

Article

# Application of Logistic Regression for Production Machinery Efficiency Evaluation

Anna Borucka \* and Małgorzata Grzelak

Faculty of Security, Logistics and Management, Military University of Technology, 00-908 Warszawa, Poland; malgorzata.grzelak@wat.edu.pl

\* Correspondence: anna.borucka@wat.edu.pl; Tel.: +48-261-837-060

Received: 11 October 2019; Accepted: 5 November 2019; Published: 8 November 2019



**Abstract:** Production companies operate in a complex economic, technological, social and political environment. There are a number of factors contributing to a satisfactory market position, the most important one being a properly defined and implemented strategy. It needs, however, to be continuously monitored and, if necessary, modified. One of the elements subject to such evaluation is the efficiency of the production processes, which has become the genesis of this article. In response to the methods presented in the literature, a proposal using the logistic regression method for this purpose is presented. The dichotomous form of the dependent variable makes it possible to make such an evaluation in an unambiguous manner and to determine the significance and influence of selected factors on the result thereof.

**Keywords:** efficiency; production processes; machinery; production maintenance; logistic regression

## 1. Introduction

Market success of a manufacturing company is shaped primarily by the demand for the manufactured products and the rate of return on capital employed [1,2]. Poor machine efficiency and frequent downtimes can lead to a reduction in production levels, resulting in lost market opportunities, increased operating costs and reduced profits [3]. It is therefore necessary to apply appropriate methods and tools to support management and to organize maintenance services in an adequate manner to ensure that the production system operates at the assumed levels of productivity and efficiency [4].

Effectiveness is an important element in the analysis of the production process [5–7], often considered in scientific publications. It is assessed on the basis of various measures. In practice, numerous mathematical models and tools are used to support the assessment of the performance of machinery. The most frequently used measures for analyzing the efficiency of technical facilities are those resulting from three general models of operation assessment, i.e., the reliability model, the operational efficiency OEE (overall equipment effectiveness) model and the organizational and technical KPI (key performance indicators) model [8]. In addition, methods and tools for its evaluation can be classified in five main areas, i.e., operational, market, financial, technical or dynamic [9,10]. Particularly important from the point of view of machinery efficiency diagnostics is operational efficiency, and the research in this area focuses primarily on the search for opportunities to reduce the consumption of production resources. These include analysis of labor productivity growth, cost reduction, minimization of losses and shortening of production cycles. Studies available in the literature indicate the application of a number of methods and tools in this area, such as methods of productivity and profitability indicators, analysis of efficiency and degree of work stations' utilization, cost calculation of activities, study of spatial efficiency of production organization and economic evaluation of the production structure [11].

Maintaining the company's machinery stock at an appropriate level requires continuous monitoring and evaluation of the adopted effectiveness indicators. A number of companies have MES (manufacturing execution system) systems in place, which enable ongoing control of the above parameters. There are also companies (including those examined by the authors) that do not have such software and therefore proper evaluation of efficiency parameters is difficult. Such analyses are supported by mathematical tools and methods, which also include modeling with the use of logistic regression, as presented in this article. The subject of this research was a plastics manufacturing company, while the main objective was to evaluate the effectiveness of the production process based on selected factors that may significantly affect the level of machinery efficiency. The analysis was carried out on the basis of information on the performance of the company's production system recorded from 1 September 2015 to 31 August 2017.

Monitoring the effectiveness of utilization of the available machinery allows production reserves or waste in the processes underway to be identified [12–14]. The basis for successful assessment is an appropriate selection of measures and indicators. The analysis of literature made it possible to distinguish those which were of the greatest importance both in theoretical and industrial-practical aspects. Three general models should be distinguished:

- operational efficiency model OEE (overall equipment effectiveness),
- reliability model,
- organizational and technical model—KPI [15].

Within the operational efficiency model, a frequently employed parameter (which was monitored in the examined entity as well) is the overall equipment effectiveness (OEE) indicator, which is widely described in the literature [16–19]. The available studies most often present the theoretical aspects of its calculation and indicate the categories of losses that may occur during the process of machinery and equipment use in relation to ideal conditions [19–21]. Analyses are also available to demonstrate the practical implementation of this parameter in manufacturing companies [12,22,23].

The OEE index is a product of three components [23,24], i.e., readiness and efficiency of machinery and quality of the manufactured products. It is therefore a general, comprehensive assessment, most often presented in percentage form. According to Seichi Nakajime from the Japan Institute of Plant Maintenance [25,26], OEE should remain at 85.41%, but it should be stressed that each enterprise operates in a specific environment; thus, this indicator will be different for each entity, depending on its size, profile and industry, and will not take on the same value in two different operating units [9,27]. Therefore, in practice the above indicator has evolved into different forms of application depending on the sector in which a given entity operates, adjusting to the needs of the environment. The following indicators should be mentioned: OFE (overall factory effectiveness), OPE (overall plant effectiveness), OTE (overall throughput effectiveness), PEE (production equipment effectiveness), OAE (overall asset effectiveness) or TEEP (total equipment effectiveness performance) [21].

The reliability model allows measures in statistical terms to be determined, on the basis of a time analysis of the performance of technical facilities. In practice, these refer to the technical condition of machines, as well as to the activities of maintenance staff. These are MTBF (mean time between failures), MTTR (mean time to repair) or MTTF (mean time to failure) [15].

The organizational and technical KPI model includes a set of measures enabling a comprehensive assessment of the efficiency and effectiveness of the implemented processes. It includes 72 indicators classified in three areas: economic (e.g., total relative cost of maintenance), technical (availability of facilities for preventive works) and organizational (number of maintenance staff) [28].

In relation to the analyzed company, indicators associated with the operational effectiveness model, related to efficiency, will be preferable in the context of machinery stock management; therefore, they have become the subject of this analysis. Following the literature in this field [13,29], it was

assumed that efficiency in production processes is the quotient of the actual efficiency to the nominal efficiency, as specified in the following ratio (1):

$$W_Q = \frac{Q_r}{Q_O}, \quad (1)$$

where:

$W_Q$ —machinery operational efficiency indicator,

$Q_r$ —actual (achieved) efficiency (pcs./h),

$Q_O$ —theoretical efficiency, as defined in the technical documentation (pcs./h).

Thus calculated, it indicates the degree of efficient use of the production line for each operation and, as such, indicates areas for improvement. Efficiency is most often presented in percentage form. This does not always allow for its quick and unambiguous assessment. In forecasting studies, it is assumed to be a quantitative variable, which limits the availability of some modeling methods. Therefore, with regard to the analyzed company, according to the authors, a better approach would be to analyze the efficiency from the individual point of view of each company by determining its satisfactory level and reacting only if it is not achieved.

Numerous studies using logistic regression models with regard to machine maintenance are available in the literature. The main objective of the proposed tools is to assess the technical condition of technical objects along with reliability parameters [30,31], predict upcoming failures [32,33] and estimate the service life of machinery [34]. For example, Yan and Lee assessed the performance of an elevator door system in real time and identified the types of possible failures [30]. Kozłowski et al. developed a model classifying the condition of a cutting tool blade and predicting its durability [31]. Lee et al. studied the reliability of a cutting tool using a combination of logistic regression and acoustic emission methods [32]. Caesarendra combined methods of logistic regression and relevance vector machine to evaluate performance degradation and to predict failure times based on simulation and experimental data [33], whereas Chen et al., on the basis of vibration characteristics of cutting tools, developed a universal model enabling the analysis of reliability and performance for machine tools [34].

This article proposes a model of logistic regression to be used for analysis and evaluation of the level of efficiency of executed processes. The research covered the process of manufacturing garbage bags in a company operating several production plants located in Poland and Ukraine. It was carried out in three main stages, in line with the CBM (condition-based maintenance) strategy. The basis for the research were the work and inspection cards of roll making machines provided by the company, which came from one of the plants and covered the period from 1 September 2015 to 31 August 2017. These provided information in two main categories. Event data indicated what events occurred during the operation of the machine (i.e., the need to repair, replacement of worn parts or breakdowns). On the other hand, the condition monitoring data provided information about the current technical condition of the facilities and the need for preventive measures (e.g., adjustment of Teflon blades). The processing of the above information and interpretation thereof made it possible to identify factors shaping the efficiency of the machinery stock. Then, on their basis, a model for the evaluation of machinery efficiency was built. It was assumed that its satisfactory level was 90%. This value is based on the daily production cycle, which also includes the breaks required by the Labor Code, daily service and the preparation of the machine for operation. Finally, the manner in which the model can support decision making in the area of improvement of the production processes was indicated [35].

Due to the specificity of manufactured products, i.e., serial products with standard parameters, the company operates in the MTS (make to stock) production system. The plant works on a three-shift basis, with each shift lasting 8 h. The process of model parameter estimation and the results obtained are presented in subsequent sections of this article.

## 2. Logistic Regression Model

Logistic regression is a model that allows the influence of several variables  $X_1, X_2, \dots, X_k$  on the dichotomous variable  $Y$  in the mathematical form to be presented. The logistic regression model is based on a logistical function that takes the following form:

$$f(x) = \frac{e^x}{1 + e^x} = \frac{1}{1 + e^{-x}}, \tag{2}$$

where  $e$  is the Euler number, and  $x$  is the value of the explanatory variable  $X$ .

The use of logistic regression is supported by the fact that it is not required to meet many assumptions that are formulated in relation to linear regression and general linear models. These include, first of all, the linearity of the relationship between a dependent and an independent variable, as well as the normality and homoscedasticity of the distribution of independent variables. In addition, observations must be reported using metric measurement systems.

The logistic regression model can be written in several ways. Assuming that  $Y$  stands for a dichotomous variable with values 1, for the occurrence of the event we are interested in (success), and 0, for the opposite case (failure), the logistic regression model is described by Equation (3):

$$P(Y = 1|x_1, x_2, \dots, x_k) = \frac{e^{\beta_0 + \sum_{i=1}^k \beta_i \cdot x_i}}{1 + e^{\beta_0 + \sum_{i=1}^k \beta_i \cdot x_i}}, \tag{3}$$

where  $\beta_i, i = 0, \dots, k$  are logistic regression factors, while  $x_1, x_2, \dots, x_k$  are independent variables, which can be measurable or qualitative.

An equivalent form of the logistic regression equation can be written as the odds for the occurrence of the event (success) we are interested in:

$$\frac{P(Y = 1|X)}{1 - P(Y = 1|X)} = e^{\beta_0 + \sum_{i=1}^k \beta_i \cdot x_i}. \tag{4}$$

In a special case, for one independent variable the logistic regression equation takes the following form:

$$P(Y = 1|X) = \frac{e^{\beta_0 + \beta_1 \cdot x_1}}{1 + e^{\beta_0 + \beta_1 \cdot x_1}}. \tag{5}$$

If, in turn, both sides of the Equation (5) are logarithmized, the logit form of the logistic model will be obtained:

$$\text{logit } P(Y = 1|X) = \ln \frac{P(Y = 1|X)}{1 - P(Y = 1|X)} = \beta_0 + \beta_1 \cdot x_1. \tag{6}$$

The condition necessary for logistic regression is a sufficiently large sample, the number of which should be  $n > 10(k + 1)$ , where  $k$  is the number of parameters.

Important concepts related to logistic regression are the odds and the odds ratio. The odds are defined as the probability of an event occurring  $P(A)$  divided by the probability of an event not occurring,  $1 - P(A)$ :

$$(\text{Odds})S(A) = \frac{P(A)}{P(\text{non} - A)} = \frac{P(A)}{1 - P(A)}. \tag{7}$$

The odds ratio, in turn, marked  $OR$ , is defined as the odds of one event occurring  $S(A)$  divided by the odds of another event occurring  $S(B)$ :

$$OR_{AxB} = \frac{S(A)}{S(B)} = \frac{P(A)}{1 - P(A)} : \frac{P(B)}{1 - P(B)}. \tag{8}$$

### 3. Estimation of Markov Logistic Model Parameters

The first stage of the study was to define possible explanatory variables in order to determine which of them could be used in the model. The following explanatory variables were selected: shift, device, occurrence of failure (yes or no) and no production order (yes or no).

The shift predictor was analyzed first. First of all, the normality of distribution and homogeneity of the variance of the efficiency dependent variable during individual shifts was examined in order to determine the possible methods of statistical analysis. The distributions in all groups turned out to be inconsistent with the normal distribution, which is confirmed by the graphs in Figure 1 and the calculated chi-square test statistic values, presented in Table 1.

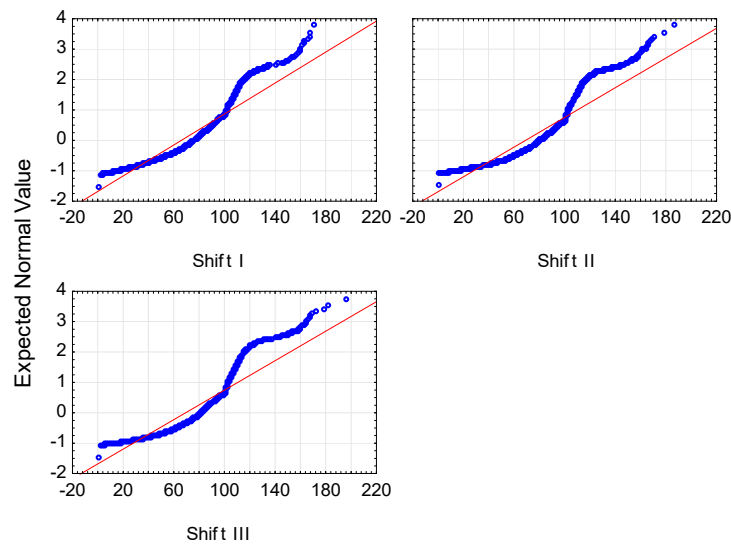


Figure 1. Graphs of normality of distribution of the efficiency variable grouped by shifts.

Table 1. Chi-square test statistic values of the efficiency variable grouped by shifts.

	Chi-Square	Degrees of Freedom	p-Value
Shift = 1st	4355.45	df = 8	$p = 0.00$
Shift = 2nd	5932.71	df = 8	$p = 0.00$
Shift = 3rd	6474.96	df = 8	$p = 0.00$

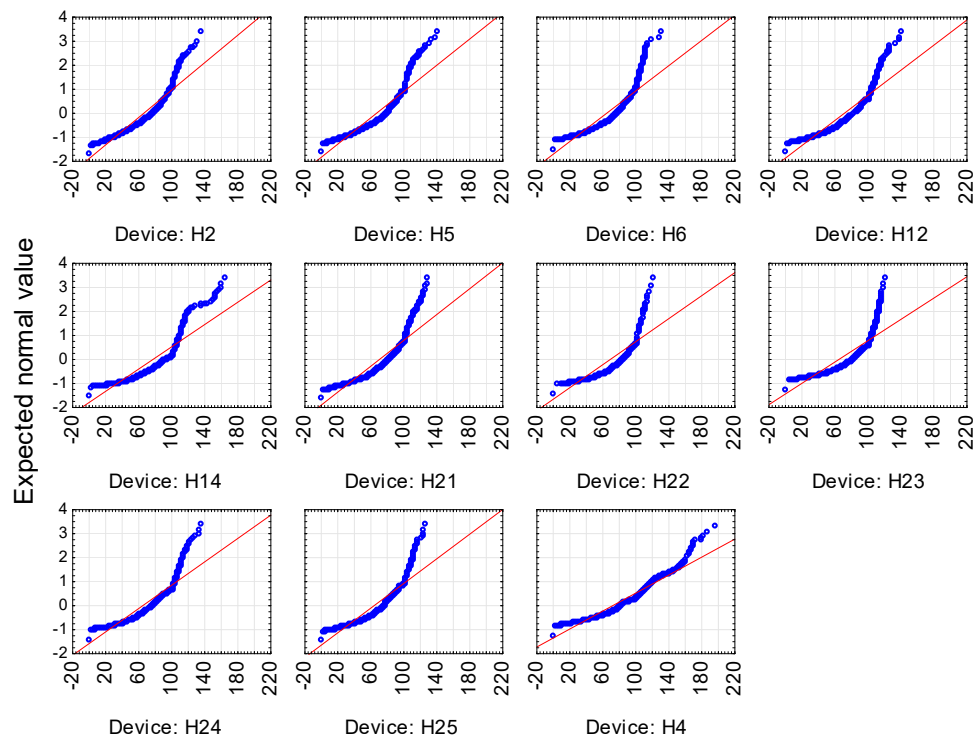
Next, the homogeneity of variance in individual groups was checked; the Levene and Brown-Forsythe test was used for this purpose. The obtained results are presented in Table 2.

Table 2. Results of the Levene and Brown-Forsythe tests of the efficiency variable grouped by shifts.

Average—1st	Average—2nd	Average—3rd	Levene F(1,df)	Levene p	Brn-Fors F(1,df)	Brn-Fors p
66.54	69.84		1.54	0.21	0.01	0.9
66.54		70.06	1.07	0.3	0.35	0.55
	69.84	70.06	0.039	0.84	0.47	0.49

Although the homogeneity of variance was confirmed in all groups, due to the lack of normality of distributions, the Mann–Whitney test was used to examine the significance of differences between individual averages, and the results thereof are presented in Table 3.





**Figure 3.** Graphs of normality of distribution of the efficiency variable grouped by the device variable.

Since the average efficiency varied for virtually every pair of devices, a decision was made not to combine them and to include each of them in the study. After defining the form of independent variables, the impact of each of them on the dependent variable, i.e., efficiency, was checked, but presented in a dichotomous form, as an assessment of whether the level achieved was satisfactory for the company. In line with the expectations of the Management Board, it was assumed that the assessment was positive if the productivity was equal to or above 90%. In other cases, the assessment would be negative. The chi-square test allowed for a statistical and substantive study of the relationship between variables. In all cases, the calculated test statistic did not allow the zero hypothesis on the lack of relationships between variables to be accepted. It was therefore rejected in favor of the alternative hypothesis of the existence of a relationship, the strength of which was measured using Yule’s  $\Phi$  (for binary tables) and Cramér’s  $V$  coefficient (for tables more complex than  $2 \times 2$ ). The obtained results are presented in Table 5.

The observed relationships between variables, although significant, are not strong. This is also confirmed by the graphs of interaction of individual dependent variables with the explained variable (Figure 4). Nevertheless, from the point of view of the analyzed company, the diagnosed bonds should not take place at all. A uniform and efficient operation of all devices is expected, so even minor deviations are undesirable and require further investigation.

The calculations carried out (Table 5) and the charts (Figure 4) confirm that the model variables were selected correctly. This allows the parameters of the logistic regression model to be estimated, the values of which are presented in Table 6.

**Table 5.** Results of the tests of significance and strength of the relationship between the predictors and the efficiency variable.

Statistics	Chi-Square	df	p
<b>Shift</b>			
Pearson’s Chi <sup>2</sup>	82.35	df = 1	p = 0.00
NW Chi <sup>2</sup>	83.12	df = 1	p = 0.00
Yates Chi <sup>2</sup>	82.10	df = 1	p = 0.00
Fi for 2 × 2 tables	0.06		
Contingency coefficient	0.06		
<b>Device</b>			
Pearson’s Chi <sup>2</sup>	538.57	df = 10	p = 0.00
NW Chi <sup>2</sup>	532.84	df = 10	p = 0.00
Contingency coefficient	0.15		
Cramér’s V	0.15		
<b>Failure</b>			
Pearson’s Chi <sup>2</sup>	786.17	df = 1	p = 0.00
NW Chi <sup>2</sup>	1062.94	df = 1	p = 0.00
Yates Chi <sup>2</sup>	784.68	df = 1	p = 0.00
Fi for 2 × 2 tables	−0.18		
Contingency coefficient	0.18		
<b>Order</b>			
Pearson’s Chi <sup>2</sup>	1756.69	df = 1	p = 0.0000
NW Chi <sup>2</sup>	2576.74	df = 1	p = 0.0000
Yates Chi <sup>2</sup>	1754.97	df = 1	p = 0.0000
Fi for 2 × 2 tables	−0.27		
Contingency coefficient	0.26		

**Table 6.** Parameters of the logistic regression model and their evaluation.

Effect	Modeled Probability Evaluation of Effectiveness—Positive Result Distribution: Binomial, Binding Function: LOGIT						
	Effect Level	Parameter	Estimated Standard Error	Wald’s Test Statistics	p	95.00% CI	−95.00% CI
Absolute term		$\beta_0 = -7.549$	0.288	689.03	0.00	−8.112	−6.985
Shift	1st	$\beta_1 = -0.292$	0.031	90.59	0.00	−0.352	−0.232
Device	H2	$\beta_2 = -1.241$	0.067	343.29	0.00	−1.373	−1.110
Device	H5	$\beta_3 = -1.066$	0.067	256.69	0.00	−1.196	−0.936
Device	H6	$\beta_4 = -1.153$	0.068	291.56	0.00	−1.286	−1.021
Device	H12	$\beta_5 = -0.668$	0.065	104.14	0.00	−0.796	−0.539
Device	H21	$\beta_6 = -0.809$	0.065	153.40	0.00	−0.937	−0.681
Device	H22	$\beta_7 = -0.496$	0.067	55.25	0.00	−0.627	−0.365
Device	H23	$\beta_8 = -0.525$	0.067	60.84	0.00	−0.656	−0.393
Device	H24	$\beta_9 = -0.873$	0.067	168.66	0.00	−1.004	−0.741
Device	H25	$\beta_{10} = -1.167$	0.068	296.30	0.00	−1.300	−1.034
Device	H4	$\beta_{11} = -0.317$	0.072	19.323	0.00	−0.458	−0.176
Order	no	$\beta_{12} = 5.047$	0.251	403.059	0.00	4.554	5.539
Failure	no	$\beta_{13} = 3.08$	0.135	520.933	0.00	2.816	3.345



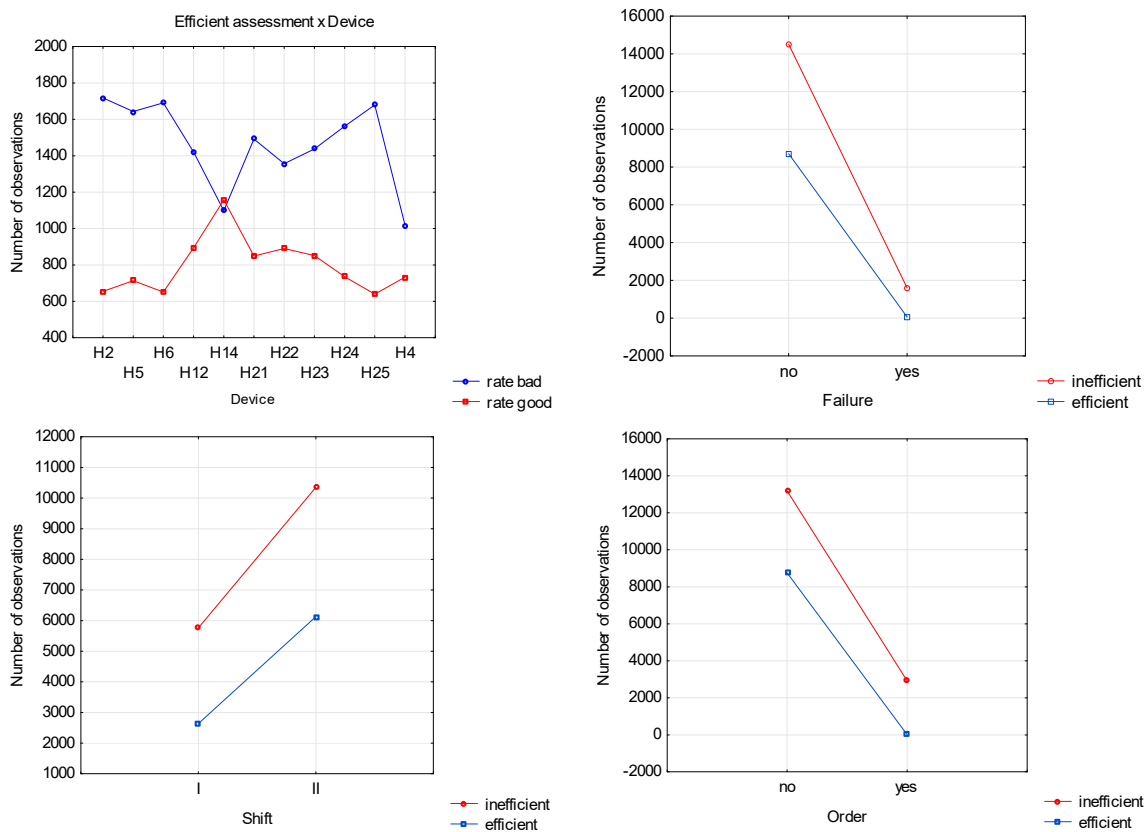


Figure 4. Interaction charts of dependent variable and predictors.

All calculated parameters turned out to be statistically significant, which is confirmed by the calculated Wald’s statistic value and the associated probability value  $p$ , which for each line is lower than the assumed level of significance  $\alpha = 0.05$ . (Table 6). This means that all the distinguished factors significantly affect the evaluation of production efficiency. This allows the equation of the logistic regression model to be written in the following form:

$$P(\text{efficiency} = \text{yes}|X) = \frac{e^a}{1 + e^a}, \tag{9}$$

where

$$a = -7.549 - 0.292 \text{ shift I} - 1.241 * H2 - 1.066 * H5 - 1.153 * H6 - 0.668 * H12 - 0.809 * H21 - 0.496 * H22 - 0.525 * H23 - 0.873 * H24 - 1.167 * H25 - 0.317 * H4 + 5.047 * \text{no order} + 3.08 * \text{no failure}. \tag{10}$$

The logistic regression curve is shown in Figure 5.

The logistic regression equation presented above can also take equivalent forms:

- logistic regression logit function:

$$\text{logit } P(\text{efficient} = 1|X) = \ln \frac{P(\text{efficient}=1|X)}{1-P(\text{efficient}=1|X)} = -7.549 - 0.292 \text{ 1st shift} - 1.241 * H2 - 1.006 * H5 - 1.153 * H6 - 0.668 * H12 - 0.809 * H21 - 0.496 * H22 - 0.525 * H23 - 0.873 * H24 - 1.167 * H25 - 0.317 * H4 + 5.047 * \text{no order} + 3.08 * \text{no failure}, \tag{11}$$

- in the form of the odds:

$$P \frac{P(\text{efficient} = 1|X)}{1 - P(\text{efficient} = 1|X)} = e^a, \tag{12}$$

where

$$a = -7.549 - 0.292 \text{ shift } I - 1.241 * H2 - 1.066 * H5 - 1.153 * H6 - 0.668 * H12 - 0.809 * H21 - 0.496 * H22 - 0.525 * H23 - 0.873 * H24 - 1.167 * H25 - 0.317 * H4 + 5.047 * \text{no order} + 3.08 * \text{no failure}. \tag{13}$$

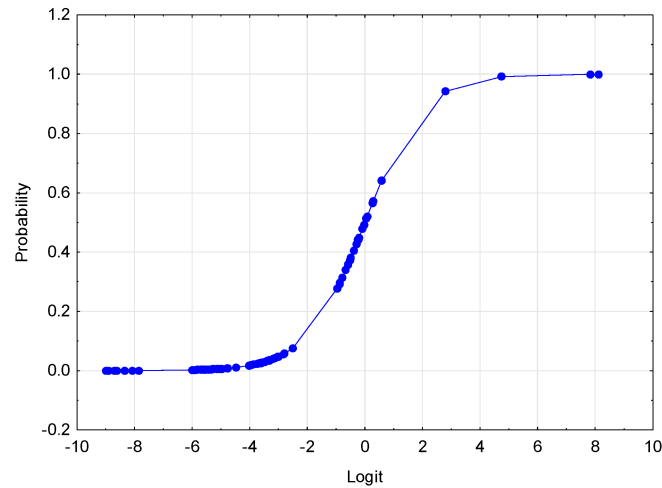


Figure 5. Logistic regression curve.

An important element of the evaluation of the studied process is the calculation of the odds of an event occurrence (in this case a satisfactory result of efficiency). The sign at the estimated parameter of the logistic regression model indicates whether the analyzed odds are greater (plus) or smaller (minus) in relation to the reference level. The scale of this change is indicated by the unit odds ratio shown for each parameter in Table 7.

Table 7. Odds ratios for individual predictors.

Effect	Effect Level	Odds Ratio	95.00% CI	−95.00% CI	<i>p</i>
Shift	1st	0.747	0.704	0.793	0.00
Device	H2	0.289	0.253	0.330	0.00
Device	H5	0.344	0.302	0.392	0.00
Device	H6	0.316	0.276	0.360	0.00
Device	H12	0.513	0.451	0.583	0.00
Device	H21	0.445	0.392	0.506	0.00
Device	H22	0.609	0.534	0.694	0.00
Device	H23	0.592	0.519	0.675	0.00
Device	H24	0.418	0.366	0.477	0.00
Device	H25	0.311	0.273	0.356	0.00
Device	H4	0.729	0.633	0.839	0.00
No order—yes no	no	155.489	95.002	254.487	0.00
Failure—yes no	no	21.7679	16.7083	28.3595	0.00

The odds ratio for the 1st shift is 0.75, which means that compared to the 2nd shift, the odds of achieving satisfactory efficiency is 0.75 times lower. In other words, the odds for the proper efficiency is 1.34 times higher in the case of the 2nd shift. The H14 machine was used as a reference when analyzing the impact on the efficiency of individual devices. Its efficiency is the highest in the test sample, so all the model coefficients obtained are negative, which means less odds of achieving a positive result. The odds ratios are given in column 3 of Table 6. The worst result was obtained for the H2 machine with an odds ratio of 0.289, which means an almost 3.5-fold increase in the odds of achieving satisfactory efficiency when replacing H2 with H14.

The results for the other machines, showing how much the efficiency of each machine should be increased in order to obtain the efficiency evaluation as in the case of the H12 machine, are shown in Table 8.

**Table 8.** Odds ratios for individual predictors.

Machine Number	H2	H5	H6	H12	H21
OR	3.460	2.904	3.169	1.950	2.246
Machine Number	H22	H23	H24	H25	H4
OR	1.642	1.690	2.393	3.212	1.372

The last two parameters in Table 7 refer to the absence of an order or failure, as their occurrence has a negative impact on the efficiency. Where there is no downtime, the odds of achieving the expected efficiency are 155 times greater than otherwise. Similarly, the occurrence of a failure has a similar effect, but its absence is not as spectacular. There is a 21-fold increase in the odds if the failure does not occur.

The presented model can also be used for predictive purposes, allowing for forecasting the probability of achieving the predicted success (here, the assumed efficiency). It is therefore important to assess the quality of the prediction. For this purpose, it is helpful to determine the so-called cut-off point  $\pi_0$ . This parameter allows the observed dichotomous values of a dependent variable to be compared with the continuous probability values calculated on the basis of the model. This value falls within the range (0, 1) and is defined as follows [36] when:

$$\hat{\pi}(x) = \hat{P}(Y = 1|x) > \pi_0, \tag{14}$$

it is assumed that an event has occurred ( $\hat{y} = 1$ ). In the opposite situation, when

$$\hat{\pi}(x) \leq \pi_0, \tag{15}$$

it is assumed that an event has not occurred ( $\hat{y} = 0$ ).

Prediction ideally occurs when sensitivity and specificity are equal to 1, which means no false positive or negative results. In real life research, the point corresponding to a case where a model best discriminates occurrences is called the optimal cut-off point. It is determined using the Youden’s index ( $J$ ), which takes the following form:

$$J = sensitivity + specificity - 1. \tag{16}$$

The optimum cut-off point corresponds to the case where the  $J$  value reaches its maximum. For the case under consideration, the proposed cut-off point is shown in Table 9.

**Table 9.** Cut-off point of the logistic regression model concerned.

Number of Observations	Cut-off Point	True Positive	True Negative	False Positive	False Negative	SE	1-SP	Youden’s Index
5	0.51	2864	13,903	2210	5905	0.33	0.14	0.19

For the proposed cut-off point, the sensitivity is 0.32, and the specificity is 0.86. There are 16,767 well classified cases (2864 true positive and 13,903 true negative) and 8115 badly classified cases (2210 false positive and 5905 false negative cases).

Based on the above table it is possible to assess the effectiveness of model prediction in relation to successes and failures, using tools among which one can distinguish such statistics as accuracy, sensitivity or specificity, ROC (receiver operating characteristic) curve or values of rank correlations.

The simplest measure is accuracy, calculated according to the following formula:

$$Accuracy = ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{17}$$

where:

- TP—number of true positive results,
- TN—number of true negative results,
- FP—number of false positive results,
- FN—number of false negative results.

For the model concerned,

$$ACC = \frac{2864 + 13903}{2864 + 5905 + 2210 + 13903} = 0.674 = 67.4\%. \tag{18}$$

However, the sensitivity *SE* and specificity *SP* are most often considered in such analyses but treated as pairs, which, after being marked on the plane and after connecting the points with segments, form the so-called ROC curve. For the analyzed model this curve is presented in Figure 6.

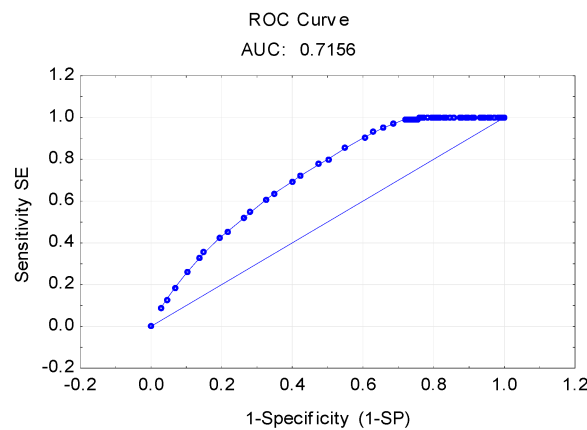


Figure 6. ROC curve for the model concerned.

The most important parameter for assessing the ROC curve is AUC—area under the ROC curve. It takes values from 0 to 1. The interpretation of the result was based on the Kleinbaum and Klein classification (Table 10), according to which discrimination is sufficient [37].

Table 10. Cut-off point of the logistic regression model concerned.

AUC Value	Score
0.9 < AUC < 1.0	Excellent discrimination
0.8 < AUC < 0.9	Good discrimination
0.7 < AUC < 0.8	Sufficient discrimination
0.6 < AUC < 0.7	Weak discrimination
0.5 < AUC < 0.6	Insufficient discrimination

The model can therefore be considered satisfactory, although it is recommended rather for qualitative analysis of processes and modification of production strategy on its basis.

Production in the company in question is carried out in a continuous three-shift system. Regardless of the production plan, the plant is fully manned and all machines are in operation at all times. The acceptable level of efficiency assumed by the management should be 90%, which allows all scheduled and expected downtime to be taken into account. The study indicates that in most cases

this level is not reached. In almost 25,000 out of over 156,000 observations, this level was not reached. It turned out that the efficiency of the first shift was lower compared to the second and third ones, which suggests the need to diagnose the causes of circumstances or even to restructure the shift system. The use of individual machines also has a negative impact on efficiency. It turns out that many of them represent much lower efficiency than the one taken as a reference point, the efficiency of which was the highest (but also less than 100%). Of course, a lack of orders also leads to a significant reduction in productivity. Such a result, next to failures that occur, indicates the need for a detailed analysis of the machinery. It might be advisable to exclude several machines from the production process so that production is closely correlated with demand. Unused machines could constitute a reserve in case of failure and thus increase the level of readiness of the machinery.

#### 4. Conclusions

The reliability of machinery and equipment is an essential part of the proper functioning of a business. Modern technologies support the maintenance of an adequate level of readiness and suitability of the technical infrastructure. They not only facilitate production control and implementation of modern operating strategies, but also ensure continuous monitoring of processes and detection of any disturbances or failures. The activities carried out in this area boil down to balancing the maintenance of full operational efficiency and continuity of production and ensuring an acceptable level of costs of these activities.

This is a difficult task, particularly in the case of older generation machinery stock, which is deprived of support from computerized production management systems—as the one presented in this article. In any case, however, the aim is to ensure that machines function perfectly without failures and that products are manufactured without defects. On the other hand, it is also important to ensure the efficiency of the equipment use and balanced workload, which proved to be a problem in the analyzed company. This was the reason for undertaking research in this field and the basis for mathematical analysis of the effectiveness of the manufacturing process.

The lack of IT systems for controlling and monitoring the production process results in the inability to archive data on an ongoing basis, which makes it much more difficult to control processes, and sometimes it is conducive to abandonment thereof. Poor quality of recorded documentation imposes significant limitations on the use of mathematical tools as well; therefore, the authors wanted to present simultaneously that even in such a situation it is possible to create mathematical models that would improve the efficiency of the machinery stock. The available data proved to be sufficient to achieve the research objective, which was to formulate a model providing an unambiguous answer as to whether the efficiency of the equipment used in production is acceptable from the point of view of the assumptions made by the company (in this case 90%).

This was made possible through the use of logistic regression, which, above all, does not require the meeting of assumptions made by other mathematical models, e.g., linear regression and general linear models. The advantage of this method is also the form of the dependent variable. The predictor here is a dichotomous variable and its values can be interpreted as the probability of an event occurring. The organization of production in the analyzed company has remained unchanged for many years. All machines work three shifts every day. Regardless of the orders placed and the market demand, full staff is employed. Machines are taken out of the process only in cases of random incidents. Lack of modification of the adopted procedures and control of the implemented processes results—as demonstrated in the research—in the process not being effective and the use of machinery not being optimal.

The logistic regression model made it possible to identify the causes influencing machinery efficiency. It turned out that the load is not identical during every shift, the productivity is much lower during the first shift in comparison to the other shifts. Restructuring the shift system and limiting the production process to only two shifts or modifying the working time could increase the productivity of machinery, optimize the use of human resources and reduce the costs of the production process. The load on the individual machines also appeared to be disproportionate. Increased use of one piece

of equipment may result in an increase in the frequency of breakdowns, increase the costs of repairs and reduce the total life of the equipment, so it would be advisable to evenly distribute production across all equipment. The reduction in productivity was also caused by a decline in the number of production orders. The lack of production control with regard to orders received and the maintenance of a continuous three-shift readiness exposes the company to costs and favors the aforementioned disproportionate workload.

The studied company has no IT systems in place that enable comprehensive monitoring of production processes. Therefore, the unquestionable advantage of the proposed model is the provision of additional information allowing decisions to be made on the production and use of machinery. They can also encourage the implementation of modern solutions and the abandonment of traditional, outdated methods of recording and archiving data.

In companies that use specialist MES systems based on real-time information on manufacturing execution at subsequent workstations, the proposed model can improve monitoring of productivity drops below the adopted level and activate preventive actions. Additionally, the implementation of data obtained from the IT system may allow to the model to be extended with additional parameters, which is important from the point of view of individual companies.

The aim of the article was to investigate the possibility of developing a model for the analysis and evaluation of the level of efficiency of ongoing production processes, as well as to indicate the method of logistic regression as a tool supporting decision-making in this respect. The model developed for the analyzed company indicated the need for a strict correlation between the demand for a product and the production process. Adopted strategies require verification and modification.

The proposed model may also serve as a basis for setting directions for improvement of the production process, by maximizing the use of the machinery stock and reducing the idle time of both employees and equipment. Re-application of the logistic regression model constructed on the basis of the observation of the process after introduction of changes allows the effectiveness and efficiency of implemented solutions to be evaluated.

**Author Contributions:** Conceptualization, A.B. and M.G.; Formal analysis, A.B.; Methodology, A.B.; Resources, M.G.; Writing—original draft, A.B. and M.G.; Writing—review & editing, A.B.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Świć, A.; Gola, A. Economic Analysis of Casing Parts Production in a Flexible Manufacturing System. *Actual Probl. Econ.* **2013**, *141*, 526–533.
2. Gola, A. Reliability analysis of reconfigurable manufacturing system structures using computer simulation methods. *Ekspluat. I Niezawodn. Maint. Reliab.* **2019**, *21*, 90–102. [[CrossRef](#)]
3. Jasiulewicz-Kaczmarek, M.; Saniuk, A. How to make maintenance processes more efficient using lean tools? *Adv. Intell. Syst. Comput.* **2017**, *605*, 9–20.
4. Antosz, K.; Stadnicka, D. Evaluation measures of machine operation effectiveness in large enterprises: Study results. *Ekspluat. I Niezawodn. Maint. Reliab.* **2015**, *17*, 107–117. [[CrossRef](#)]
5. Chien, C.-F.; Dou, R.; Fu, W. Strategic capacity planning for smart production: Decision modeling under demand uncertainty. *Appl. Soft Comput.* **2018**, 68900–68909. [[CrossRef](#)]
6. Iscioglu, F.; Kocak, A. Dynamic reliability analysis of a multi-state manufacturing system. *Ekspluat. I Niezawodn. Maint. Reliab.* **2019**, *21*, 451–459. [[CrossRef](#)]
7. Liu, T.; Chen, C.; Kuo, Y.; Kuo, M.; Hsueh, P. Developing a monitoring system of production effectiveness and traceability for semi-automatic production process. In Proceedings of the International Conference on Fuzzy Theory and Its Applications (iFUZZY), Pingtung, Taiwan, 12–15 November 2017; pp. 1–5.
8. Loska, A. Przegląd modeli ocen eksploatacyjnych systemów technicznych. *Komput. Zinteg. Zarz.* **2011**, *2*, 37–46.

9. Kang, N.; Zhao, C.; Li, J.; Horst, J.A. A Hierarchical structure of key performance indicators for operation management and continuous improvement in production systems. *Int. J. Prod. Res.* **2016**, *54*, 6333–6350. [[CrossRef](#)]
10. Koliński, A.; Śliwczyński, B.; Golińska-Dawson, P. The assessment of the economic efficiency of production process-simulation approach. In Proceedings of the 24th International Conference on Production Research (ICPR 2017), Poznań, Poland, 30 July–3 August 2017; pp. 283–289.
11. Chen, W.; Wang, Z.; Chan, F.T.S. Robust production capacity planning under uncertain wafer lots transfer probabilities for semiconductor automated material handling systems. *Eur. J. Oper. Res.* **2017**, *261*, 929–940. [[CrossRef](#)]
12. Gola, A.; Kosicka, E.; Daniewski, K.; Mazurkiewicz, D. Analiza błędów przy ocenie wskaźnika OEE na przykładzie linii rozlewu butelkowego. *Innow. W Zarz. I Inż. Prod.* **2016**, *2*, 654–662.
13. Aleš, Z.; Pavlu, J.; Legát, V.; Mošna, F.; Jurča, V. Methodology of overall equipment effectiveness calculation in the context of Industry 4.0 environment. *Ekspluat. I Niezawodn. Maint. Reliab.* **2019**, *21*, 411–418. [[CrossRef](#)]
14. Wilczarska, J. Efektywność i bezpieczeństwo użytkowania maszyn. *Inż. I Apar. Chem. Chem. Eng. Equip.* **2012**, *51*, 41–43.
15. Szwedzka, K.; Jasiulewicz-Kaczmarek, M. Determining maintenance services using production performance indicators. *Res. Logist. Prod.* **2016**, *6*, 361–374. [[CrossRef](#)]
16. Jasiulewicz-Kaczmarek, M.; Piechowski, M. Practical Aspects of OEE in Automotive Company—Case Study. In Proceedings of the 3rd International Conference on Management Science and Management Innovation (MSMI 2016), Guilin, China, 13–14 August 2016; pp. 213–218.
17. Oechsner, R.; Pfeffer, M.; Pfitzner, L.; Binder, H.; Muller, E.; Vonderstrass, T. From overall equipment efficiency (OEE) to overall Fab effectiveness (OFE). *Mat. Sci. Semicond. Process.* **2002**, *5*, 333–339. [[CrossRef](#)]
18. Wang, T.J.; Pan, H.C. Improving the OEE and UPH data quality by Automated Data Collection for the semiconductor assembly industry. *Expert Syst. Appl.* **2011**, *38*, 5764–5773. [[CrossRef](#)]
19. Ylipää, T.; Skoogh, A.; Bokrantz, J.; Gopalakrishnan, M. Identification of maintenance improvement potential using OEE assessment. *Int. J. Prod. Perform. Manag.* **2017**, *66*, 126–143. [[CrossRef](#)]
20. Jonsson, P.; Lesshammer, M. Evaluation and improvement of manufacturing performance measurement systems—the role of OEE. *Int. J. Oper. Prod. Manag.* **1999**, *19*, 55–78. [[CrossRef](#)]
21. Muchiri, P.; Pintelon, L. Performance measurement using overall equipment effectiveness (OEE): Literature review and practical application discussion. *Int. J. Prod. Res.* **2008**, *46*, 3517–3535. [[CrossRef](#)]
22. Antosz, K.; Pacana, A. Wskaźnikowa ocena efektywności funkcjonowania maszyn na przykładzie wybranego przedsiębiorstwa—studium przypadku. *Innow. W Zarz. I Inż. Prod.* **2016**, *2*, 634–6461.
23. Panagiotis, H.T. Evaluation of overall equipment effectiveness in the beverage industry: A case study. *Int. J. Prod. Res.* **2013**, *51*, 515–523.
24. Gola, A. Economic Aspects of Manufacturing Systems Design. *Actual Probl. Econ.* **2014**, *6*, 205–2012.
25. Kubik, S.; Kornicki, K. *OEE dla operatorów. Całkowita Efektywność Wyposażenia*; ProdPress.com: Wrocław, Poland, 2009.
26. Panagiotis, H.T. Evaluation of maintenance management through the overall equipment effectiveness of a yogurt production line in a medium-sized Italian company. *Int. J. Prod. Qual. Manag.* **2015**, *16*, 123–135.
27. Luo, H.; Wang, K.; Kong, X.T.R.; Lu, S.; Qu, T. Synchronized production and logistics via ubiquitous computing technology. *Robot. Comput. Integr. Manuf.* **2017**, *45*, 99–115. [[CrossRef](#)]
28. Muchiri, P.; Pinteleon, L.; Martin, H.; De Meyer, A.M. Empirical analysis of maintenance performance measurement in Belgian industries. *Int. J. Prod. Res.* **2010**, *20*, 5905–5924. [[CrossRef](#)]
29. Kalliski, M.; Krahé, D.; Beisheim, B.; Krämer, S.; Engell, S. Resource Efficiency Indicators for Real-Time Monitoring and Optimization of Integrated Chemical Production Plants. In *Computer Aided Chemical Engineering*; Gernaey, K.V., Huusom, J.K., Gani, R., Eds.; Elsevier: Amsterdam, The Netherlands, 2015; Volume 37, pp. 1949–1954.
30. Yan, J.; Lee, J. Degradation assessment and fault modes classification using logistics regression. *J. Manuf. Sci. Eng.* **2005**, *127*, 912–914. [[CrossRef](#)]
31. Kozłowski, E.; Mazurkiewicz, D.; Żabiński, T.; Pruncal, S.; Sęp, J. Assessment model of cutting tool condition for real-time supervision system. *Ekspluat. I Niezawodn. Maint. Reliab.* **2019**, *21*, 679–685. [[CrossRef](#)]
32. Yan, J.; Ko, M.; Lee, J. A prognostic algorithm for machine performance assessment and its application. *Prod. Plan. Control* **2004**, *15*, 796–801. [[CrossRef](#)]

33. Caesarendra, W.; Widodo, A.; Yang, B.S. Application of relevance vector machine and logistic regression for machine degradation assessment. *Mech. Syst. Signal Process.* **2010**, *24*, 1161–1171. [[CrossRef](#)]
34. Chen, B.; Chen, X.; Li, B. Reliability estimation for cutting tool based on logistics regression model using vibration signals. *Mech. Syst. Signal Process.* **2011**, *25*, 2516–2537. [[CrossRef](#)]
35. Jardine, A.K.S.; Lin, D.; Banjevic, D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech. Syst. Signal Process.* **2006**, *20*, 1483–1510. [[CrossRef](#)]
36. Stanisław, A. *Modele Regresji Logistycznej. Zastosowanie W Medycynie, Naukach Przyrodniczych I Społecznych*; Statsoft Polska: Kraków, Poland, 2016.
37. Kleinbaum, D.G.; Klein, M. *Logistic Regression A Self-Learning Text*; Springer: Berlin, Germany, 2010.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).