Using the Kaplan–Meier Estimator to Assess the Reliability of Agricultural Machinery

Karol Durczak 1,*, Jarosław Selech 2, Adam Ekielski 3, Tomasz Żelaziński 3, Marcin Waleński 1 and Kamil Witaszek 1

1 Department of Biosystems Engineering, Faculty of Environmental and Mechanical Engineering, Poznań University of Life Sciences, Wojska Polskiego 50 Street, 60-627 Poznan, Poland; marcin.walenski@up.poznan.pl (M.W.); kamil.witaszek@up.poznan.pl (K.W.)
2 Faculty of Civil and Transport Engineering, Poznan University of Technology, 3 Piotrowo st., 60-965 Poznan, Poland; jaroslaw.selech@put.poznan.pl
3 Department of Production Management and Engineering, Warsaw University of Life Sciences, ul. Nowoursynowska 164, 02-787 Warsaw, Poland; adam_ekielski@sggw.pl (A.E.); tomasz_zelazinski@sggw.pl (T.Z.)

* Correspondence: karol.durczak@up.poznan.pl

Abstract: Kaplan–Meier analyses can be used in many disciplines, e.g., agricultural engineering. Agricultural machinery and vehicles can be regarded as objects that ‘die’ because, like living creatures, they failed, although after repair they can be used until scrapped. This article presents an example of using the Kaplan–Meier estimator to plot the reliability function curves of five different models of Zetor farm tractors. The research shows that the median operating time for one of the tested models, which is about 200 engine-operating hours, is 20% lower than for the entire population of analyzed Zetor tractors. This means that the quality of the model, which is very popular in Poland, differs significantly from the other models of this manufacturer. The method cannot be validated, due to a lack of similar functions for other brands of tractors. Progressive automation and digitization of agriculture can contribute to improving the reliability of agriculture work. The user can focus on the correct performance of agrotechnical treatments, and modern control systems will signal in real time, about identified or approaching costly failures.

Keywords: agricultural engineering; tractors; censored data; smart decision-making; agriculture 4.0; automation

1. Introduction

One of the main criteria in the global assessment of technical facilities, including agricultural machinery and vehicles, is their reliability, which determines the probability of the occurrence of a failure, over time. The reliability theory has been well known and well documented for several decades [1–7]. It developed, quickly, after World War II, due to the huge (gargantuan) amount of data, on damage to all types of combat equipment (tanks, planes, guns, etc.). As a consequence of the quantification of its reliability, which turned out to be very low compared to the expectations of soldiers, there was rapid development of technology and elimination of design errors. Today, the quality of combat equipment is excellent because its reliability has improved. On the contrary, there has never been such a revolution or evolution in agricultural equipment. However, due to intense competition in the agricultural-machinery market (e.g., there are more than 50 farm-tractor manufacturers), it is necessary to improve quality and reliability. Each self-respecting manufacturer of agricultural machinery has its own design office and tries to eliminate defects at this stage of machine life, e.g., by using better materials or equipping products with new safety systems.

However, if we take cars as an example, their quality and reliability are monitored. Automotive organizations and associations, such as ADAC, DEKRA, TÜV, GTÜ, and
Warranty Direct, as well as research centers [8] have been publishing car reliability or unreliability rankings for many years. However, such rankings are not available for agricultural machinery, which is an equally large group. For example, in Poland, there are about 1.5 million registered farm tractors, 152,000 combine harvesters, 28,000 beet harvesters, 80,000 potato harvesters, 12,000 forage harvesters, 496,000 field sprayers and 52,000 orchard sprayers.

It is difficult or impossible to access data on the failures of agricultural machinery because this information is restricted, as it is considered sensitive and proprietary. Lack of access to these data is a major obstacle, to creating relevant reliability ratings.

The staff of the Department of Biosystems Engineering at Poznan University of Life Sciences, noticed this problem and decided to help to make such rankings. In 2017, the IFOP platform (Independent Farmers’ Opinion Poll) was established on the Internet (Figure 1).

Figure 1. The start-up screen of the IFOP platform, at http://www.nbor.pl/index.php?lang=en (accessed on 1 September 2020).

The global quality index $Q$ is determined in a multi-stage procedure (Figure 2).

The collection of at least 30 questionnaires referring to a particular brand (or another criterion, e.g., age, power), i.e., a large random sample, is a prerequisite to implement stage 5. Additionally, various numbers of datasets, e.g., brands, were verified with the chi-squared test of independence [9].

It constantly monitors and collects users’s (mainly farmers’s) opinions on their satisfaction with agricultural machinery. One of the most important groups for the main criteria is the reliability criterion, $R$. Farmers can evaluate up to 54 detailed criteria, regarding a tractor, including 11 reliability criteria. Data are carefully collected, and algebraic-heuristic methods, based on weighted averages, are applied to process them. The results are, successively, published in Top Agrar Polska magazine (the factual and media partner of the IFOP project) and at important national agricultural events (Polagra Premiery and Agro Show). Therefore, we can be certain to say that this is an exception in the world. There is no other similar initiative. For this reason, it is difficult to argue whether the final results are fully subjective or if they concern only Poland.
Figure 2. An algorithm of the quantification of the farming machinery and vehicles quality indicator, Q.

Long-term research on the assessment of the suitability of agricultural equipment for work has resulted in a number of publications on the subject. This problem was extensively discussed in [10]. There have, also, been numerous other ideas for the quantification of quality (with simplified methods), using tools, methods, and concepts taken from quality engineering [11–16].

The reliability evaluation methods used in the studies were based on classical statistical relationships. Tractors were excluded from further use, regardless of the number of failures identified. Incomplete information on the sample under analysis was the basis for the evaluation. In this situation, the survival function may be helpful.

In 1958, Edward L. Kaplan and Paul Meier collaborated to publish a seminal paper, on how to deal with incomplete observations [17]. Subsequently, the Kaplan–Meier (K–M) curves and estimates of survival data became a familiar way of dealing with differing survival times (times-to-event), especially when not all the subjects continued the study. Examples of when times-to-events may be important include end-point variables.

The K–M estimate is one of the best options to be used, to measure the fraction of subjects living for a certain amount of time after treatment. In clinical trials or community trials, the effect of an intervention is assessed by measuring the number of surviving or saved subjects after that intervention, over a period of time. The time from a defined point to the occurrence of a given event, for example death, is called survival time, whereas the analysis of group data is called survival analysis [18].

The following situations are possible:
- some of the ‘participants’ observed in the study leave the trial before the end of the study,
- the study ends before the ‘death’ of some ‘participants’ (this situation, often, occurs in tests of the reliability of technical facilities, which had no time to fail, due to their excellent reliability and long periods of operation),
- during the study, new ‘participants’ meeting the participation criterion were added.

For example, ‘survival’ times do need to relate to actual survival, with death being the event; the ‘event’ can be any event of interest. K–M analyses are, also, used in non-medical disciplines [19].
The survival function \( P(T > t) \) can be estimated as \( S(x) = l_x/l_0 \), where \( l_x \) refers to the number of ‘participants’ ‘living’ up to age \( x \), if we have complete (complete) information on the sample, under analysis. Due to the cost and time-consuming nature of such studies, it is, often, impossible to have full observations, and incomplete, so-called censored data can, also, be used. In such situations, the K–M function enables survival analysis, by estimation of \( S(t) \).

The K–M estimator is so important in statistical analysis that it has been implemented in many pieces of commercial software, such as R, Reliasoft, JMP, or Statistica, which is used here. In addition, Statistica enables the choice of one of three ways, to determine the survival times of the tested sample:

- select one variable with failure times, e.g., the number of weeks, days, years, etc.,
- select two variables with the start and end dates,
- select six variables with dates: day (1–31), month (1–12), and year when the observation began (e.g., when the patient was admitted to the hospital) and day, month, and year when the observation ended (due to the patient’s death/failure or the end of the observation, for example, when the patient was discharged from the hospital).

When processing the data, the survival-analysis module calculates the number of days between the dates and performs an analysis based on this measure.

In order to calculate the confidence intervals for the K–M estimator, it is necessary to estimate its variance [20–22]. Another variance estimator was proposed by Peto. Many statistical tests were proposed to compare the survival function, e.g., the log-rank test (Mantel–Haenszel procedure), Breslow test (generalized Wilcoxon test), Tarone-Ware test, and the Peto test. A very important characteristic of the moment of ‘death’ is its median, i.e., the number defined as the moment at which half of the population ‘survives’.

This methodology perfectly fits into the real picture of research on the reliability of technical facilities, e.g., agricultural equipment. If we have a group of objects with similar functions (e.g., ploughs, seed drills, sprayers, combine harvesters, or tractors) and information about the number of objects leaving the group, new objects joining the group, and failed objects (temporarily excluded from further operation), it is possible to calculate the values of the K–M estimator, i.e., the reliability of these objects as a function of time \( R(t) \).

The K–M estimator of the survival function can be used to quantify technical facilities in each automated farm. The simulations were carried out on tractors, which are a basic source of tractive energy, for most hooked-up and suspended agricultural machinery.

2. Materials and Methods

Actual data on the failures of 74 farm tractors of the Czech brand Zetor Tractors a.s. were obtained from an authorized national service company and used for a simulation study. Table 1 shows a database of 200 records for the Forterra, Proxima, Proxima Plus, and Proxima Power models, in order of repairs done.

The time of individual visits to the repair shop (and the time the failure occurred, column number 4) is given in EOH (engine-operating hours).

If some of the data are censored, the K–M estimator can be used to estimate the survival function. If we assume that:

- the duration of the study was assumed from \( t_0 = 0 \) to \( T_{\text{max}} > 0 \),
- all times when:
  - \( m \)—number of new tractors joining the study population,
  - \( c \)—number of tractors leaving the study population,
  - \( d \)—number of tractors that have failed, are between 0 and \( T_{\text{max}} \): \( 0 < t_1 < t_2 < \ldots < t_k \leq T_{\text{max}} \),
- at time \( t_0 = 0 \) there are no failed tractors,
- at time \( t_i \) for \( i = 1, 2, \ldots, k \):
  - \( m_i \geq 0 \) new tractors joining the study population
  - \( c_i \geq 0 \) tractors leaving the study population,
Agronomy 2022, 12, 1364

- $d_i \geq 0$ tractors have failed to operate,
- then $m_i + c_i + d_i \geq 1$.

To use the K–M estimator for $i = 1, 2, \ldots, k$, by $n_i$ we denote the number of agricultural tractors being studied, shortly after the start of the study $t_{i-1}$ ($n_i$—is also, the number of tested tractors just before time $t_i$).

Then $n_1 = n_0$ and for $i = 2, 3, \ldots, k$:

$$n_i = n_{i-1} + m_{i-1} - c_{i-1} - d_{i-1}$$  \hspace{1cm} (1)

The unbiased estimator of the probability of surviving the period from time $t_i$, on the survival condition until time $t_{i-1}$, for $i = 1, 2, \ldots, k$ is:

$$\hat{p}_i = 1 - \frac{d_i}{n_i}.$$  \hspace{1cm} (2)

Thus, the K–M estimator of the survival function for time $t$ is:

$$\hat{S}(t) = \prod_{i \geq 0; t_i \leq t} \hat{p}_i$$  \hspace{1cm} (3)

where:
- $\hat{p}_0 = 1$
- $\prod i \geq 0 : t_i \leq t$ — the product of all $\hat{p}_i$, when $i$ takes all such indices greater than or equal to zero that $t_i \leq t$.

Especially for $i = 1, 2, \ldots, k$, there is $\hat{S}(t_0) = 1$ and a very practical dependency:

$$\hat{S}(t_1) = \hat{S}(t_{t-1}) \cdot \hat{p}_i.$$  \hspace{1cm} (4)

Table 1. Failures data for 200 of different models of Zetor tractors, in time $t_i$ [23].

<table>
<thead>
<tr>
<th>No.</th>
<th>Model</th>
<th>VIN</th>
<th>$t_i$ [EOH]</th>
<th>Symptoms of Failure</th>
<th>Broken Part</th>
<th>Cause of Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Forterra 95</td>
<td>… 1337</td>
<td>… 330</td>
<td>air escapes</td>
<td>quick coupler damaged sealing flange</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Forterra 115</td>
<td>… 6708</td>
<td>323</td>
<td>wrong sensor indication</td>
<td>air pressure sensor</td>
<td>short circuit</td>
</tr>
<tr>
<td>3</td>
<td>Forterra 125</td>
<td>… 2534</td>
<td>221</td>
<td>worn out mounts</td>
<td>wheel disc</td>
<td>inaccurate machining</td>
</tr>
<tr>
<td>4</td>
<td>Forterra 105</td>
<td>… 1286</td>
<td>130</td>
<td>no light</td>
<td>headlight</td>
<td>short circuit</td>
</tr>
<tr>
<td>5</td>
<td>Forterra 125</td>
<td>… 2408</td>
<td>220</td>
<td>worn out mounts</td>
<td>wheels</td>
<td>wheel disc</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
<td>…</td>
</tr>
<tr>
<td>196</td>
<td>Proxima 85</td>
<td>… 2657</td>
<td>… 610</td>
<td>shock absorption failure</td>
<td>cabin</td>
<td>gas spring</td>
</tr>
<tr>
<td>197</td>
<td>Proxima 85</td>
<td>… 3028</td>
<td>487</td>
<td>shock absorption failure</td>
<td>cabin</td>
<td>gas spring</td>
</tr>
<tr>
<td>198</td>
<td>Proxima 85</td>
<td>… 3028</td>
<td>487</td>
<td>engine hour meter failure</td>
<td>panel</td>
<td>short circuit</td>
</tr>
<tr>
<td>199</td>
<td>Proxima Plus 85</td>
<td>… 1690</td>
<td>4</td>
<td>voltage drop</td>
<td>battery</td>
<td>faulty battery cell</td>
</tr>
<tr>
<td>200</td>
<td>Forterra 125</td>
<td>… 2743</td>
<td>304</td>
<td>shock absorption failure</td>
<td>gas spring</td>
<td>damaged surface of sealing flange</td>
</tr>
</tbody>
</table>

3. Results

There are both failures $d_i$ of one, several parts, and entire systems, at the same time $t_i$, that is, with the same number of engine-operating hours. There are cases where two or three failures need to be repaired in a given tractor while servicing. For this reason, $\sum d_i = 140$ (which is less than the number of records).

Being renewable facilities, repaired tractors are reused, so $n_i$ is constant and amounts to 74 throughout the test period, that is, from $t_1 = 0$ EOH to $t_{121} = 1498$ EOH. Probability and survival function values were calculated, on the basis of dependencies 1–4 (Table 2).
Table 2. Sample data necessary to calculate the K-P estimator and its values, for the i-th times.

<table>
<thead>
<tr>
<th>i</th>
<th>$t_i$ [EOH]</th>
<th>$n_i$</th>
<th>$m_i$</th>
<th>$c_i$</th>
<th>$d_i$</th>
<th>$p_i$</th>
<th>$S(t_i)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>74</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>74</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>74</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>74</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0.97</td>
<td>0.93</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>74</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.99</td>
<td>0.92</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>74</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.99</td>
<td>0.91</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>117</td>
<td>999</td>
<td>74</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.99</td>
<td>0.16</td>
</tr>
<tr>
<td>118</td>
<td>1005</td>
<td>74</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.99</td>
<td>0.15</td>
</tr>
<tr>
<td>119</td>
<td>1041</td>
<td>74</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.99</td>
<td>0.15</td>
</tr>
<tr>
<td>120</td>
<td>1474</td>
<td>74</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.99</td>
<td>0.15</td>
</tr>
<tr>
<td>121</td>
<td>1498</td>
<td>74</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0.99</td>
<td>0.15</td>
</tr>
</tbody>
</table>

The results enable a classic presentation of the survival function (reliability) for the entire population of Zetor tractors and the selected model (Figures 3 and 4).

The shape of this function representing points has an exponential nature, with a negative shape parameter $\lambda = 0.002$ EOH$^{-1}$, so it is similar to the classic reliability diagram of technical facilities. Therefore, their average life expectancy is $1/(0.002$ EOH$^{-1}$) = 500 EOH. This value is very close to the average operation time of a tractor in a Polish farm, which is 12,000 EOH/20 years = 600 EOH, according to the methodology of Muzalewski [24]. Meanwhile, in Western Europe, such as Switzerland (method developed by FAT Tänikon, currently Agroscope Reckenholz-Tänikon), tractors generally work about 1000 EOH a year for 10 years [25].
Figure 3. The K–M survival function for Zetor brand farm tractors and the regression function. Own calculations made in an Excel spreadsheet.

Figure 4. The K–M survival function for models of Zetor Forterra tractors, within the range of 0–1498 EOH. Own calculations made in an Excel spreadsheet.

Further analysis is impossible, due to the lack of similar diagrams for other brands (regardless of methodology).

Such diagrams representing a specific model are more useful in agricultural practice. The proposed methodology requires a large random sample, i.e., \( n \geq 30 \) pieces of one model. Only the Forterra models (42 tractors of types 95, 105, 115, and 125) met the requirement for the available data, of 200 failure records. Figure 2 shows the survival function diagram, for \( i = 89 \), of this model.

Again, as expected, there is an exponential distribution here. However, again, it is a single diagram, which cannot be compared with other models from this manufacturer (there are 25 Proxima tractors, only Proxima Plus tractors, and 5 Proxima Power tractors). The median failure-free operation time of the Forterra model is lower (about 200 EOH) than for all Zetor models \( (\text{med} \approx 250 \text{ EOH}) \). It is recommended that the manufacturer improve the quality of this model, which is popular in Poland.

It is possible to plot both the survival function and the reliability function, with the Statistica program (Figure 5).

The proposed reliability rating method based on the K–Me survival function estimator more accurately reflects the susceptibility of technical facilities (tractors) to fail. The Reliability module from Statistica (ver. 13.1) treats all objects under analysis as non-renewable. The authors of this software indicate the potential possibilities of using the K–M estimator, placing it consciously in the Process Analysis/Weibull and Reliability/Failure Time Analysis Table. However, we operate here on the basis of a “black box”. The graph is obtained after entering the damage data, specifying the time of occurrence.

It is not the right approach for agricultural machinery and vehicles, which are, by definition, repairable (although a farm tractor consists of about 4000 parts, most of which are non-renewable, e.g., screws, nuts, bearings, cables, covers, operating fluids, etc.). The method is more accurate, but requires more detailed data, not only the failure times. It allows for the number of failures \( (d_i) \), objects leaving \( (c_i) \), and objects joining the test \( (m_i) \).

Both methods determine probability, which means that the extreme values they can take are 1 (one, start of the test) and 0 (zero, end of test, all objects have been scrapped).

Due to the fact that it is impossible to refer to similar failure diagrams for other brands of tractors, there are no reliability rankings or analyses with conclusions, for potential users and service technicians.
Again, as expected, there is an exponential distribution here. However, again, it is a single diagram, which cannot be compared with other models from this manufacturer (there are 25 Proxima tractors, only Proxima Plus tractors, and 5 Proxima Power tractors). The median failure-free operation time of the Forterra model is lower (about 200 EOH) than for all Zetor models ($\text{med} \approx 250 \text{ EOH}$). It is recommended that the manufacturer improve the quality of this model, which is popular in Poland.

It is possible to plot both the survival function and the reliability function, with the Statistica program (Figure 5).

Figure 5. The Kaplan–Meier estimate for Zetor tractors for $n = 121$ and $p = 0.95$ (data from Table 1).

4. Conclusions

The Kaplan–Meier method enables estimation of the survival function, directly from the continuous survival or failure times. The K–M method has an advantage over the life-table method, in the analysis of data on survival time or rate of failure-free operation, since estimators do not depend on the data-grouping method (grouping into a certain number of time intervals). The K–M method and the life-table method are identical, when life-table intervals contain, at most, one observation.

The K–M estimator works, only if there is detailed data on repairs to technical facilities. Spreadsheets (e.g., MS Excel) or programs for statistical calculations and the visualisation of results, such as the Statistica or R packages, can be used to determine the survival-time curves for agricultural machinery and vehicles.

Due to the lack of rankings based on such diagrams, it is difficult to make purchase decisions. When potential customers are choosing a particular brand/model, they are usually guided by the subjective opinions of long-term users or the opinions of experts, who describe specific models in the professional literature. Service technicians should, also, be interested in the reliability ratings of various brands of farm tractor, since they provide services and have direct contact with users. A good position in the ranking will, surely, translate into a better financial situation for the producers and dealers of these tractors. Rankings stimulate competition and, thus, improve the quality of products, which is the basis of every human activity.

The high reliability of agricultural machinery and tractors guarantees the performance of agrotechnical treatments, at optimal times. Hence, manufacturers and users should know the basics of the theory of reliability and, above all, the methods of its improvement at every stage of “machine life”. Automation and digitization of agriculture
4.0), by taking over the functions and relieving the operator, should, definitely, increase the reliability of work on the farm. However, it is, still, necessary to wait a few years for specific effects and, then, using the K–M method (or another, better method) described in this work, to estimate the numerical indicators, allowing for the creation of reliability rankings that are relevant to agricultural practice.

Improving the reliability, of currently produced agricultural machines and vehicles, will allow for digitization and automation, and, thus, the production of autonomous tractors and combines, as well as intelligent cultivation, spraying, and fertilizing machines.

To sum up, the K–M estimator, perfectly, describes the operation process of tractors (continuous in time and discrete in states), which, as repairable (renewable) objects, after repair, can successfully perform the work for which, according to the constructors, they were intended. In the case of tractors, as a tractive force (energy machine) for mounted, trailed, or suspended machines, which are mounted at the rear (traditionally) but, also, at the front and side. Thus, the proposed K–M method, originally developed for medicine, allows to successfully quantify the reliability of farm tractors.

Author Contributions: Conceptualization, K.D., A.E. and T.˙Z.; methodology, K.D., J.S. and M.W.; software, K.D., M.W. and K.W.; validation, K.D., J.S., M.W. and A.E.; formal analysis, K.D. and J.S.; data curation, K.D. and M.W.; writing—original draft preparation, K.D. and J.S.; writing—review and editing, K.D. and J.S.; visualization, M.W.; supervision, K.D. and J.S.; project administration, K.D. and K.W.; funding acquisition, K.D. All authors have read and agreed to the published version of the manuscript.

Funding: The publication was co-financed within the framework of the Polish Ministry of Science and Higher Education program: Regional Initiative Excellence, for the years 2019–2022 (No. 005/RID/2018/19), with a financing amount of PLN 12,000,000.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available on request.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References


