Software Failure and Reliability Assessment Tool: Report

jesh

September 01, 2020

1 Data Selection and Analysis

1.1 Sampling the input SYS1 data

The table below shows the first 15 points of the input SYS1 data. 'FC' and 'CFC' (failure count and cumulative failure count) represent the index of a given failure. 'FT' (failure time) represents some units of testing time for which a failure will occur at. 'IF' (inter-failure time) represents the amount of time units between two adjacent failures. 'FN' denotes the number of a given failure.

Table 1: First 15 points of the input data

FN	IF	FT
1	3	3
2	30	33
3	113	146
4	81	227
5	115	342
6	9	351
7	2	353
8	91	444
9	112	556
10	15	571
11	138	709
12	50	759
13	77	836
14	24	860
15	108	968

1.2 Cumulative Failures

The following figure shows the SYS1 data as the cumulative number of failures ('FN') detected as a function of test time ('FT'). An increasing trend indicates periods where more faults are detected. Ideally, the cumulative number of failures will plateau to a horizontal line, indicating the occurrence of zero new faults.
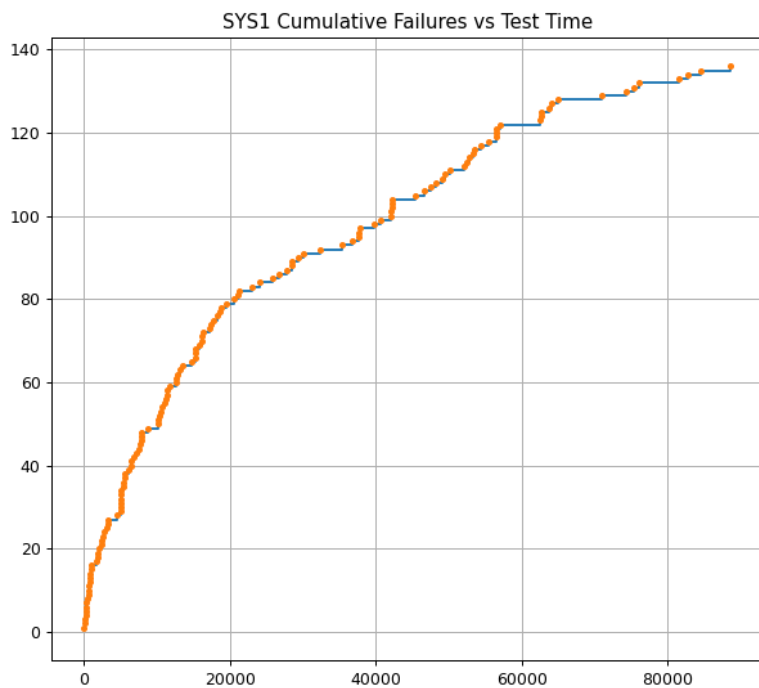

Figure 1: Cumulative Failures vs. Cumulative Test Time

1.3 Inter-failure Time

The following figure shows the SYS1 inter-failure times (i.e. times between failures, 'IF') as a function of testing time ('FT'). An increasing trend indicates periods where fewer faults are detected. Ideally, the inter-failure time increases maximally, indicating the detection of zero new failures.
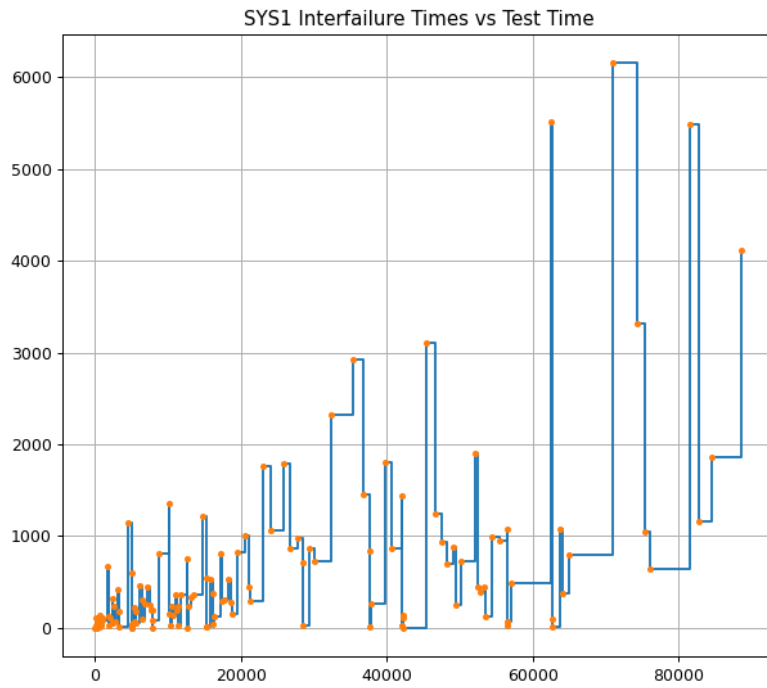

Figure 2: Inter-failure Time vs. Cumulative Test Time

3

1.4   Failure Intensity

The following figure shows the SYS1 data as the number of faults per unit time versus testing time ('FT'). A decreasing trend indicates periods of fewer fault detection. Ideally, the failure intensity should approach zero.
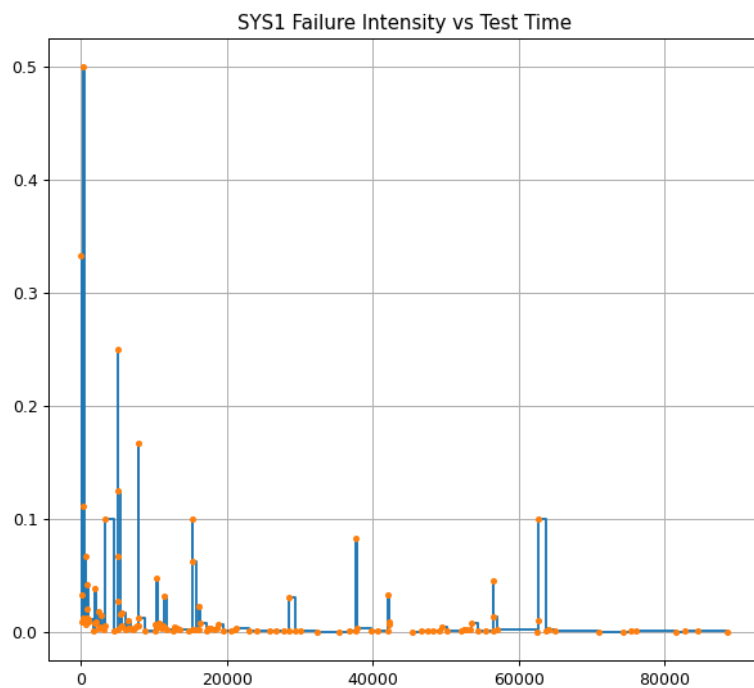

Figure 3: Failure Intensity vs. Cumulative Test Time

1.5 Laplace Trend Test

The following figure shows the Laplace test statistic for reliability growth as a function of cumulative test time (FT). A decreasing trend indicates reliability growth, while an increasing trend indicates reliability deterioration. The Laplace test statistic on the y-axis corresponds to the critical values of a normal distribution. This means that if the trend falls below a specific level, then we cannot reject the null hypothesis that the failure data suggests reliability growth at a specified level of confidence. The six black dot-dash style lines correspond to the 90%, 95%, 99%, 99.9%, 99.9999%, and 99.999999% confidence levels respectively. The red line is user-specified and has been set to 90%. Reliability growth is desired because software reliability growth models assume curves that exhibit increasing inter-failure times. If reliability growth is not present than the model fitting step may fail or produce predictions that are inaccurate. Therefore, the Laplace test statistic provides an objective quantitative measure for the analyst to decide if predictions may or may not be accurate.
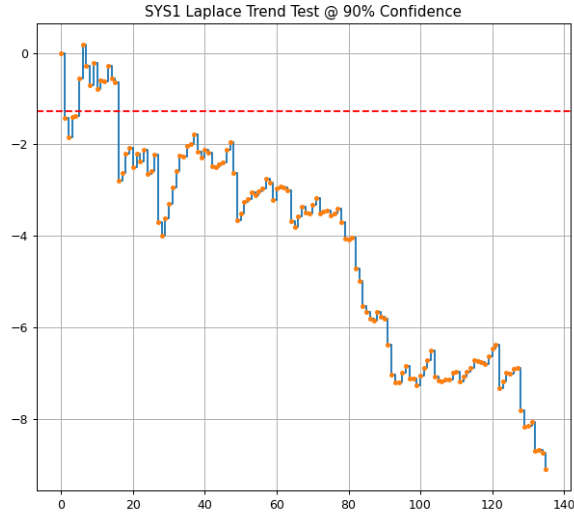

Figure 4: Laplace Trend Test at 90% Confidence, SYS1

1.6   Running Arithmetic Average

The running arithmetic average plots the average of the first k inter-failure times. An increasing trend indicates reliability growth, while a decreasing trend indicates reliability deterioration. This intuitively displays the direct relationship between increasing inter-failure times and the presented average.
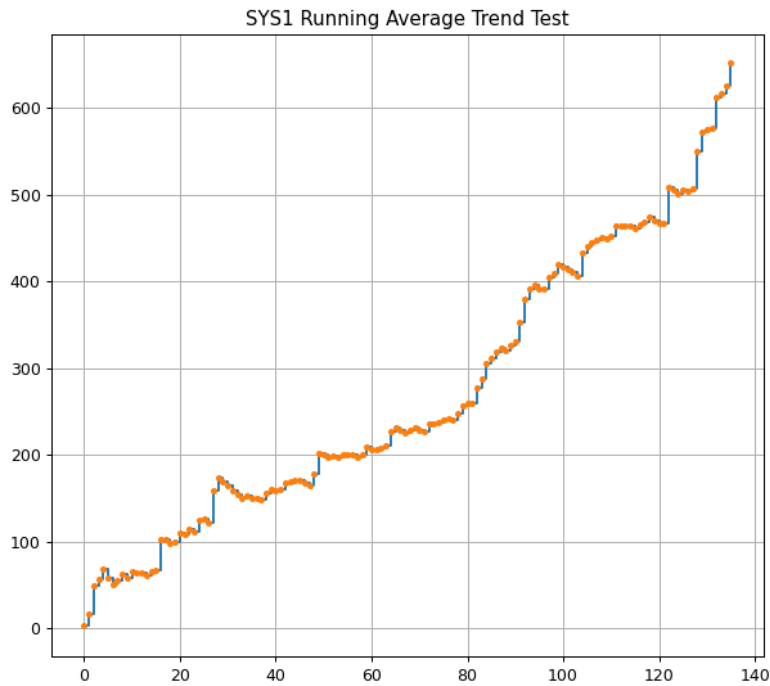

Figure 5: Running Arithmetic Average Test, SYS1

2 Apply Models to Data

2.1 Cumulative Failures

The following figure shows the fit of Geometric and Jelinski-Moranda models to the cumulative number of failures detected in the SYS1 data. The original data is also shown.
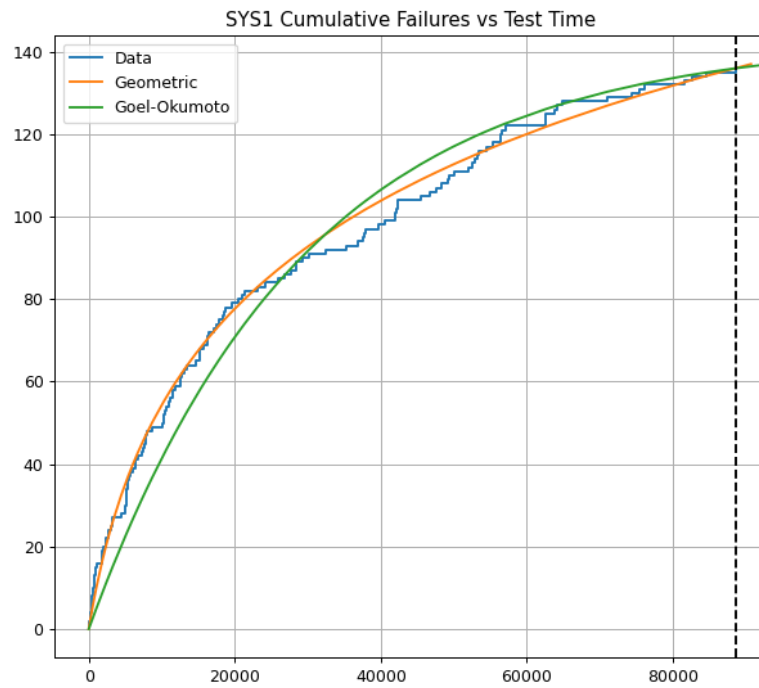

Figure 6: Cumulative Failures vs. Cumulative Test Time

7

2.2 Inter-failure Time

The following figure shows the fit of Geometric and Jelinski-Moranda models to the inter-failure times in the SYS1 data. The original data is also shown.
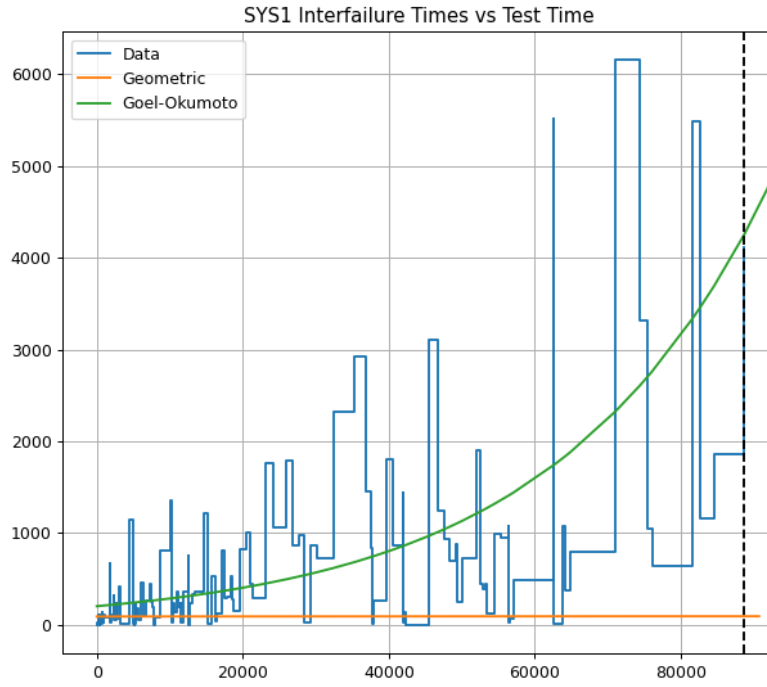

Figure 7: Inter-failure Times vs. Cumulative Test Time

2.3   Failure Intensity

The following figure shows the fit of Geometric and Jelinski-Moranda models to the failure intensity of the SYS1 data. The original data is also shown.
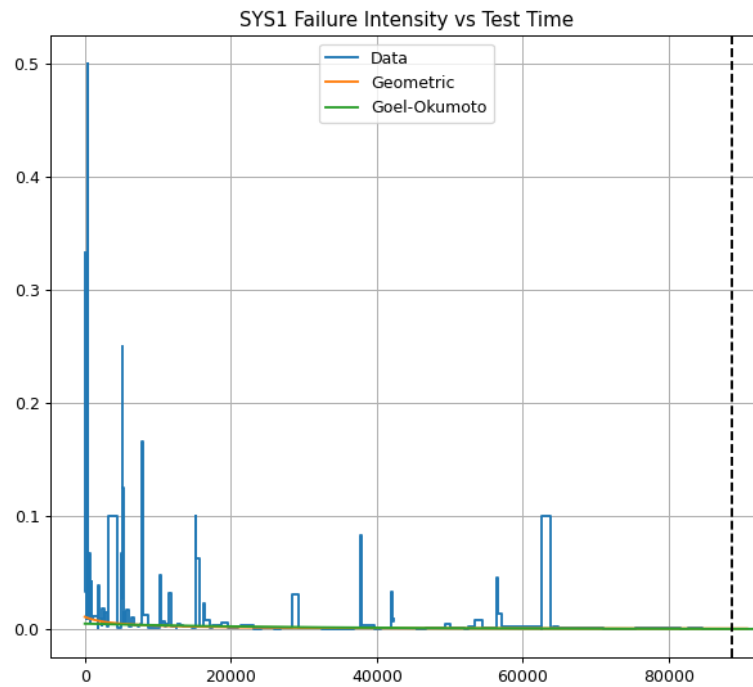

Figure 8: Failure Intensity vs. Cumulative Test Time

2.4   Reliability Growth

The following figure shows the reliability growth curve for the fit of the Geometric and Jelinski-Moranda models to the SYS1 data. This plot indicates a model's prediction that the software will be reliable (i.e. exhibit zero failures) for a duration of 4116 time units, as a function of testing time ('FT'). Assessing a model's reliability is a subjective choice made by the analyst, aided by statistical measures of goodness of fit. Omitting a reliability estimate within a report may be suggested by Laplace test statistic lacking reliability growth.
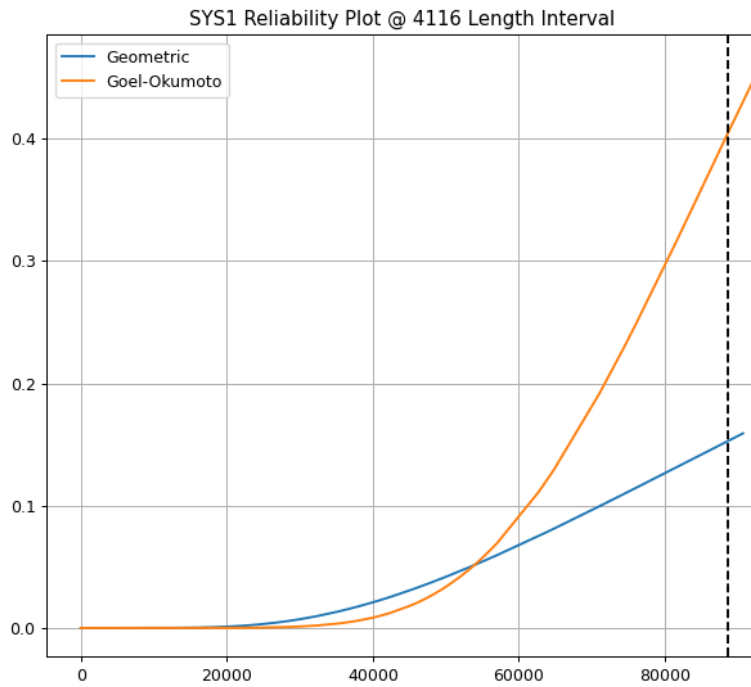

Figure 9: Reliability Growth vs. Cumulative Test Time

3 Querying Model Results

The following table shows inferences enabled by the models, including the time to achieve a reliability of 95% (i.e. the probability of zero failures for 4116 time units), plus the expected number of failures in the next 4116 and the expected time to observe an additional 10 failures.

Table 2: Model Result Query

Time, $R = 95\%$	Failures, $t = 4116$	Time, 10 Failures

4   Model Evaluation

The following table shows the measures of goodness of fit computed for the Geometric and Jelinski-Moranda models. The Aiaike Information Criterion ('AIC') is an information theoretic measure. Lower values are preferred. The model achieved the lowest AIC with SYS1 data. A difference of 2.0 or more in the AIC values of two models indicates a significant statistical preference. The Predictive Sum of Squares Error ('PSSE') used 90% of the SYS1 data to fit the models and computed the sum of squares between the remaining 10% of the data not used to fit the models. Lower values are preferred. The model achieved the lowest PSSE value on the SYS1 data. The measure of goodness of fit can help select a model, but the choice is ultimately subjective for the analyst.

Table 3: Model Evaluation

Akaike Information Criterion	Predictive SSE, 90% Data