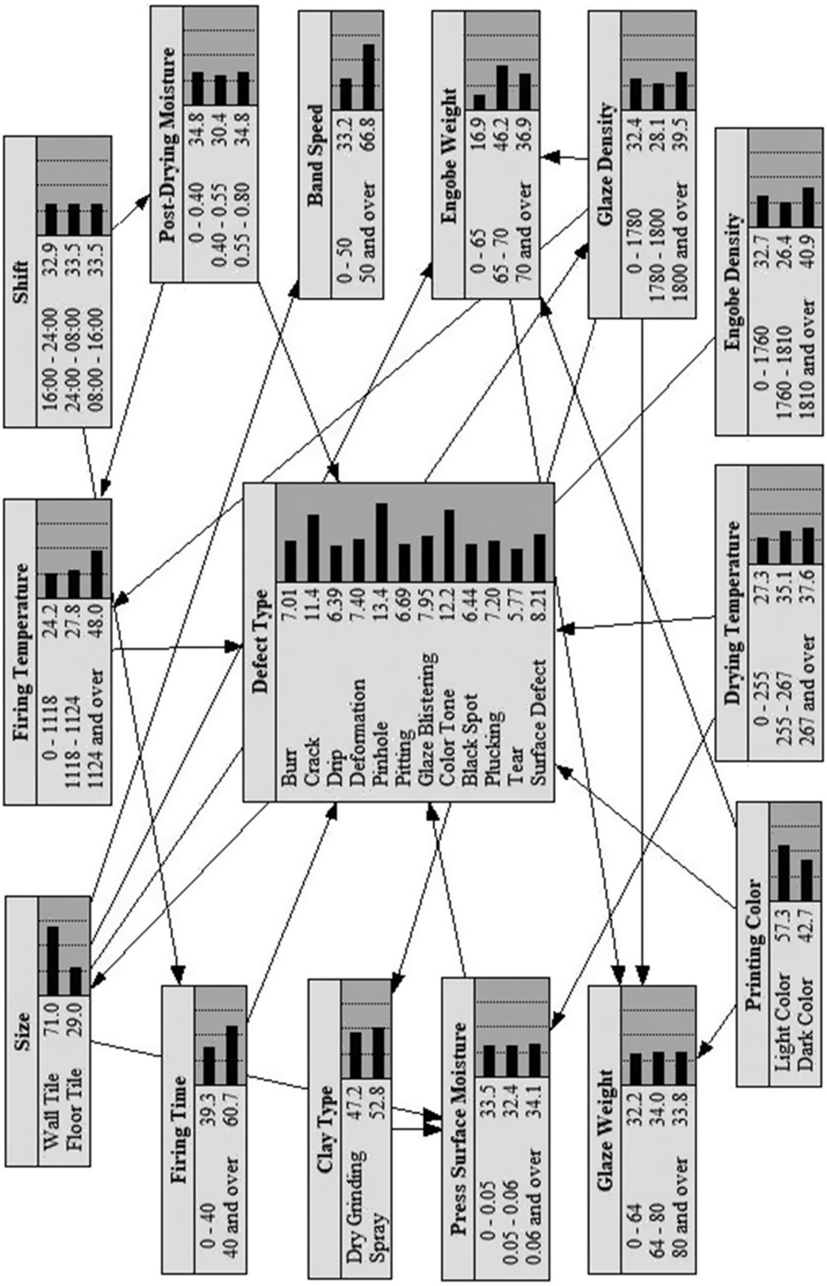

Figure 2. The constructed Bayesian network model based on Bayesian search algorithm.

Table 5. The AUC scores of the constructed Bayesian network model.

Defect type	AUC Scores
Crack	0.86
Pinhole	0.86
Colour tone	0.86
Glaze blistering	0.83
Plucking	0.82
Surface defect	0.82
Burr	0.81
Deformation	0.81
Drip	0.80
Black spot	0.80
Tear	0.79
Pitting	0.77

Table 6. AUC scores of the Bayesian network model and the logistic regression model.

	AUC Scores	
Defect type	The Bayesian network model	The logistic regression model
Crack	0.86	0.53
Pinhole	0.86	0.62
Colour tone	0.86	0.48
Glaze blistering	0.83	0.63
Plucking	0.82	0.62
Surface defect	0.82	0.46
Burr	0.81	0.52
Deformation	0.81	0.67
Drip	0.80	0.63
Black spot	0.80	0.58
Tear	0.79	0.60
Pitting	0.77	0.51

bilateral or multilateral dependencies among the variables simultaneously, which allows a more complex analysis.

As for the disadvantages of Bayesian network models, the most important disadvantage of Bayesian network models against logistic regression models is that there is not a generally accepted standard approach to the creation of Bayesian network models. In addition, the construction of Bayesian network models is more difficult than the construction of logistic regression models. However, due to their less complex structures, logistic regression models are easier to build and interpret compared to Bayesian networks

Sensitivity analysis

Entropy, which is commonly used in information theory, was introduced by Shannon (1948). In Bayesian networks, it is possible to perform a sensitivity analysis to determine to what extent the levels of any variable are affected by the change in the levels of other variables. Entropy score, which shows the irregularity of a system, can be used for sensitivity analysis. With the help of entropy score, it is possible to see how likely an unexpected situation will occur in a system.

Let X be a discrete random variable with possible values $x_1, \ldots, x_n$ and $P(X)$ be its probability mass function. Then, entropy score $H(X)$ can be calculated with the following equation.

$$H(x) = -\sum_{i=1}^{n} P(x_i) log_b P(x_i \tag{8}$$

Table 7. The entropy scores of the factors causing the defects on the ceramic tiles.

Variables	Entropy
Firing Temperature	0.03670
Drying Temperature	0.03549
Press Surface Moisture	0.03372
Post-Drying Moisture	0.03303
Printing Color	0.01940
Glaze Density	0.01325
Firing Time	0.01285
Size	0.01059
Glaze Weight	0.00869
Band Speed	0.00862
Engobe Density	0.00822
Clay Type	0.00753
Engobe Weight	0.00748
Shift	0.00006

where *b* is the base of the logarithm that is used? The entropy scores for the levels of the defect type variable with respect to the other variables are given in Table 7.

Table 7 presents the factors causing various defects on the ceramic tiles and the corresponding entropy scores sorted in a descending order. According to the entropy values, the defects on the ceramic tiles are mostly affected by the firing temperature. The second most effective factor appears to be the drying temperature, while the third important factor is emerging as the press surface moisture. The finding that the firing temperature is the most effective factor causing the defects is in accordance with the finding by Jin et al. (2017) that the most important factor leading to shell deformation on industrial ceramics is firing temperature.

Analyses regarding the defect types

To find the most likely causes that lead to an increase in a defect type, first, the related defect type should be selected on the Bayesian network as evidence by marking it. Then, the corresponding conditional probabilities of the factor levels, which are calculated by the Bayesian network, are examined and the ones with the highest probabilities are recorded as the most likely production settings that cause the defect type under consideration.

To find the suggested production settings that decrease the occurrence probability of a certain defect type, however, it is reasonable to select and try the other settings of the most likely factors causing the related defect. For instance, if the most likely cause of a certain defect type appears to be the high levels of a production factor, then the low or moderate levels of this production factor should be tried by selecting and marking them on the Bayesian network as evidence, respectively. Afterwards, the occurrence probability of the defect type, which is calculated by the Bayesian network, should be examined. Finally, the production factor level causing the smallest defect probability can be suggested as the optimal production setting for the defect type under consideration.

The most likely causes and the suggested production settings, which were provided by the estimated Bayesian network model for all types of defects, are presented briefly in Table 8.

In addition, some factors that are not included in Table 8 but are also effective in the occurrence of some defects and some remarkable results from the analyses are given for each defect type in the following paragraphs.

It is evident that the probability of the burr defect occurrence increases also when the engobe weights are at the medium levels and the engobe densities are at the high levels. Additionally, the burr defect is more commonly seen on the light-coloured ceramics and the ceramics being produced using spray type of clay.

The crack defect is the third most common type of defect observed on the ceramic tiles. The high levels of the glaze weight variable were also determined to be among the factors increasing the occurrence of the crack defect. On the contrary, the low levels of the press surface moisture and the moderate levels of the engobe weight seem to be increasing the emergence of this defect type. In addition, the crack defect is more likely to occur on the light-coloured ceramics. Likewise, when the spray type of clay is used, the risk of the crack defect occurrence increases.

The analysis results suggest that the medium levels of the firing and the drying temperatures also play a role in increasing the drip defect occurrences. Unlike the other defect types, the drip defect is more likely to occur at the low levels of the glaze weight. Moreover, unlike the burr and the crack defects, the drip defect is also more likely to be observed when dry grinding type of clay is used. On the other hand, it appears that the drip defect is more frequently seen on the light-coloured ceramic tiles, just like the burr and the crack defects.

The low levels of the engobe density and the press surface moisture also appear to increase the deformation defect occurrence. It is also observed that the deformation defect is more likely to occur on the light-coloured tiles.

When the analysis results regarding the pinhole defect are examined, it is possible to say that pinhole is the most common defect type observed on the ceramic tiles. In addition, the pinhole defect is observed more frequently at the low levels of the post-drying moisture factor. Another remarkable result is that the pinhole defect tends to occur at approximately equal rates, when the tiles are produced using dry grinding or spray type of clay.

Table 8. The most likely causes of the defects and the suggested settings.

Defect type	Most likely causes	Suggested settings
Burr	The high levels of the firing temperature (47.5%)	0–1118°C (the low levels)
	The high levels of the press surface moisture (45.9%)	0–0.05 (the low levels)
	The low levels of the post-drying moisture (37.8%)	0.55–0.80 (the high levels)
Crack	The high levels of the firing time (69.8%)	0–40 min (the low levels)
	The high levels of the firing temperature (59.8%)	0–1118°C (the low levels)
	The high levels of the post-drying moisture (48.5%)	0–0.40 (the low levels)
Drip	The high levels of the firing time (53.1%)	0–0.40 min (the low levels)
	The moderate levels of the engobe weight (51.8%)	70 g and over (the high levels)
	The moderate levels of the press surface moisture (50%)	0.06 and over (the high levels)
Deformation	The high levels of the firing temperature (65.7%)	0–1118°C (the low levels)
	The low levels of the firing time (50.4%)	40 min and more (the high levels)
	The low levels of the glaze density (44.6%)	1780–1800 g/l (the moderate levels)
Pinhole	The high levels of the firing time (63.2%)	0–40 min (the low levels)
	The high levels of the drying temperature (59.8%)	0–255°C (the low levels)
	The moderate levels of the engobe weight (48.1%).	70 g and over (the high levels)
Pitting	The high levels of the firing time (64.9%)	0–40 min (the low levels)
	The high levels of the firing temperature (48.8%)	1118–1124°C (the moderate levels)
	The low levels of the engobe weight (46.7%)	70 g and over (the high levels)
Glaze blistering	The high levels of the firing time (59.6%)	0–40 min (the low levels)
	The moderate levels of the engobe weight (48.9%)	70 g and over (the high levels)
	The high levels of the engobe density (44.1%)	0–1760 g (the low levels)
Color tone	The high levels of the firing time (65.6%)	0–40 min (the low levels)
	The high levels of the firing temperature (50.4%)	1118–1124°C (the moderate levels)
	The moderate levels of the drying temperature (44.9%)	267°C and over (the high levels)
Black spot	The high levels of the firing time (57.8%)	0–40 min (the low levels)
	The moderate levels of the engobe weight (47.6%)	70 g and over (the high levels)
	The high levels of the firing temperature (45.2%)	0–1118°C (the low levels)
Plucking	The high levels of the firing temperature (57.6%)	1118–1124°C (the moderate levels)
	The high levels of the press surface moisture (57.6%)	0–0.05 (the low levels)
	The high levels of the firing time (54.2%)	0–40 min (the low levels)
Tear	The moderate levels of the engobe weight (51.9%)	70 g and over (the high levels)
	The low levels of the firing time (50.6%)	40 min and over (the high levels)
	The high levels of the engobe density (46.1%)	0–1760 g/l (the low levels)
Surface defect	The high levels of the firing time (67.6%)	0–40 min (the low levels)
	The firing temperature (54.3%)	0–1118°C (the low levels)
	The moderate levels of the engobe weight (44.5%)	0–65 g (the low levels)

It appears that keeping the post-drying moisture at the low levels slightly increases the probability of the pitting defect occurrence. Like the pinhole defect, the pitting defect occurs with approximately equal probabilities, when dry grinding or spray type of clay is used during the production.

Unlike the other defect types, the glaze blistering defect is the only defect type that is more likely to be observed on the dark coloured ceramic tiles. Moreover, the glaze blistering defect is also among the defect types that occur approximately equally in the use of dry grinding or spray type of clay.

According to the results of the analysis, the colour tone defect is the second most common defect type occurring on the ceramic tiles. The low levels of the post-drying moisture, the engobe density, and the engobe weight also increase the colour tone defect. Additionally, the colour tone defect is more likely to occur on the light-coloured ceramic tiles.

Considering the factors leading to the black spot defect, it appears that also the press surface moisture and the engobe density variables have increasing effects on the occurrence of this defect type. However, unlike the other defect types, it turns out that the low levels of any variables do not lead to a black spot defect. It is also seen the black spot defect is more likely to be observed on the light-coloured ceramics.

The high levels of the glaze weight and the drying temperature factors are among the causes increasing the probability of observing the plucking defect. Similarly, the moderate levels of the engobe weight variable also increase the occurrence probability of the plucking defect. In addition, the plucking defect is more likely to be observed on the light-coloured ceramics.

Speaking of the tear defect, the moderate levels of the post-drying moisture variable appear to be among the causes increasing the tear defect formation. Unlike the other defect types, the tear defect is more likely to occur at the low levels of the firing temperature. This defect type is observed on the light-coloured ceramics more frequently. In addition, use of dry grinding type of clay is another factor increasing the formation of the tear defect.

The surface defect seems to be increased by the high levels of the variables except for one variable. The only variable whose low levels increase the probability of observing the surface defect appears to be the post-drying moisture. Moreover, the surface defect is more likely to occur on the light-coloured ceramics.

In addition to the analyzes and recommendations given specifically for each error type in the above lines, some general comments, comparisons, and recommendations that will apply to all the defect types can be given as follows.

According to the findings obtained from the analyses, it is observed that the high levels of the band speed significantly increase the probability of every defect type except the deformation defect. Although, the clay-type factor has partial effects on the defects in general, the spray type of clay appears to cause more defects than the dry grinding type of clay. Another result shows that the light-coloured ceramics are more exposed to all kinds of defects except for the glaze blistering. This difference is especially evident in the black spot, the colour tone, the deformation, the drip, and the burr defects. The effect of the shift hours in the factory, however, was found to be insignificant on the occurrence of the defects. Additionally, the high levels of the glaze density increase all the defect types but the deformation defect. As far as the firing temperature is concerned, it quite significantly increases all the defects except the tear defect. Moreover, firing temperature is the most effective factor causing the defects on the ceramic tiles. This result is parallel to the findings of Jin et al. (2017) who report that the most important factor leading to deformation on industrial ceramic shells was found to be firing temperature. Firing time variable gives similar results too. The high levels of the engobe weight and the engobe density variables were found to be increasing the defects in general. In addition, the high levels of the glaze density also increase the defects. Moreover, it was observed that the high levels of the drying temperature variable increase some of the defects, yet they have an adverse effect on some defect types.

Finally, based on the findings summarised above, it can generally be said that it would be a logical approach not to keep the band speed at high levels to prevent the defects. In addition, it should not be preferred to keep the ceramic tiles in kilns at high temperatures and for long times. Keeping the firing temperatures and the firing times at optimum levels will be effective in reducing the defects too. The same interpretation regarding the firing time and temperature can be made also for the glaze weight variable.

In a ceramic factory, production settings need to be adjusted frequently, in accordance with the needs, amounts, properties, and the qualities regarding the outputs. To determine the optimum production line settings leading to minimum levels of the ceramic tile defects, trial and error method would be quite cumbersome, time consuming, and costly. Instead, referring to artificial intelligence methods, like machine learning, will solve this problem faster and more efficiently. Our study shows that the estimated Bayesian network model, which is based on machine learning, performed better than the logistic regression model, which is a classical method. Thus, when the estimated Bayesian network model is run by appropriate software, the optimal settings can be determined easily and quickly, by propagating the evidences through the network, depending on the production needs or plans of the factory.

Conclusions

In conclusion, in this study, the possible reasons of the defects encountered on the ceramic tiles were investigated and analysed, after building two candidate models and selecting the most appropriate of them. Before the analysis part, a logistic regression model, which has been conventionally used in the literature and a Bayesian network model, which is comparatively a new approach, were constructed. When the model validation results of the Bayesian network model and the logistic regression model were examined by comparing their AUC scores, it appeared that the Bayesian network model was more successful in estimating the defect types. Indeed, Bayesian network models are very flexible and advantageous models compared to the conventional additive models like regression. Unlike in regression models, every variable in Bayesian networks can affect each other simultaneously. Thus, instead of using logistic regression, preferring a Bayesian network model is more convenient to analyse the factors causing defective tiles, as these factors usually interact with each other at the different stages of the production.

After the model comparison stage, using the constructed Bayesian network, three different types of analyses were performed. Firstly, various types of defects caused by various production conditions were examined. Secondly, the probabilities of the most likely causes of the defect types were estimated. Afterwards, with the help of the calculated entropy scores, the effect degrees of the production factors on the defect types were determined. Finally, various suggestions to prevent the defects were provided. As a continuation of this study, it is possible to examine the causes of the same or other kinds of defects seen on ceramic tiles by using other kinds of models and compare them with the results obtained in this study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was not funded by anyone or any establishment.

ORCID

Volkan Sevinç http://orcid.org/0000-0003-4643-443X
Meryem Merve Kırca http://orcid.org/0000-0001-5932-768X

References

Acid, S., & De Campos, L. M. (2001). A hybrid methodology for learning belief networks: BENEDICT. *International Journal of Approximate Reasoning*, *27*(3), 235–262. https://doi.org/10.1016/S0888-613X(01)00041-X

Agarwal, S., & Rao, P. V. (2005). A new surface roughness prediction model for ceramic grinding. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, *219*(11), 811–819. https://doi.org/10.1243/095440505X32832

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csáki (Eds.), *2nd international symposium on information theory* (pp. 267–281). Akadémia Kiadó.

Chen, X., Li, D., Wu, H., Tang, Y., & Zhao, L. (2011). Analysis of ceramic shell cracking in stereolithography-based rapid casting of turbine blade. *International Journal of Advanced Manufacturing Technology*, *55*(5), 447–455. https://doi.org/10.1007/s00170-010-3064-x

Cooper, G. F., & Herskovits, E. (1992). A Bayesian method for the induction of probabilistic networks from data. *Machine Learning*, *9*(4), 309–347. https://doi.org/10.1007/BF00994110

Dash, D., & Druzdzel, M. J. (1999, July 30August 1). A hybrid anytime algorithm for the construction of causal models from sparse data. *Fifteenth Conference on Uncertainty in Artificial Intelligence*, Stockholm, Sweden.

Dong, G., Sun, S., Wu, N., Chen, X., Huang, P., & Wang, Z. (2022). A rapid detection method for the surface defects of mosaic ceramic tiles. *Ceramics International*, *48*(11), 15462–15469. https://doi.org/10.1016/j.ceramint.2022.02.080

Druzdzel, M. J., & Sowinski, T. (2019). *GeNIe* (Version 2.5) [Computer software]. http://www.bayesfusion.com/

Everhart, W., Lekakh, S., Richards, V., Chen, J., Li, H., & Chandrashekhara, K. (2013). Corner strength of investment casting shells. *International Journal of Metalcasting*, *7*(1), 21–27. https://doi.org/10.1007/BF03355541

Everhart, W. A., Lekakh, S. N., Richards, V. L., Smith, J., Li, H., Chandrashekhara, K., Zhao, H., & Nam, P. (2012). Foam pattern aging and its effect on crack formation in investment casting ceramic shells. *Proceedings of American Foundry Society (AFS) Conference* Nashville, Tennessee, USA, 1–8.

Frank, E., Hall M. A., & Witten I. H. (2021). Weka (Version 3.8.6) [Computer software]. https://www.cs.waikato.ac.nz/ml/weka/

Hand, D. J. (1997). *Construction and assessment of classification rules*. John Wiley & Sons.

Hanzaei, S. H., Afshar, A., & Barazandeh, F. (2017). Automatic detection and classification of the ceramic tiles' surface defects. *Pattern recognition*, *66*, 174–189. https://doi.org/10.1016/j.patcog.2016.11.021

Heckerman, D., Geiger, D., & Chickering, D. M. (1995). Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning*, *20*(3), 197–243. https://doi.org/10.1007/BF00994016

Hosmer, D. W., Jr., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons.

Hsu, S. M., & Shen, M. (2004). Wear prediction of ceramics. *Wear*, *256*(9–10), 867–878. https://doi.org/10.1016/j.wear.2003.11.002

Ighe, R., Tamgole, A., & Tiwari, D. (2021). Measuring the dimensions, detecting the defects and measuring flatness of the tiles. *New Arch-International Journal of Contemporary Architecture*, *8*(2), 1053–1059.

Jiang, J., & Liu, X. Y. (2004). Burning-out process of ceramic moulds. *International Journal of Cast Metals Research*, *17*(2), 121–127. https://doi.org/10.1179/136404604225014864

Jin, S., Liu, C., Lai, X., Li, F., & He, B. (2017). Bayesian network approach for ceramic shell deformation fault diagnosis in the investment casting process. *International Journal of Advanced Manufacturing Technology*, *88*(1), 663–674. https://doi.org/10.1007/s00170-016-8795-x

Kim, S. S., Kato, K., Hokkirigawa, K., & Abe, H. (1986). Wear mechanism of ceramic materials in dry rolling friction. *Journal of Tribology*, *108*(4), 522–526. https://doi.org/10.1115/1.3261255

Kong, H., & Ashby, M. F. (1992). Wear mechanisms in brittle solids. *Acta Metallurgica et Materialia*, *40*(11), 2907–2920. https://doi.org/10.1016/0956-7151(92)90455-N

Li, G., Liu, X., Tao, B., Jiang, D., Zeng, F., & Xu, S. (2021). Research on ceramic tile defect detection based on YOLOv3. *International Journal of Wireless and Mobile Computing*, *21*(2), 128–133. https://doi.org/10.1504/IJWMC.2021.120013

Liu, C., Wang, F., Jin, S., Li, F., & Lai, X. (2019). Permafrost analysis methodology (PAM) for ceramic shell deformation in the firing process. *International Journal of Metalcasting*, *13*(4), 953–968. https://doi.org/10.1007/s40962-019-00317-0

Liu, H., Zhou, S., Lam, W., & Guan, J. (2017). A new hybrid method for learning bayesian networks: Separation and reunion. *Knowledge-Based Systems*, *121*, 185–197. https://doi.org/10.1016/j.knosys.2017.01.029

Margaritis, D. (2003). *Learning Bayesian network model structure from data* [Unpublished doctoral dissertation]. Carnegie-Mellon University.

Marshall, D. B., Lawn, B. R., & Evans, A. G. (1982). Elastic/Plastic indentation damage in ceramics: The lateral crack system. *Journal of the American Ceramic Society*, *65*(11), 561–566. https://doi.org/10.1111/j.1151-2916.1982.tb10782.x

Norsys Software Corp. (2019). *Netica* (Version 607) [Computer software]. https://www.norsys.com/netica.html

Pearl, J. (1985). Bayesian networks: A model of self-activated memory for evidential reasoning. *Proceedings of the 7th conference of the Cognitive Science Society* Irvine, California, USA, 15–17.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Singh, M., & Valtorta, M. (1995). Construction of Bayesian network structures from data: A brief survey and an efficient algorithm. *International Journal of Approximate Reasoning*, *12*(2), 111–131. https://doi.org/10.1016/0888-613X(94)00016-V

Spirtes, P., & Glymour, C. (1991). An algorithm for fast recovery of sparse causal graphs. *Social Science Computer Review*, *9*(1), 62–72. https://doi.org/10.1177/089443939100900106

Stephen, O., Maduh, U. J., & Sain, M. (2021). A machine learning method for detection of surface defects on ceramic tiles using convolutional neural networks. *Electronics*, *11*(1), 55. https://doi.org/10.3390/electronics11010055

Ting, B. Y., & Winer, W. O. (1989). Friction-induced thermal influences in elastic contact between spherical asperities. *Journal of Tribology*, *111*(2), 315–322. https://doi.org/10.1115/1.3261916

Triantafillou, S., Tsamardinos, I., & Roumpelaki, A. (2014). Learning neighborhoods of high confidence in constraint-based causal discovery. In L. C. van der Gaag & A. J. Feelders (Eds.), *Probabilistic Graphical Models PGM 2014 Lecture Notes in Computer Science* (Vol. 8754 (: Springer, Cham), pp. 487–502).

Tsamardinos, I., Aliferis, C. F., Statnikov, A. R., & Statnikov, E. (2003). Algorithms for large scale Markov blanket discovery. *FLAIRS conference* St. Augustine, Florida, USA, *2*, 376–380.

Tsamardinos, I., Brown, L. E., & Aliferis, C. F. (2006). The max-min hill-climbing Bayesian network structure learning algorithm. *Machine Learning*, *65*(1), 31–78. https://doi.org/10.1007/s10994-006-6889-7

Verma, T., & Pearl, J. (1990). Equivalence and synthesis of causal models. *Proceedings of the Sixth Annual Conference on Uncertainty in Artificial Intelligence* Cambridge, Massachusetts, USA, 255–270.

Wan, G., Fang, H., Wang, D., Yan, J., & Xie, B. (2022). Ceramic tile surface defect detection based on deep learning. *Ceramics International*, *48*(8), 11085–11093. https://doi.org/10.1016/j.ceramint.2021.12.328

Wang, Y., & Hsu, S. M. (1996). Wear and wear transition modeling of ceramics. *Wear*, *195*(1–2), 35–46. https://doi.org/10.1016/0043-1648(95)06750-7

Wang, K., Li, Z., & Wang, X. (2022). Concatenated network fusion algorithm (CNFA) based on deep learning: Improving the detection accuracy of surface defects for ceramic tile. *Applied Sciences*, *12*(3), 1249. https://doi.org/10.3390/app12031249

Wereszczak, A. A., Breder, K., Ferber, M. K., Kirkland, T. P., Payzant, E. A., Rawn, C. J., Krug, E., Larocco, C. L., Pietras, R. A., & Karakus, M. (2002). Dimensional changes and creep of silica core ceramics used in investment casting of superalloys. *Journal of Materials Science*, *37*(19), 4235–4245. https://doi.org/10.1023/A:1020060508311

Yaramakala, S., & Margaritis, D. (2005). Speculative Markov blanket discovery for optimal feature selection. *Fifth IEEE International Conference on Data Mining*, 809–812. https://doi.org/10.1109/ICDM.2005.134

Younas, M., Nawaz, Q., Hamid, I., Gilani, S. M. M., & Iqbal, M. J. (2022). A novel approach of ceramic tile crack detection using morphological operations. *Mehran University Research Journal of Engineering and Technology*, *41*(2), 146–154. https://doi.org/10.22581/muet1982.2202.14