DirtSlurper3100

Ambra Balin   Marceli Morawski   Arthur Sliwinski   Davide Zanni

1913492   1816683   [Arthur's sn]   2371782

October 17, 2025

Abstract

1 Introduction

The DirtSlurper3100 is a state-of-the-art robotic vacuum cleaner manufactured by IButler. The company provides a five-year warranty for all its products, during which customers can receive free repairs for damages resulting from normal wear and tear, excluding any intentional damage. As the managers of IButler observed a higher-than-expected number of product returns, an investigation was conducted to assess whether the lifetimes of three key components met the manufacturer's stated expectations.
The components examined were the battery that powers the device, the IR sensor responsible for mapping the environment, and the impact sensor, which enables the robot to detect and react to physical obstacles. The manufacturer claimed that 90% of IR sensors would operate for at least 2,000 days and that the battery would retain a minimum of 80% of its capacity after 2,400 hours of continuous use. These claims represent key factors investigated in this report. Finally, since the impact sensor is a component developed by IButler and may be sold to a satellite company, its lifetime specifications will also be evaluated.
To better understand the behavior of these sensors, their survivability will be evaluated under different environmental conditions. The analysis will consider the effects of pet presence in the household and the number of carpets within the environment. To conduct these experiments, a selection of methods from survival analysis was applied. Both non-parametric estimators, such as the Kaplan–Meier estimator, and parametric approaches, including the Weibull accelerated failure-time (AFT) model,

were used for the initial inference. For the environmental analysis, hypothesis testing was performed using the log-rank test.

2 Data

The original dataset consists of 6473 entries. The first column in the dataset notes the date of registration of the vacuum cleaner. This date represents the day the appliance was supplied to its user. The total usage time contains information about how much the robot cleaner was used before one of its components broke down. The dataset also includes information on whether the household of the registered vacuum cleaner has any pets. This variable is binary and denoted with either "YES" or "NO". Additionally, a carpet score is also given using numbers 1 through 9. A low value means there is little area carpeted, whereas a higher value means there is more floor area that is carpeted. There is another binary variable that states whether the product was sent in for repairs or not. For each product that has been sent for repair, the date for this event is also mentioned in the column "Sent for repair". If a robot was not sent for repair, it will have this column marked as "-". Finally, there is a column for each part that is important for the vacuum cleaner: battery status, impact status, and IR status. For each entry in which the DirtCleaner3100 is sent for repair, it will have either "OK" or "Damage" for each part, depending on the assessment for each one of them. In case the product was not sent for repair, all three columns will be marked with "-" as well. There were no missing values, indicating good overall data quality.

The data contains all products that have been registered from the 1st of January 2015 up to and including the 31st of December 2019. The data is cut off at the latter date, so the dataset is right-censored.

2.1 Preprocessing

Since the data is censored, we created new columns to mark which entries are censored and which are not. There are 4 such columns. In the first case, the assumed event is the failure of the device; hence, the censored cases are those when the vacuum cleaner does not have a failure date. Additionally, we also look over the failure of each component, in which case the censored entries also include those that have a failed component that is not of interest. That is, for focusing on the battery, the only non-censored entries are those that have been sent for repair, and the value of 'Battery' is 'Damaged'. All cases where all three components have 'OK' despite having a failure date are regarded as censored since the device failed, but not because of an event we are interested in. Similar preprocessing was applied to the other two sensors, following the same methodology.

3 Methodology and Modeling Assumptions

In this section, both non-parametric and parametric estimation methods are presented and discussed. The statistical tools employed for inference and hypothesis testing are also described in detail, outlining their purpose and application within this analysis. Finally, the key assumptions underlying these methods are introduced.

3.1 Non-parametric estimate - Kaplan-Meier

The Kaplan–Meier estimator of the survival function, along with 95% pointwise confidence intervals calculated using Greenwood's formula, was applied to estimate the survival probabilities for each sensor and for the overall machine. This method was also used to infer survival at specific time points—2,000 days for each sensor and to support hypothesis testing.

Definition (Kaplan–Meier Survival Function Estimator):

$$d_j = \sum_{i=1}^{n} \delta_i \mathbf{1}\{y_i = t_j\} \quad \text{(number of items that fail at time } t_j\text{)}$$

$$r_j = \sum_{i=1}^{n} \mathbf{1}\{y_i \geq t_j\} \quad \text{(number of items "at risk" at time } t_j\text{)}$$

$$\hat{S}(x) = \prod_{j:t_j \leq x} \left(1 - \frac{d_j}{r_j}\right), \quad \text{where an empty product is defined as } 1.$$

Definition Variance (Greenwood Formula): For a given (deterministic) point $x \geq 0$,

$$\mathbb{V}(\hat{S}_{\mathrm{KM}}(x)) \approx \hat{S}^2(x) \sum_{j:t_j \leq x} \frac{d_j}{r_j(r_j - d_j)}.$$

3.2 Parametric estimate - Weibull AFT model

A parametric estimate used for inference for given sensors. Weibull accelerated failure-time (AFT) distribution, estimates the scale parameter α and the shape parameter β. Under the AFT parameterization, the Weibull survival function is defined as:

$$S(t) = \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right], \qquad \beta = \frac{1}{\hat{\sigma}}, \quad \alpha = \exp(\hat{\mu}).$$

DESCRIBE EVERY SYMBOL USED IN THE FORMULA

3.3 Inference

Add here for each sensor what estiamtes and sanity checks were used. Mention the assumptions if needed.

3.4 Hypothesis Testing

Available data were used to analyze how presence of pets and carpets affects the survival of each component. At a high level, the analysis was conducted by Kaplan–Meier (KM) estimates of the survival function per environmental factor, which were later evaluated using hypothesis testing with a log-rank test. Each test was conducted under the null hypothesis by log-rank test that both survival functions were identical, with the alternative hypothesis stating that they differed. A confidence level of 0.05 was applied to all tests.

$$\begin{aligned} &H_0 : S_1(t) = S_2(t), \quad \forall t \geq 0 \\ &H_1 : S_1(t) \neq S_2(t), \quad \text{for at least one } t \end{aligned}$$

Presence of pets. In total, three different scenarios were tested: the overall lifetime of the battery, the impact sensor, and the IR sensor. The presence of pets, encoded as a binary variable, resulted in two separate survival function estimates for each scenario.

Carpet level. For these tests, the data for carpets has been separated into three different groups: values 1-2, values 3-5, and values 6-9. These groups of data are named 'Low', 'Medium', and 'High' to represent the amount of carpet in the room where the vacuum cleaner is used. The tests have been conducted on the same scenarios as the pets.

3.5 Assumptions

The main assumption used in this analysis is that when a DirtSlurper has not been sent in for the repairs, none of it's components have failed, at least until the end of the study. This further implies the presence of right censoring.

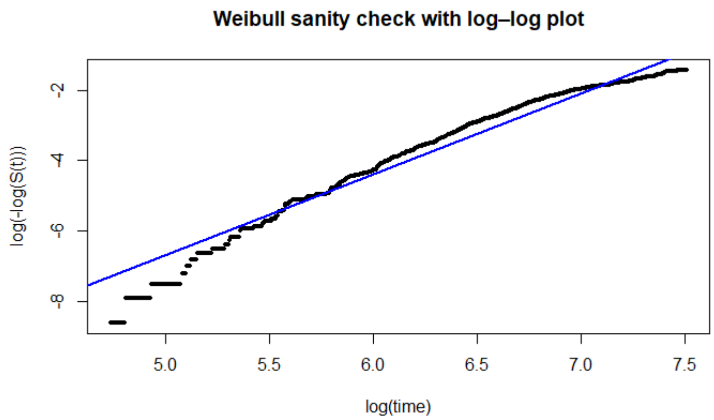

Figure 1: The full complementary log-log plot of $\log(-\log \hat{S}(t))$ versus $\log(t)$ shows small deviations from linearity at the tails, but the overall trend remains approximately linear, supporting the Weibull assumption

Lastly, there is a presence of truncation, as some of the produced robots, might not have been sent in time for the repairs, despite being reported for damage.

4 Results and Recommendations

4.1 Results from inference

Put here the results from infernce of each sesnor

4.1.1 IR sensor

4.1.2 Battery

4.1.3 Impact sensor

4.2 Results from Hypothesis Testing

4.2.1 Presence of pets

Effect of presence of pets on battery lifetime

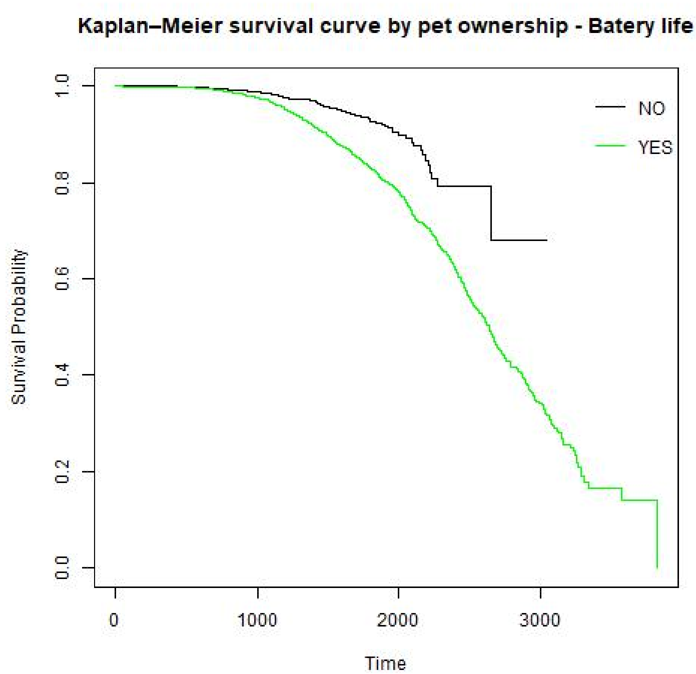

Figure 2: Plot of Kaplan–Meier survival functions for batteries in households with and without pets.

The p-value of corresponding test was equal to $4X10^{-13}$. This test showed enough evidence to reject the null hypothesis. By looking at the plot of both functions, it shows that the battery survival is worse, in presence of animal. This effect may be attributed to factors such as increased exposure to hair, dust, or physical interference caused by animal activity.

Effect of presence of pets on impact sensor lifetime

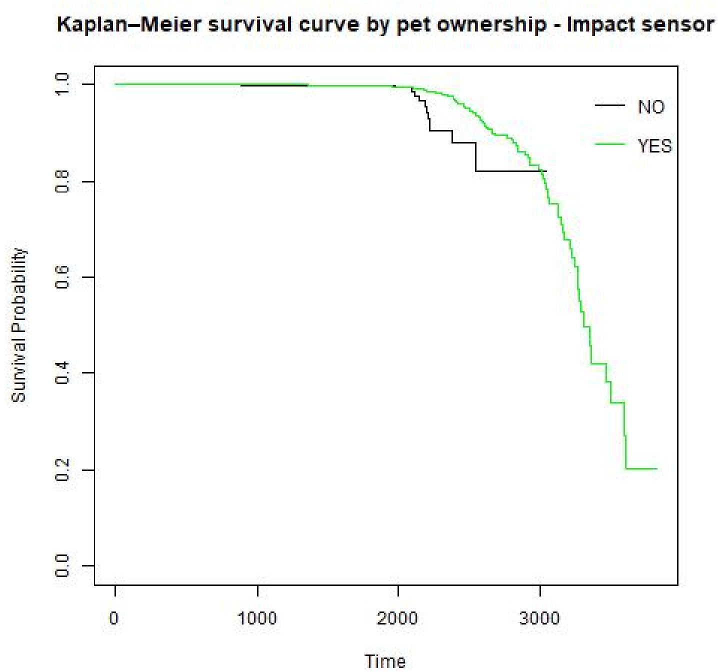

Figure 3: Plot of Kaplan–Meier survival functions for impact sensor in households with and without pets.

The p-value of the corresponding test was 0.005, providing sufficient evidence to reject the null hypothesis. Both the test results and the plotted survival functions indicate that robots operating in households with pets tend to have impact sensors that function longer compared to those in households without pets. However, due to the limited number of observations, the estimated functions are not sufficient to interpret this behaviour as causal.

Effect of presence of pets on IR sensor lifetime

The p-value of the corresponding test was $6x10^{-6}$, providing sufficient evidence to reject the null hypothesis. Both the test results and the plotted survival functions indicate that robots operating in households with pets tend to have IR sensor have

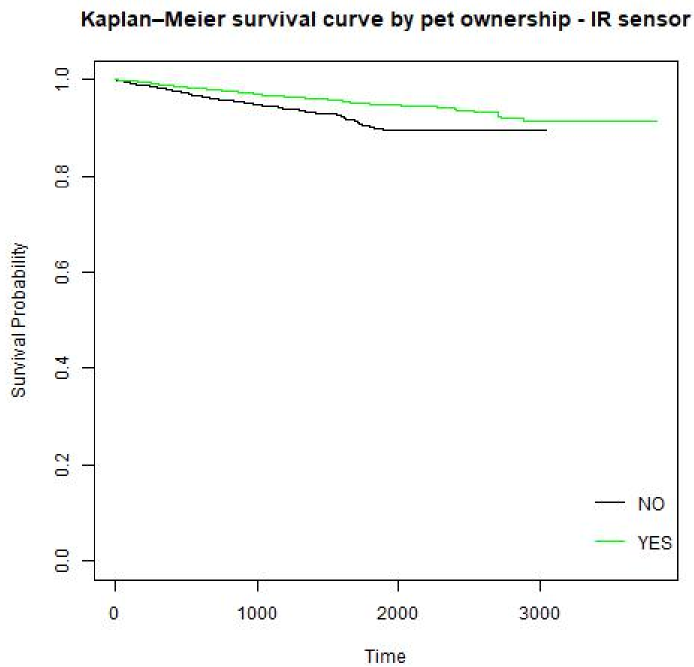

Figure 4: Plot of Kaplan–Meier survival functions for IR sensor in households with and without pets.

slightly higher survival probability over time than those in non-pet households. Yet, the difference between these survival functions seems to be negligible as there are not many failures across time.

4.2.2 Carpet levels

Effect of carpeted area on battery lifetime

The p-value for this test is $7e - 08$, so the test shows enough results to reject the null hypothesis. As seen in Figure 5, the higher the carpeted area the sooner the decline in battery survivability. A possible explanation for this behaviour is the extra effort the device needs to exert to run over carpeted areas.

Effect of carpeted area on impact sensor lifetime

The p-value for this test is $6e - 08$, so for this test, the null hypothesis can also be rejected. The impact sensor seems to benefit more from a carpeted area, as seen in figure 6. High levels of carpet give a constant survival line, which implies that all the entries are censored, that is, the sensor never failed when the carpet score was

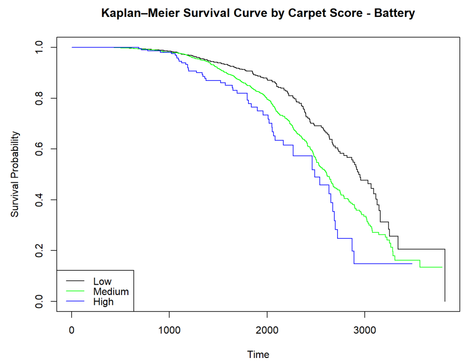

Figure 5: Kaplan-Meier estimator for three levels of carpet for the battery

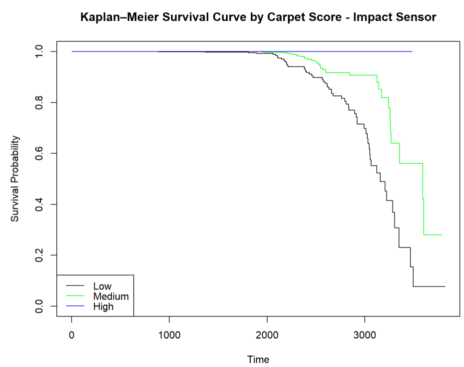

Figure 6: Kaplan-Meier estimator for three levels of carpet for the impact sensor

6 or above. A possible explanation for this behaviour is that carpeted areas slow down the device, leading to less impact

Effect of carpted area on IR sensor

4.3 Kaplan-Meier

KM estimates (with 95% CI):

$$\begin{aligned} S(365) &= 0.988\ [0.985,\ 0.991], \\ S(730) &= 0.937\ [0.929,\ 0.944], \\ S(1000) &= 0.886\ [0.875,\ 0.896] \\ S(1095) &= 0.871\ [0.859,\ 0.883]. \end{aligned}$$

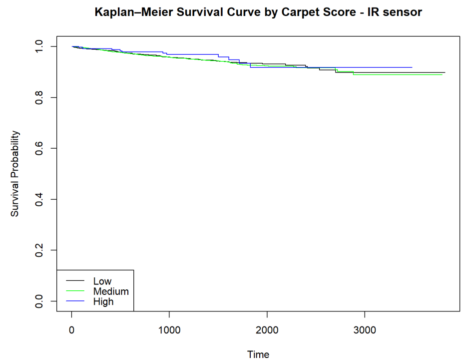

Figure 7: Kaplan-Meier estimator for three levels of carpet for the IR sensor

This indicates very high early-life reliability; around 11-12% fail by approximately 1000 days.

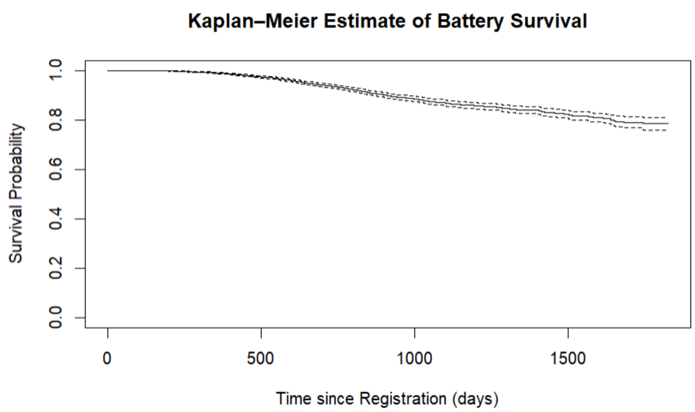

Figure 8: Kaplan-Meier estimate of $S(t)$ with 95% Greenwood confidence band.

4.3.1 Weibull fit and L_{10}

The Weibull AFT fit yields shape $\beta \approx 1.86$ and scale $\alpha \approx 3419$ days. The fact that $\beta > 1$ means the chance of failure increases over time, which is obvious. The 10% quantile is

$$L_{10} \approx 1020 \text{ days } (\approx 2.79 \text{ years}),$$

with bootstrap two-sided 95% CI $[975, 1067]$ days ($[2.67, 2.92]$ years). The one-sided lower confidence limits for L_{10} are 984 days (95%) and 991 days (90%), indicating the lower end of the plausible range for the 10% failure time. The fitted Weibull curve follows the KM survival estimates very closely, showing that the parametric model provides an accurate description of the observed battery lifetime data.

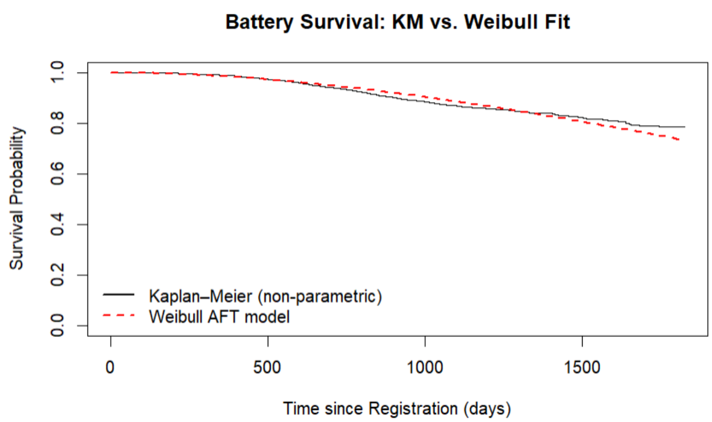

Figure 9: KM (black) vs. Weibull AFT (red, dashed).

4.4 Impact of usage intensity

Daily High vs. Low usage curves separate early and remain apart. The log-rank test gives $\chi^2 \approx 499$ (df = 1), $p < 2 \cdot 10^{-16}$, so they indicate a significant difference between the two groups. In particular, we reject equal survivor functions and we can say that heavy daily use substantially accelerates battery failure.

4.4.1 Summary table

Quantity	Estimate	95% CI lower	95% CI upper
$S(365)$	0.988	0.985	0.991
$S(730)$	0.937	0.929	0.944
$S(1000)$	0.886	0.875	0.896
Weibull β	1.86	(AFT estimate)	
Weibull α (days)	3419	(AFT estimate)	
L_{10} (days)	1020	975	1067
Log-rank χ^2	499	$p < 2 \cdot 10^{-16}$	

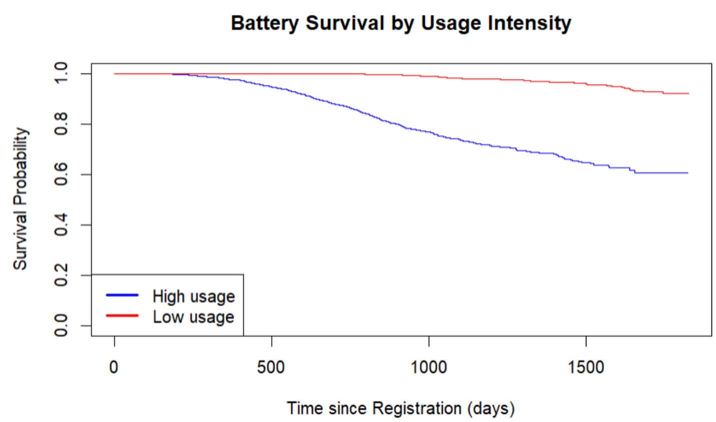

Figure 10: Battery survival by usage intensity (blue = High usage; red = Low usage).

5 Conclusion

The KM and Weibull AFT results both show high early-life survival and that the risk of failure increases gradually as they age. The estimated L_{10} is about 1020 days, meaning that around 10% of batteries are expected to fail after this time. This meets the manufacturer's target of $L_{10} \geq 1000$ days based on the best estimate. However, the lower confidence limits (95% = 984 days; 90% = 991 days) fall slightly below 1000, so the guarantee cannot be confirmed with full statistical confidence. Overall, the results continue to largely support the manufacturer's claims, especially when excluding devices intended for extreme use.

Usage intensity has a strong effect on battery life. Devices used more often tend to fail sooner, as confirmed by the log-rank test. In practice, heavy users should expect to replace batteries earlier than light users. From a service point of view, the L_{10} of about 2.79 years suggests preparing for battery replacements from the third year of normal use, and sooner for heavily used devices.

6 EXTRA PARTS

From the fitted (β, α) parameters, we derive the characteristic lifetime metric L_{10}, representing the time by which 10% of the batteries are expected to fail:

$$L_{10} = \alpha \left[-\log(0.9)\right]^{1/\beta}.$$

The uncertainty around L_{10} is calculated using a nonparametric bootstrap procedure with $B = 500$ resamples.

6.1 Usage effect

We define daily usage rate as `Total.usage.time/time_days`. Using the sample median as cut-off, we form "High usage" and "Low usage" groups and compare survival with a log-rank test.