Software Reliability Modeling

Technical Report SERC-2016-TR-104

Date: February 28, 2016

Principal Investigator:

Lance Fiondella, University of Massachusetts Dartmouth

Period of Performance:

March 1, 2015 to February 29, 2016

This material is based upon work supported, in whole or in part, by the U.S. Department of Defense through the Systems Engineering Research Center (SERC) under Contract H98230-08-D-0171 (Task Order 033, RT 48). SERC is a federally funded University Affiliated Research Center managed by Stevens Institute of Technology

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Department of Defense.

NO WARRANTY
THIS STEVENS INSTITUTE OF TECHNOLOGY AND SYSTEMS ENGINEERING RESEARCH CENTER MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. STEVENS INSTITUTE OF TECHNOLOGY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. STEVENS INSTITUTE OF TECHNOLOGY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

This material has been approved for public release and unlimited distribution.

Table of Contents

List of Figures

EXECUTIVE SUMMARY

The work performed under RT-139 includes:

(i) Design and implementation of an open source software reliability tool (SRT) to enable the automatic application of software reliability models. Two failure rate and three failure counting software reliability models have been implemented as well as two trend tests to assess if failure data exhibits reliability growth and two measures of goodness of fit. Plots of data and model fits including: cumulative failures, time between failure, failure intensity, and reliability growth. Inferences such as the time to achieve a target reliability or detect a specified number of additional failures have also been implemented. The source code is available from https://github.com/lfiondella/SRT and a web-based instance of the tool is accessible from http://www.sasdlc.org/. This report provides instructions for users. The contributors guide is available from: https://github.com/lfiondella/SRT/blob/master/documentation/contributors guide/SRT_contributors_guide.pdf.

(ii) Statistical algorithms based on the Expectation Maximization (EM) and Expectation conditional maximization (ECM) algorithms that are both stable and fast were developed to enhance the robustness of software reliability model fitting. Traditional algorithms suffered from instability, which required a high level of expertise to apply models. The details of these algorithms are reported in the technical publications that resulted from this research and given in Appendix A: List of Resulting Publications.

(iii) The PI will visit the Technical Lead Thierry Wandji at NAVAIR during the summer of 2016 through a summer fellowship from the Office of Naval Research. Together, they will use the tool on software failure data from DoD systems collected during testing. This will allow them to refine the tool and further enhance its suitability for use to ensure the reliability of software integrated into DoD systems.

INTRODUCTION

The Software Reliability Tool (SRT) is an application to estimate and predict the reliability of a software system during test and operation. It allows users to answer the following questions about a software system during test:

1. Is the software ready to release (has it achieved a specified reliability goal)?

2. How much more time and test effort will be required to achieve a specified goal?

3. What will be the consequences to the system's operational reliability if not enough testing resources are available?

The SRT runs under the R statistical programming framework and can be used on computers running Windows, Mac OS X, or Linux.

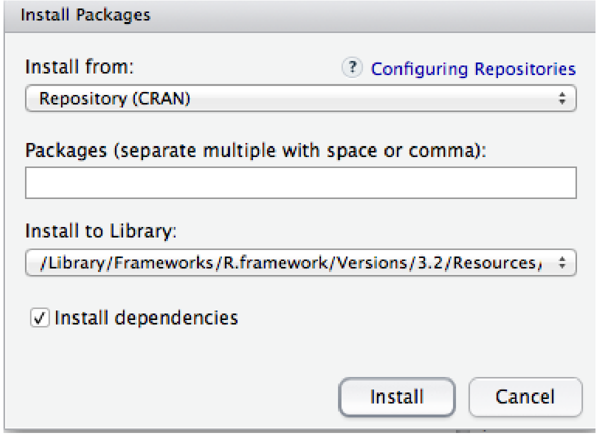
INSTALLATION

An automated installation script has been prepared and is available from the GitHub repository at: https://github.com/lfiondella/SRT/blob/master/install_script.R

The manual installation procedure is as follows:

- Make sure that your platform is either 64-bit Windows 7 or later, Mac OS X 10.9 or later, or a version of Linux capable of running R Studio.
- Perl 5 version 16 or later.
- Install R and RStudio on your machine. You can download both RStudio and R at rstudio.com. For Windows, Mac OS X, and Linux, you will need version R version 3.0 or later and RStudio 0.99.482 or later.
    - R (https://cran.r-project.org/) needs to be installed before RStudio can be installed. Once RStudio has been installed, the following packages need to be installed.
        - shiny
        - gdata
        - ggplot
        - DT
        - rootSolve
        - knitr
        - rmarkdown

- You can install the packages using the “Install packages…” menu item in RStudio’s “Tools” menu. That will bring up the following dialog box:



Install Packages

Install from: Repository (CRAN)

Packages (separate multiple with space or comma):

Install to Library: /Library/Frameworks/R.framework/Versions/3.2/Resources,

☑ Install dependencies

Figure 1 - Install Packages Dialog Box

Specify the package(s) you want to install in the first line of the dialog box. It is not necessary to change the installation location, so just leave the second line alone. Make sure that the “Install dependencies” box is checked – if is not, the packages you install may not work the way they should.

Once you have installed RStudio, you will want to configure your system so that R scripts (files with an “.R” extension) are opened with RStudio. This will make it easy to start the SRT.

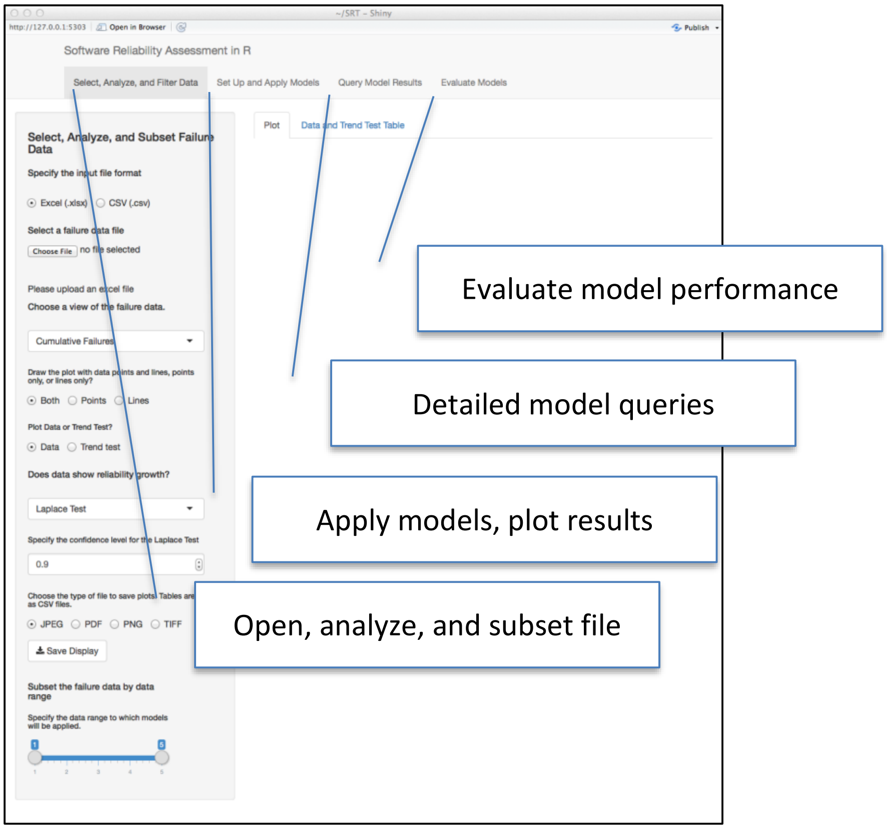
- Make a directory called “SRT” on a portion of your disk to which you have write access.
- Download the SRT files from https://github.com/lfiondella/SRT and copy the SRT files to the folder you have created in the previous step.

Using the SRT

Applying software reliability models consists of four main subtasks:

- Open a failure data history file and determine whether the data exhibits reliability growth that would justify model application.
- Apply one or more software reliability models and view the results.
- Query models to answer the following questions:
  - How many additional failures will be observed if testing is continued for an additional amount of time specified by the user?
  - How much additional time is required to observe a given number of additional failures?
  - How much additional testing time will be needed to achieve a specified reliability goal (or has that goal already been achieved)?
- Compare the models with one another to determine which one is the most likely to provide accurate predictions.

A self-contained area within the user interface has been designed for each of these subtasks. **Figure 2** indicates the initial view within The SRT after it has been opened.



Select, Analyze, and Subset Failure Data

Specify the input file format

● Excel (.xlsx) ○ CSV (.csv)

Select a failure data file

no file selected

Please upload an excel file

Choose a view of the failure data.

Cumulative Failures

Draw the plot with data points and lines, points only, or lines only?

● Both ○ Points ○ Lines

Plot Data or Trend Test?

● Data ○ Trend test

Does data show reliability growth?

Laplace Test

Specify the confidence level for the Laplace Test

0.9

Choose the type of file to save plots. Tables are as CSV files.

● JPEG ○ PDF ○ PNG ○ TIFF

Subset the failure data by data range

Specify the data range to which models will be applied.

1 – 5

Evaluate model performance

Detailed model queries

Apply models, plot results

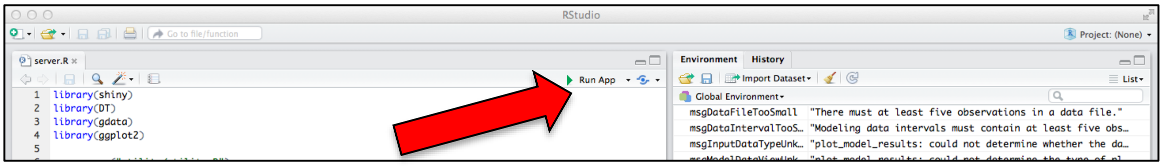
Open, analyze, and subset file

Figure 2 - Initial View within the SRT

Each of the four subtasks is contained within its own tab. Controls for each tab are grouped in a gray sidebar at the left side of the SRT window; plots and tables are displayed in the larger clear area at the center and right. Workflow proceeds from left to right, beginning with the leftmost tab and progressing to the right most tabs.

STARTING THE SRT

The easiest way to start the SRT is to open it with RStudio and then start it with RStudio's "Run App" button. Find the folder in which you have installed the SRT and double-click on the icon representing either "server.R" or "ui.R." Make sure you have associated R script files (".R" files) with RStudio as mentioned in the INSTALLATION Section. Once you have opened The SRT, the top of the RStudio window will resemble Figure 2:



```
library(shiny)
library(DT)
library(gdata)
library(ggplot2)
```

```
msgDataFileTooSmall    "There must at least five observations in a data file."
msgDataIntervalTooS…   "Modeling data intervals must contain at least five obs…
msgInputDataTypeUnk…   "plot_model_results: could not determine whether the da…
```

Figure 3 - Top of RStudio Window with The SRT Opened

To start The SRT, simply click on the "Run App" button indicated by the red arrow in **Figure 3.** The initial The SRT screen shown in **Figure 5** will then appear.

OPENING AND ANALYZING DATA

The first thing you will want to do once you have started The SRT is to open a set of failure history data. This is data that has been collected during testing from your problem reporting system which contains at least one of the following pieces of information:

- For each failure: the time elapsed between each failure observed. Equivalently, at what time was each failure observed?
- For each test interval: the length of the interval and number failures observed within it.

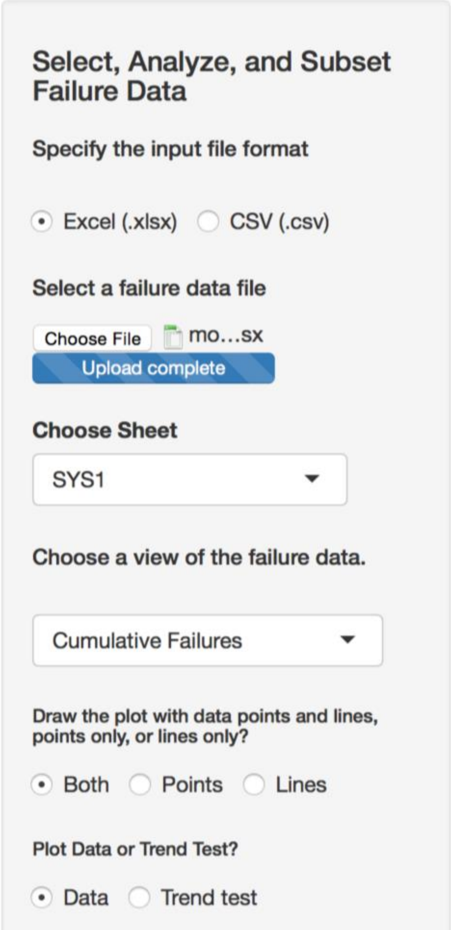
Any of these types of failure history can be read by The SRT. Examples of each of these are shown in Table 1 and Table 2 below. Table 1 shows part of a set of input data in the form of times between successive failures/failure times. The column labeled "FN" indicates the failure number, the column labeled "IF" is the time between each failure observed, and the column labeled "FT" is the total time since the start of testing. "CFC" denotes cumulative failure count.

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

Table 1 - Times Between Failures/Failure Times Input Data

T	FC	CFC
1	6	6
3	1	7
4.5	1	8
6	0	8
7	1	9
8	3	12
9	0	12
10.5	5	17
11	6	23
12	1	24

Table 2 – Failure Counts Input Data



Select, Analyze, and Subset Failure Data

Specify the input file format

● Excel (.xlsx) ○ CSV (.csv)

Select a failure data file

Choose File mo…sx

Upload complete

Choose Sheet

SYS1

Choose a view of the failure data.

Cumulative Failures

Draw the plot with data points and lines, points only, or lines only?

● Both ○ Points ○ Lines

Plot Data or Trend Test?

● Data ○ Trend test

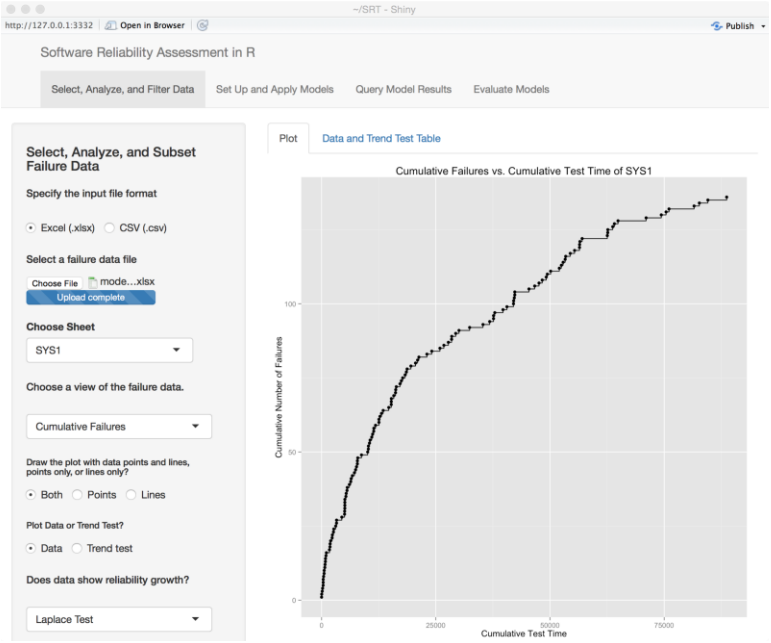
Figure 4 - Open and Plot Failure Data

Table 2 shows part of a set of failure data in the form of failure counts and test intervals. Each row represents a test interval: the column labeled "T" indicates the total amount of testing time – the difference between the "T" value in a row and the "T" value in the preceding row is the length of a test interval. The column labeled "FC" indicates the number of failures observed in that interval, and the column labeled "CFC" indicates the total failures observed.

The columns do not have to be in the order shown in the tables. However, they do have to be present, and the header rows have to be present for each type of data.

Open a Data File

Figure 5 – **Default Initial View of Failure Data** shows the controls you will use to open and plot a failure data file. These controls are on the "Select, Analyze, and Filter Data" tab. First, select whether to read an Excel file or CSV file. The default is Excel. Next, click the "Choose File" button to select the file you want. An "open file" dialog box will appear. If you open an Excel file, the "Choose Sheet" pull-down menu will appear, which allows you to select the sheet in the Excel file containing the data you want to work with. You can view the failure data in different ways by selecting an option from the "Choose a view of the failure data" pull-down menu. The default view is cumulative failures vs. elapsed test time. An example is shown in Figure 5. You can also view the input data in tabular form by clicking on the "Data and Trend Test Table" shown in Figure 5.


~/SRT - Shiny
http://127.0.0.1:3332
Open in Browser
Publish
Software Reliability Assessment in R
Select, Analyze, and Filter Data
Set Up and Apply Models
Query Model Results
Evaluate Models
Plot
Data and Trend Test Table
Cumulative Failures vs. Cumulative Test Time of SYS1
Cumulative Number of Failures
100
50
0
0
25000
50000
75000
Cumulative Test Time

Select, Analyze, and Subset Failure Data

Specify the input file format

● Excel (.xlsx) ○ CSV (.csv)

Select a failure data file

Choose File mode...xlsx

Upload complete

Choose Sheet

SYS1

Choose a view of the failure data.

Cumulative Failures

Draw the plot with data points and lines, points only, or lines only?

● Both ○ Points ○ Lines

Plot Data or Trend Test?

● Data ○ Trend test

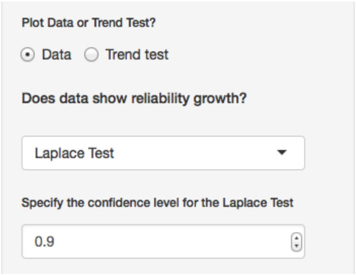
Does data show reliability growth?

Laplace Test

Figure 5 – Default Initial View of Failure Data

You can also decide whether the plot should show both data points and lines, lines only, or data points only. Use the radio buttons labeled "Draw the plot with data points and lines, points only, or lines only?" to make this choice. The default is both data points and lines.

ANALYZING FAILURE DATA FOR TRENDS



Plot Data or Trend Test?

● Data ○ Trend test

Does data show reliability growth?

Laplace Test

Specify the confidence level for the Laplace Test

0.9

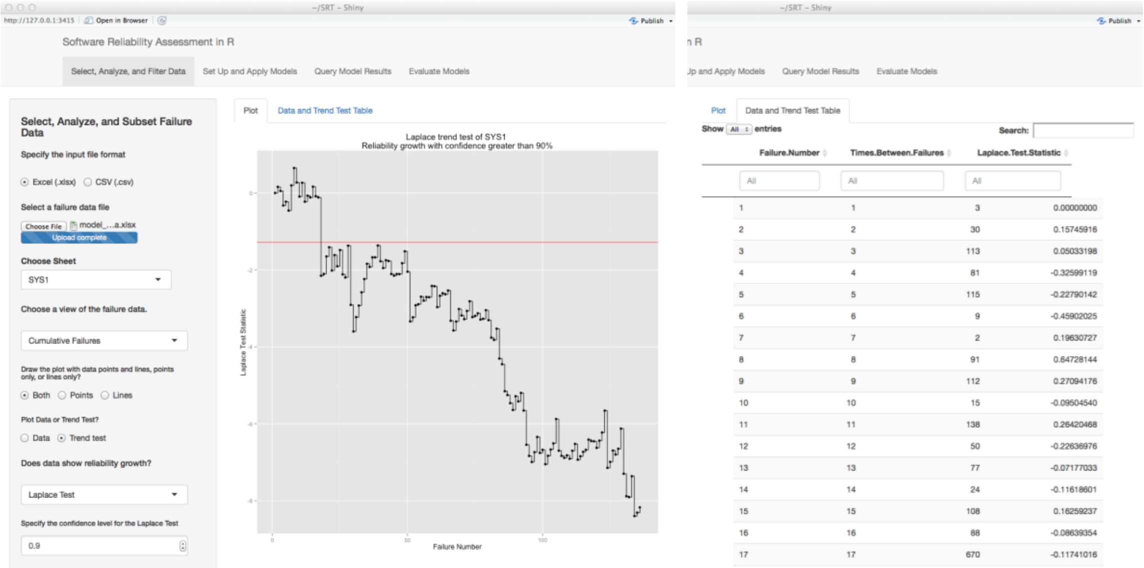
Figure 6 - Applying Trend Tests

Before you apply one or more models to the failure data file you have opened, it is important to check if it is appropriate to apply models. Since all software reliability models assume that the reliability of the software system being analyzed increases during testing as new defects are identified and removed, The SRT implements two trend tests to help you determine whether the failure data you are assessing exhibits reliability growth. To begin, click on the radio button labeled "Trend test" under the "Plot Data or Trend Test?" label shown in Figure 6. Selected trend test from the pull-down menu under the "Does data show reliability growth" label shown in Figure 6. The Laplace test is the default trend test. It possesses a statistical interpretation and allows the user to specify a confidence level between 0 and 1 to quantify a desired level of significance that the data exhibits (or does not exhibit) reliability growth. To specify the desired level of significance, enter the confidence value in the input area under the "Specify the confidence level for the Laplace Test" label in Figure 6. This input area and label appear if you have selected the Laplace test.

The second trend test, the running arithmetic average, simply computes and plots a running average of the times between failures. A plot exhibiting a positive slope indicates that the times between failures are increasing, suggesting reliability growth during testing.

This indicates that it is suitable to apply software reliability models to the data. A plot exhibiting a negative slope indicates the times between failures are decreasing with testing, suggesting decreasing software reliability. Software reliability models should **not** be applied to such failure data.

An example of the Laplace test is shown in **Figure 7** below.



~/SRT – Shiny
http://127.0.0.1:3415
Open in Browser
Publish
Software Reliability Assessment in R
Select, Analyze, and Filter Data
Set Up and Apply Models
Query Model Results
Evaluate Models
Plot
Data and Trend Test Table
Laplace trend test of SYS1
Reliability growth with confidence greater than 90%
Laplace Test Statistic
0
-2
-4
-6
-8
0
50
100
Failure Number

Select, Analyze, and Subset Failure Data

Specify the input file format

● Excel (.xlsx) ○ CSV (.csv)

Select a failure data file

Choose File model_....a.xlsx

Upload complete

Choose Sheet

SYS1

Choose a view of the failure data.

Cumulative Failures

Draw the plot with data points and lines, points only, or lines only?

● Both ○ Points ○ Lines

Plot Data or Trend Test?

○ Data ● Trend test

Does data show reliability growth?

Laplace Test

Specify the confidence level for the Laplace Test

0.9

Show All entries

Search:

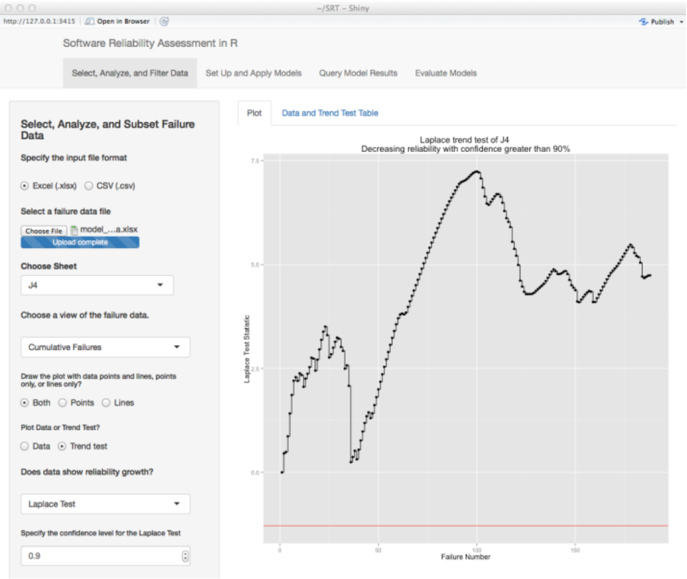
	Failure.Number	Times.Between.Failures	Laplace.Test.Statistic
	All	All	All
1	1	3	0.00000000
2	2	30	0.15745916
3	3	113	0.05033198
4	4	81	-0.32599119
5	5	115	-0.22790142
6	6	9	-0.45902025
7	7	2	0.19630727
8	8	91	0.64728144
9	9	112	0.27094176
10	10	15	-0.09504540
11	11	138	0.26420468
12	12	50	-0.22636976
13	13	77	-0.07177033
14	14	24	-0.11618601
15	15	108	0.16259237
16	16	88	-0.08639354
17	17	670	-0.11741016

Figure 7 – Laplace Trend Test Plot and Tabular Results

The left side of Figure 7 shows a plot of the Laplace test. The value of the test statistic changes with each failure in the data set; the final value is shown in the lower right corner of the plot. A decreasing trend in the value of the test statistic indicates that the data exhibits reliability growth. If the final value of the test statistic is less than the red horizontal line this means that the data exhibits reliability growth at a confidence level of at least that shown in the input area under the "Specify the confidence level for the Laplace Test" label. In this case, the title of the plot also indicates that the test indicates reliability growth at a 90% or greater confidence level. Thus, for this example, the confidence value is 0.9. In this case, the confidence may be high enough to apply models to the data. To improve your confidence, you can change the value in the "Specify the confidence level for the Laplace Test" input area. This will change the value for the horizontal red line, but the test statistic plot shown by the black line will remain the same.

The right side of Figure 7 shows the Laplace test results in tabular form, where the rightmost column is the value of the test statistic. Simply click on the "Data and Trend Test Table" tab to see this tabular form.

If the Laplace test statistic increases this indicates that the data does not exhibiting reliability growth. Figure 8 below illustrates this, as the plot increases to the upper right corner and the final value of the test statistic is above the horizontal red line. This indicates that we cannot be 90% confident the data exhibits decreasing reliability, in which case it is unlikely that a model can be applied successfully because the results may be inaccurate.



Select, Analyze, and Subset Failure Data

Specify the input file format

● Excel (.xlsx) ○ CSV (.csv)

Select a failure data file

Choose File model_....a.xlsx

Upload complete

Choose Sheet

J4

Choose a view of the failure data.

Cumulative Failures

Draw the plot with data points and lines, points only, or lines only?

● Both ○ Points ○ Lines

Plot Data or Trend Test?

○ Data ● Trend test

Does data show reliability growth?

Laplace Test

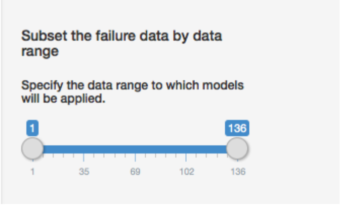
Specify the confidence level for the Laplace Test

0.9

Figure 8 – Laplace Trend Test Plot and Tabular Results – Decreasing Reliability

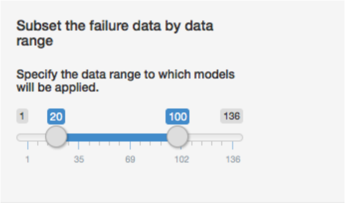
To return to a view of the failure data as shown in Figure 5, click the "Data" radio button under the "Plot Data or Trend Test" label shown in Figure 7.

SUBSET FAILURE DATA


Subset the failure data by data range

Specify the data range to which models will be applied.

Figure 9 – Subset Failure Data


Subset the failure data by data range

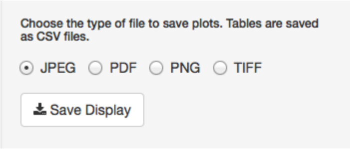
Specify the data range to which models will be applied.

Figure 10 - Subset Selection: Failures 20-100

You can select a subset of the data to work with instead of the entire data set. For example, it might be the case that only a portion of the testing represented by the data is appropriate for reliability testing. The first portion of the data set might represent regression testing, for instance, and only the portion after that is applicable to reliability testing. You can select a subset by using the double-ended sliding bar shown in Figure 9. This control is at the bottom of the control group shown in Figure 5, Figure 7, and Figure 8. The example shown in Figure 9 is for the data set shown in Figure 5, which has 136 failures. To select a starting point for the subset you want, click and drag to the right the round control on the left of the slider. If you want your subset to finish earlier than the last failure, click and drag to the left the round control on the right of the slider. If you want a subset of failure data between and including failures 20-100, move the slider controls to the positions shown in Figure 10. Once you have done this, the plot of the data will be updated to only show the subset you have selected.

If you are selecting a subset of the data, be sure that it includes 5 or more failures. If you try to include fewer than that, the SRT will display messages that you have selected too few failures in your subset and will automatically adjust the endpoints of the slider so that five failures are included in the subset.

SAVING DATA AND TREND TEST TABLES


Choose the type of file to save plots. Tables are saved as CSV files.

● JPEG ○ PDF ○ PNG ○ TIFF

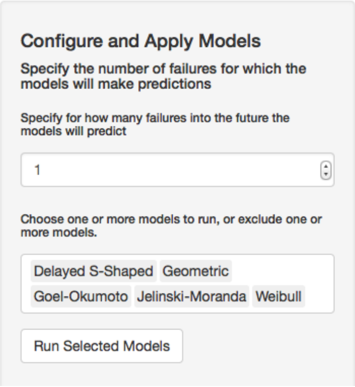
Figure 11 - Save Data Plots or Tables

To save a plot or data table being displayed, click the "Save Display" button near the bottom of the controls on the left side of the window. If you are looking at a data plot or a plot of trend test results, clicking that button will save that plot in the format you have chosen with the radio buttons shown in Figure 11. If you are looking at a tabular display of failure data or trend test results, the table you are viewing will be saved as a comma separated value (CSV) file.

APPLYING MODELS AND DISPLAYING RESULTS

Once you have opened a failure data file, the next step is to apply software reliability models to that data and view the results. You can do this on the second tab shown in Figure 2 - "Set Up and Apply Models." There are two groups of controls in the sidebar on the left side of the window. The first group allows you configure and select models, while the second group controls the way you display model results. There is also a "Save" button at the bottom of the controls that works the same way the "Save Display" button described in the Section SAVING DATA AND TREND TEST TABLES.

CONFIGURING AND APPLYING MODELS


Configure and Apply Models

Specify the number of failures for which the models will make predictions

Specify for how many failures into the future the models will predict

1

Choose one or more models to run, or exclude one or more models.

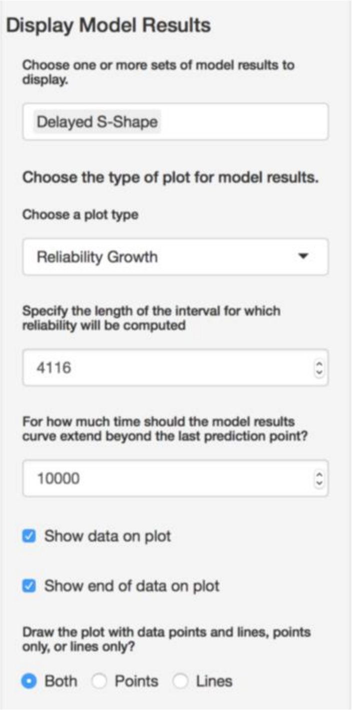
Delayed S-Shaped Geometric Goel-Okumoto Jelinski-Moranda Weibull

Figure 12 – Select, Configure, and Apply Models

Use the controls shown in Figure 12 to select models to run, configure the models, and apply the models to the failure data. Start by selecting the models you want to apply to the failure data. You can do this by choosing one or more models from the pull-down menu above the "Run Selected Models" button shown in Figure 12. The names of the models you have selected will appear below the label that reads "Choose one or more models to run, or exclude one or more models." By default, The SRT applies all of the models to the failure data, so you will see the names of all available models in this area after you have opened a failure data file. You can remove models by clicking in the area displaying the models you have selected and using the delete key to remove one or more models. The models you have deselected will reappear on the pull-down menu from which you choose them.

Once you have selected the models you want to apply, you will want to determine how far into the future you would like the models to make predictions. You can do this by entering a whole number greater than 0 into the input area labeled "Specify for how many failures into the future the models will predict." The default value is 1. Whatever number you enter, the models will make predictions for that many failures past the end of the failure data set (or subset) that you have selected. The next section illustrates how this affects the way model results are displayed.

SELECT MODEL RESULTS FOR DISPLAY AND CONTROL THE DISPLAY TYPE



Display Model Results

Choose one or more sets of model results to display.

Delayed S-Shape

Choose the type of plot for model results.

Choose a plot type

Reliability Growth

Specify the length of the interval for which reliability will be computed

4116

For how much time should the model results curve extend beyond the last prediction point?

10000

☑ Show data on plot

☑ Show end of data on plot

Draw the plot with data points and lines, points only, or lines only?

● Both ○ Points ○ Lines

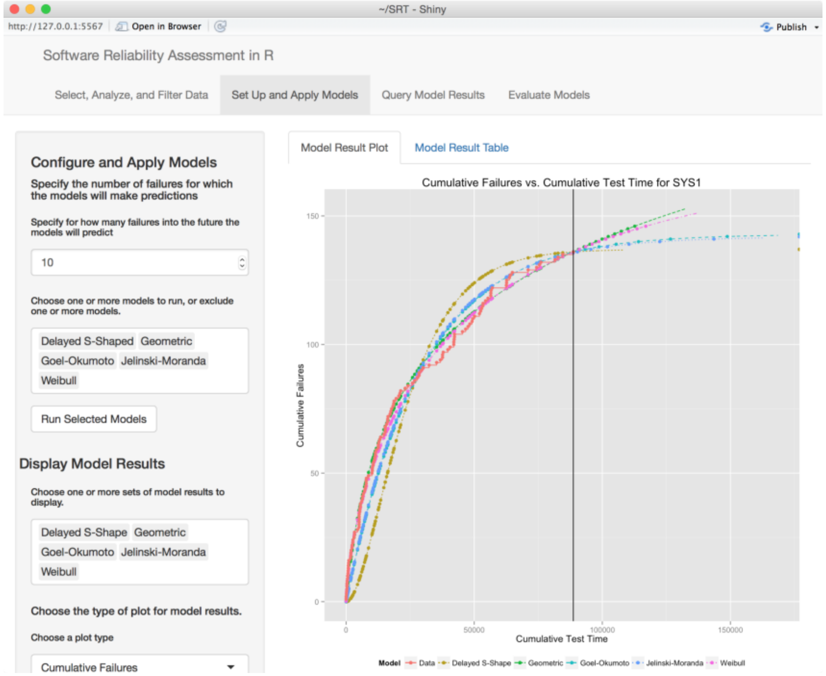
Figure 13 - Control Model Results Display

Once you have run one or more models, you can view plots of the model results and control the way the plots are shown using the controls shown in Figure 13. The names of the models that completed successfully are shown in a pull-down menu labeled “Choose one or more sets of model results to display.” You can select one or more sets of model results to be plotted. Using the “Choose a plot type” menu shown in Figure 13, you can view the model results as cumulative failures vs. elapsed time (the default view), failure intensity, times between failures, or reliability growth. If you have chosen to view reliability growth, an input area labeled “Specify the length of the interval for which reliability will be computed” appears in the display controls. This input area does not appear for the other types of displays. The input is simply the amount of time for which you would like to estimate the probability that the system will run without a failure. The default value of this input is the last non-zero value of time between successive failures in the data set you are analyzing. You can change it to any value greater than 0.

There is another input area just above the check boxes and radio buttons at the bottom of the display controls labeled “For how much time should the model results curve extend beyond the last prediction point?” When you are plotting model results, the plots show the models’ predictions into the future. The predictions can be shown as data points, curves drawn according to the model result equations, or both. If you are looking at the result curves, this input controls how far beyond the end of the last predicted failure the curves will be drawn. The default value is 10,000 time units; you can change this to any value greater than 0.

The controls at the bottom of Figure 13 determine the appearance of the model results plot. Checking the “Show data on plot” box displays the failure data on the same plot as the model results. This allows you to perform a qualitative visual comparison of the model results to the original failure data. If you check the “Show end of data on plot” box, a vertical black line is shown that passes through the final data point in the set of failure data you are analyzing. This allows you to easily distinguish the models’ predictions from the estimates the models make about the previous failure history. Finally, the radio buttons at the bottom of the controls let you plot the model results as both data points (each point represents a failure) connected by a curve drawn according to the model equations, data points only, or curves only. The default is to draw both data points and curves.

The examples below show model result plots using the data set shown in Figure 5. We have chosen to run all of the available models, and specified the models to make predictions up to 10 failures into the future. The model results curves are extended 20,000 time units into the future, instead of the 10,000 time unit default. Figure 14 shows the default plot as well as the estimated and predicted number of cumulative failures for all models.



Configure and Apply Models

Specify the number of failures for which the models will make predictions

Specify for how many failures into the future the models will predict

10

Choose one or more models to run, or exclude one or more models.

Delayed S-Shaped Geometric Goel-Okumoto Jelinski-Moranda Weibull

Display Model Results

Choose one or more sets of model results to display.

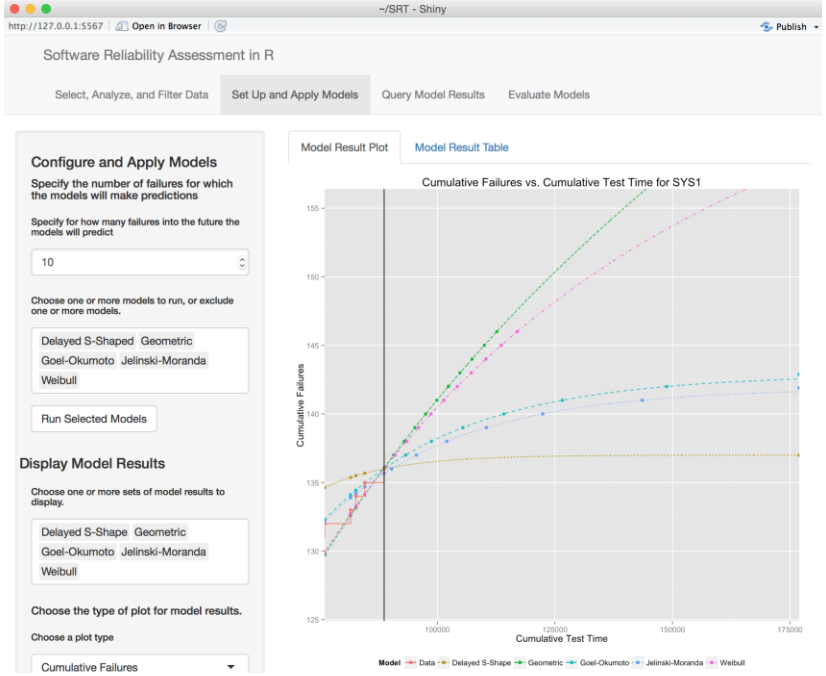
Delayed S-Shape Geometric Goel-Okumoto Jelinski-Moranda Weibull

Choose the type of plot for model results.

Choose a plot type

Cumulative Failures

Figure 14 - Default Model Results Display Showing All Models



Configure and Apply Models

Specify the number of failures for which the models will make predictions

Specify for how many failures into the future the models will predict

10

Choose one or more models to run, or exclude one or more models.

Delayed S-Shaped Geometric Goel-Okumoto Jelinski-Moranda Weibull

Display Model Results

Choose one or more sets of model results to display.

Delayed S-Shape Geometric Goel-Okumoto Jelinski-Moranda Weibull

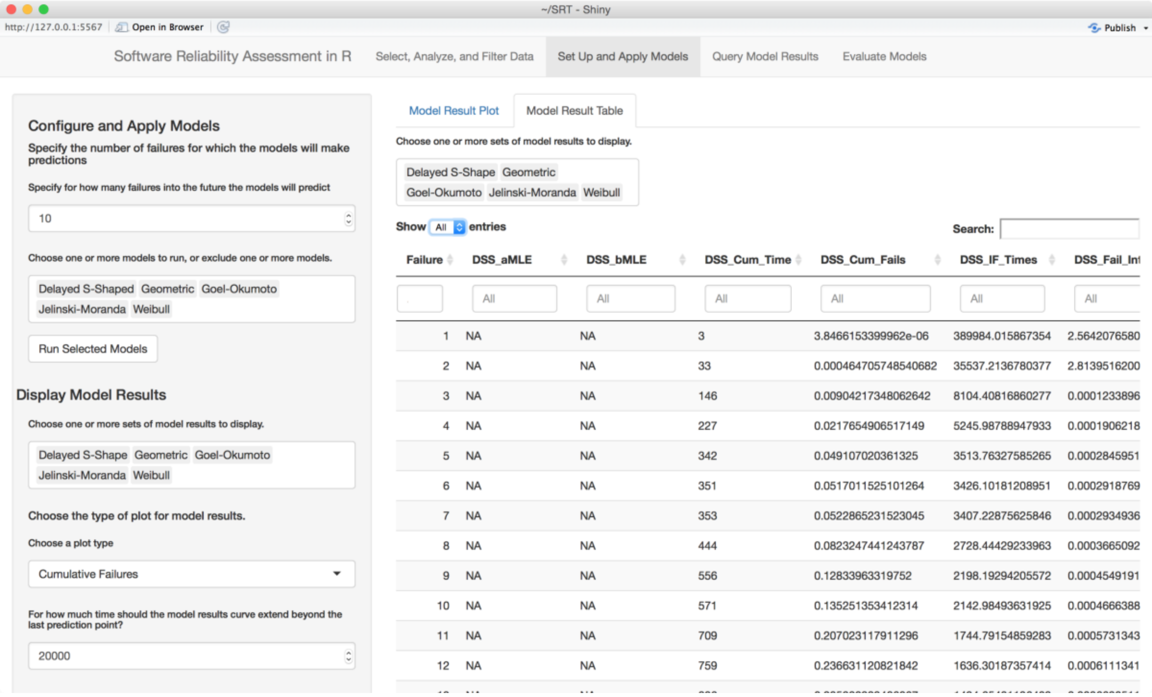
Choose the type of plot for model results.

Choose a plot type

Cumulative Failures

Figure 15 - Zooming into a Model Results Plot

You can zoom into an area of a plot using the “paint and double-click” mechanism. Click and drag over the region of the plot that you want to zoom into to define (“paint”) a rectangle, then double-click inside that rectangle. The enlarged area will appear as shown in Figure 15. To return to the original plot, just double click anywhere on the plot and the display will zoom back out to the original view.



Configure and Apply Models

Specify the number of failures for which the models will make predictions

Specify for how many failures into the future the models will predict: 10

Choose one or more models to run, or exclude one or more models.: Delayed S-Shaped, Geometric, Goel-Okumoto, Jelinski-Moranda, Weibull

Display Model Results

Choose one or more sets of model results to display.: Delayed S-Shape, Geometric, Goel-Okumoto, Jelinski-Moranda, Weibull

Choose the type of plot for model results.

Choose a plot type: Cumulative Failures

For how much time should the model results curve extend beyond the last prediction point?: 20000

Choose one or more sets of model results to display.: Delayed S-Shape, Geometric, Goel-Okumoto, Jelinski-Moranda, Weibull

Show All entries

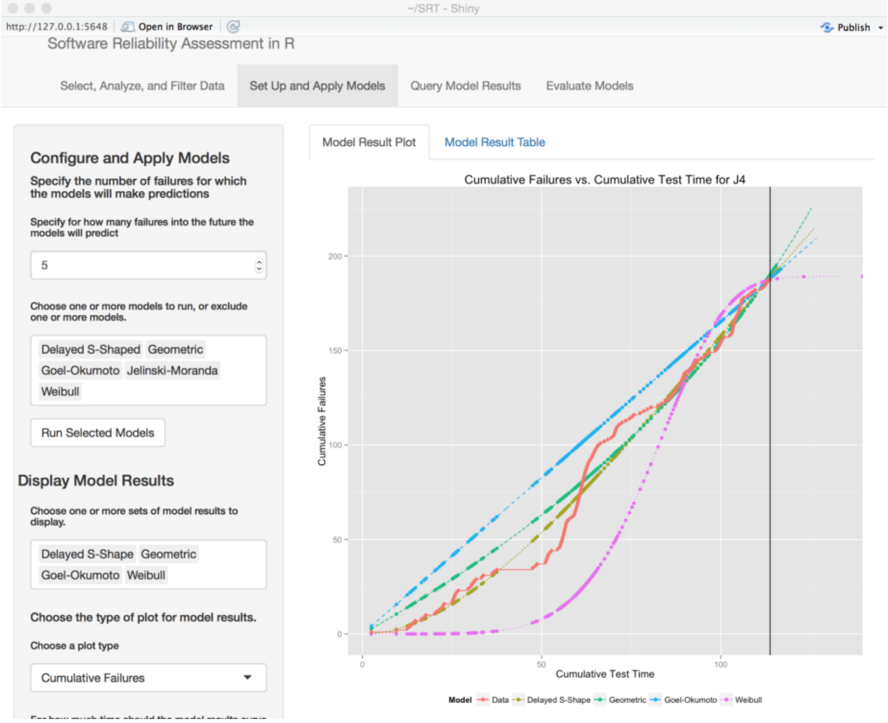
Search:

Failure	DSS_aMLE	DSS_bMLE	DSS_Cum_Time	DSS_Cum_Fails	DSS_IF_Times	DSS_Fail_Int
	All	All	All	All	All	All
1	NA	NA	3	3.8466153399962e-06	389984.015867354	2.5642076580
2	NA	NA	33	0.000464705748540682	35537.2136780377	2.8139516200
3	NA	NA	146	0.00904217348062642	8104.40816860277	0.0001233896
4	NA	NA	227	0.0217654906517149	5245.98788947933	0.0001906218
5	NA	NA	342	0.049107020361325	3513.76327585265	0.0002845951
6	NA	NA	351	0.0517011525101264	3426.10181208951	0.0002918769
7	NA	NA	353	0.0522865231523045	3407.22875625846	0.0002934936
8	NA	NA	444	0.0823247441243787	2728.44429233963	0.0003665092
9	NA	NA	556	0.12833963319752	2198.19294205572	0.0004549191
10	NA	NA	571	0.135251353412314	2142.98493631925	0.0004666388
11	NA	NA	709	0.207023117911296	1744.79154859283	0.0005731343
12	NA	NA	759	0.236631120821842	1636.30187357414	0.0006111341

Figure 16 - Tabular Display of Model Results

You can also display model results in tabular form as shown in Figure 16. Use the drop-down menu at the top of the table labeled “Choose one or more sets of model results to display” to select the model results you want to see. For each model you choose, you will see the model parameters, the cumulative time, the model’s estimates and predictions of the cumulative number of failures at that time, the estimated and predicted times between failures, the estimated and predicted failure intensity, and the estimated and predicted reliability. You can scroll up and down in the table as well as left and right.

As you gain experience using The SRT and work with an increasing number of failure data sets, it is likely that you will encounter a set of data for which one or more models are not able to produce results. One of the types of data sets that can cause this is a data set that does not exhibit reliability growth according to the trend tests (refer to the discussion in Section ANALYZING FAILURE DATA FOR TRENDS). In this case, only the names of the models that were able to run to completion will appear in the pull-down menu labeled “Choose one or more sets of model results to display” in Figure 14, Figure 15, and Figure 16. Figure 18 below provides an example of this situation, where you can see that all 5 models were chosen to run, but the Jelinski-Moranda model does not appear in the list of models you can display. For this particular data set, Figure 18 showed that the Laplace trend test did not exhibit reliability growth over time. Even in this situation, though, the tabular display of model results will allow you to select any of the models that were run, whether they completed successfully or not. This allows you to see details that might indicate why a model did not complete.



Configure and Apply Models

Specify the number of failures for which the models will make predictions

Specify for how many failures into the future the models will predict

5

Choose one or more models to run, or exclude one or more models.

Delayed S-Shaped Geometric Goel-Okumoto Jelinski-Moranda Weibull

Display Model Results

Choose one or more sets of model results to display.

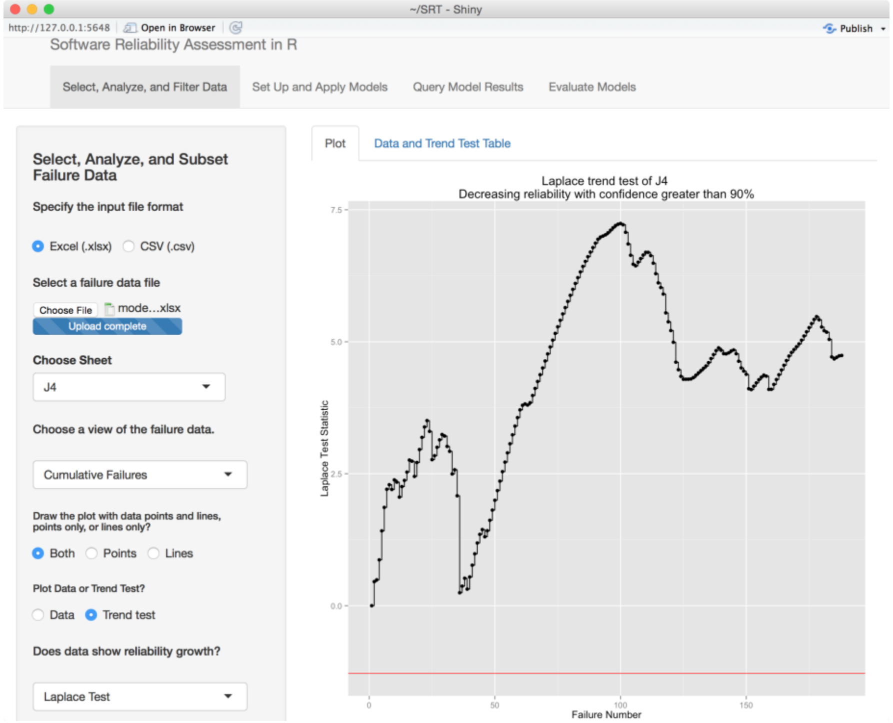
Delayed S-Shape Geometric Goel-Okumoto Weibull

Choose the type of plot for model results.

Choose a plot type

Cumulative Failures

Figure 17 - Data Set for Which One Model Could not Complete Successfully



Select, Analyze, and Subset Failure Data

Specify the input file format

● Excel (.xlsx) ○ CSV (.csv)

Select a failure data file

mode...xlsx

Choose Sheet

J4

Choose a view of the failure data.

Cumulative Failures

Draw the plot with data points and lines, points only, or lines only?

● Both ○ Points ○ Lines

Plot Data or Trend Test?

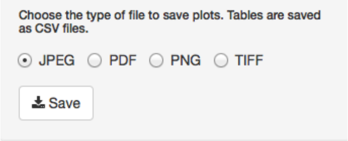
○ Data ● Trend test

Does data show reliability growth?

Laplace Test

Figure 18 – Laplace Trend Test For Data Set For Which One Model Couldn't Complete

SAVING MODEL RESULTS PLOTS AND TABLES

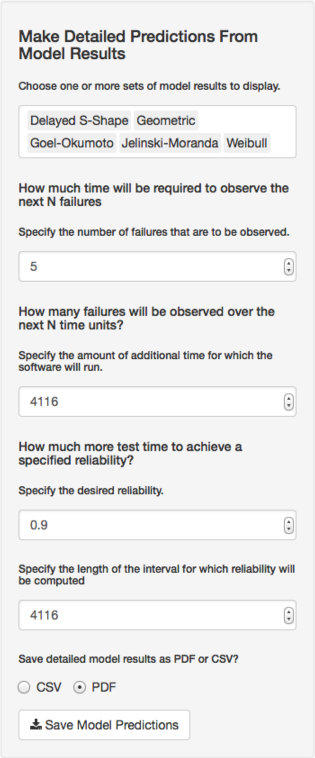

Choose the type of file to save plots. Tables are saved as CSV files.

⦿ JPEG ○ PDF ○ PNG ○ TIFF

Figure 19 - Save Model Result Plots or Tables

If you want to save the model results plot or table being displayed, click the "Save" button at the bottom of the controls on the left side of the window. If you are looking at a model results plot, clicking that button will save that plot in the format you have chosen with the radio buttons shown in Figure 19. If you are looking at a tabular display of model results, the table you are viewing will be saved as a comma separated value (CSV) file.

INTERACTIVE MODEL QUERYING


Make Detailed Predictions From Model Results

Choose one or more sets of model results to display.

Delayed S-Shape Geometric Goel-Okumoto Jelinski-Moranda Weibull

How much time will be required to observe the next N failures

Specify the number of failures that are to be observed.

5

How many failures will be observed over the next N time units?

Specify the amount of additional time for which the software will run.

4116

How much more test time to achieve a specified reliability?

Specify the desired reliability.

0.9

Specify the length of the interval for which reliability will be computed

4116

Save detailed model results as PDF or CSV?

○ CSV ⦿ PDF

Figure 20 - Detailed Model Queries

The plots and tabular displays of model results can provide a good overview of what the different models predict for the future failure behavior of the system you are analyzing. However, there are some details that are difficult to see in the plots and tables, so The SRT allows you make detailed queries of the models to answer the following questions:

- How many more failures am I going to observe in a given amount of time?
- How long will it take me to observe a given number of failures?
- How much more testing time will I need to obtain a given reliability for a specified operating time?

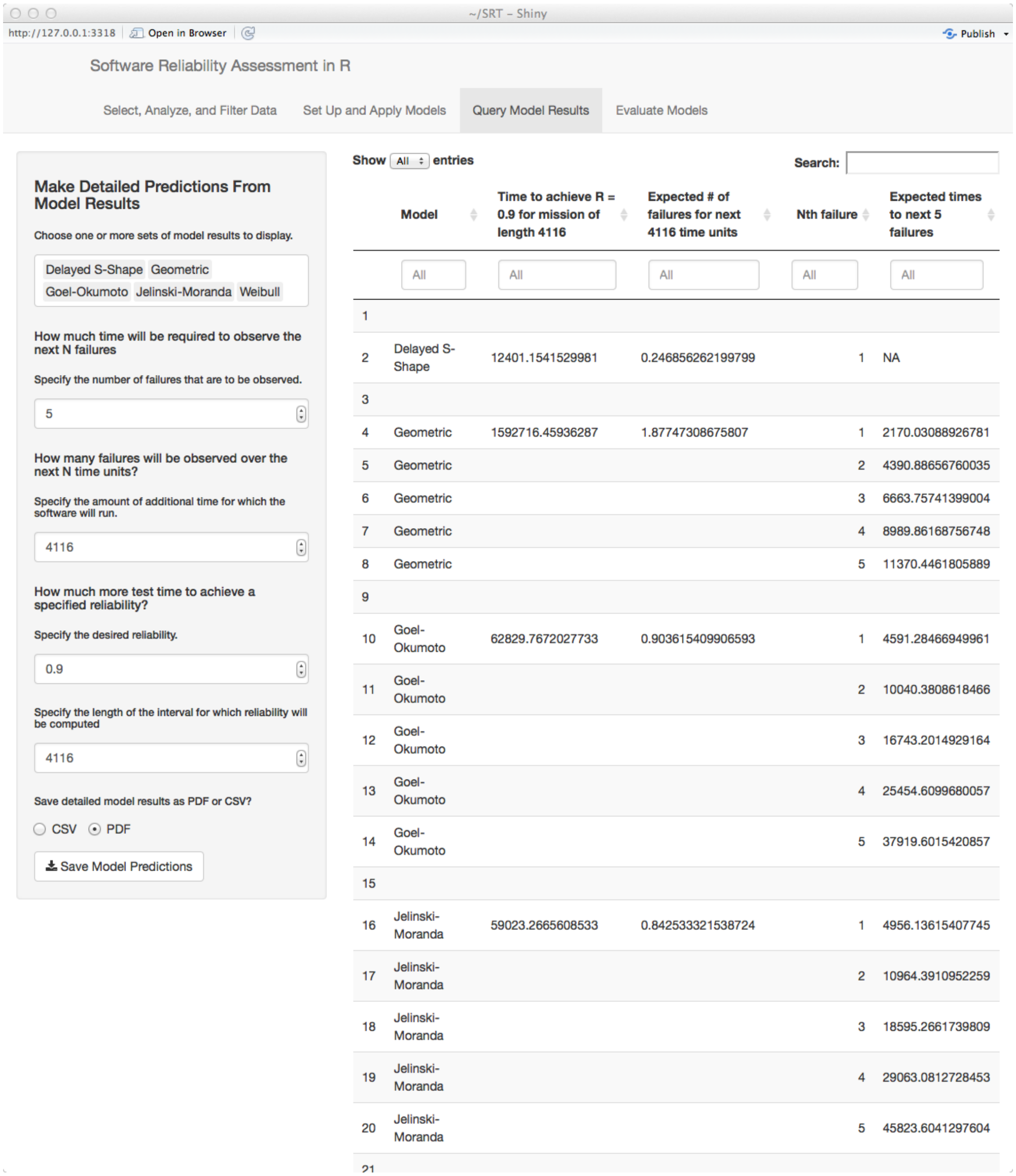
You can use the controls on the "Query Model Results" tab, shown in Figure 20, to answer these questions.

The pull-down menu at the top of the controls lets you choose one or more sets of model results to query. For this example, we have run all 5 models, and all 5 completed successfully. If there was a model that did not complete successfully, the name of that model will **not** show up in the pull-down menu.

The second control allows you determine how much more time you will need to observe a given number of failures. For this example, we will see how long it takes to observe the next 5 failures for each of the models chosen.

The third control allows you determine how many more failures we will observe in a given amount of future testing time. Since the models compute averages, we will not necessarily observe whole numbers in response to this query.

Finally, the last two controls allow you to determine how much more time you will need to obtain a specified reliability (the input area labeled "Specify the desired reliability") for a given amount of operational time (the input labeled "Specify the length of the interval for which reliability will be computed"). The response to this query is shown below in Figure 21.



Make Detailed Predictions From Model Results

Choose one or more sets of model results to display.

Delayed S-Shape Geometric Goel-Okumoto Jelinski-Moranda Weibull

How much time will be required to observe the next N failures

Specify the number of failures that are to be observed.

5

How many failures will be observed over the next N time units?

Specify the amount of additional time for which the software will run.

4116

How much more test time to achieve a specified reliability?

Specify the desired reliability.

0.9

Specify the length of the interval for which reliability will be computed

4116

Save detailed model results as PDF or CSV?

○ CSV ● PDF

Save Model Predictions

Show All entries Search:

	Model	Time to achieve R = 0.9 for mission of length 4116	Expected # of failures for next 4116 time units	Nth failure	Expected times to next 5 failures
	All	All	All	All	All
1					
2	Delayed S-Shape	12401.1541529981	0.246856262199799	1	NA
3					
4	Geometric	1592716.45936287	1.87747308675807	1	2170.03088926781
5	Geometric			2	4390.88656760035
6	Geometric			3	6663.75741399004
7	Geometric			4	8989.86168756748
8	Geometric			5	11370.4461805889
9					
10	Goel-Okumoto	62829.7672027733	0.903615409906593	1	4591.28466949961
11	Goel-Okumoto			2	10040.3808618466
12	Goel-Okumoto			3	16743.2014929164
13	Goel-Okumoto			4	25454.6099680057
14	Goel-Okumoto			5	37919.6015420857
15					
16	Jelinski-Moranda	59023.2665608533	0.842533321538724	1	4956.13615407745
17	Jelinski-Moranda			2	10964.3910952259
18	Jelinski-Moranda			3	18595.2661739809
19	Jelinski-Moranda			4	29063.0812728453
20	Jelinski-Moranda			5	45823.6041297604
21					

Figure 21 - Detailed Model Query Response

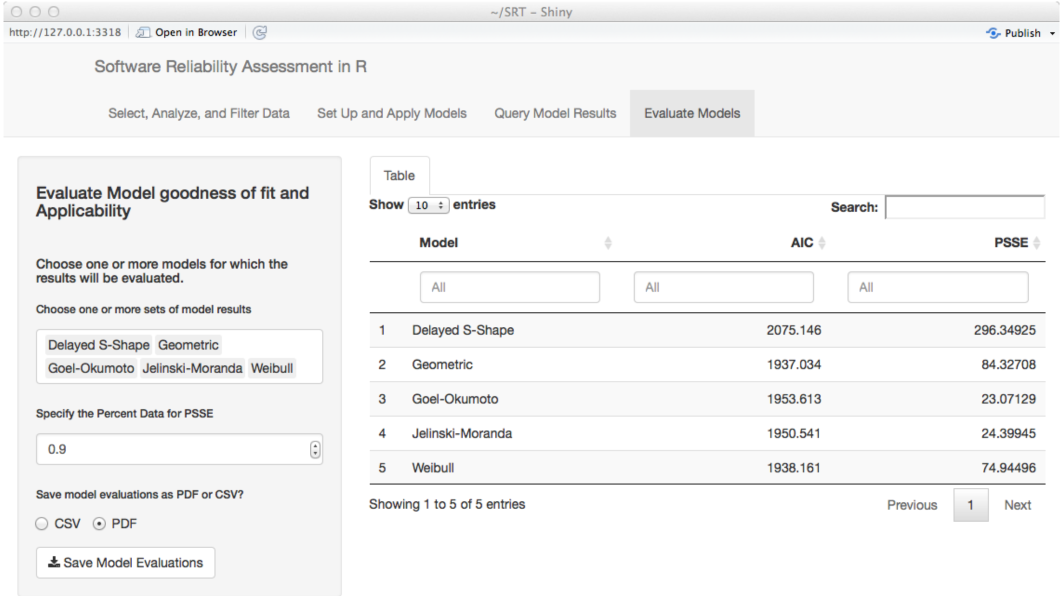
The response to the query is the 6-column table shown in Figure 21. The first column identifies the row of the table. The second column identifies the model whose results are shown in the next 4 columns. Note that there is a blank row between models. The third column tells you how much more time is needed to achieve a given reliability for a specified operating time. For this example, we use the default values for the reliability we want to achieve (0.9). The operating time is the length of the last non-zero value of times between failures in the data set we are analyzing (4116 time units). If a model's results indicate that the reliability objective has already been achieved, column 3 would say "R = 0.9 achieved."

Column 4 reports how many more failures you can expect to see in a given amount of time. Finally, columns 5 and 6 indicate how much more testing time you will need to observe the next N failures, where N ranges from 1 to the number of failures you specified in the "Specify the number of failures that are to be observed" input area. In this example, notice that the Delayed S-Shape model predicts that fewer than 5 failures remain to be discovered in the system being tested. Thus, you only see the number of rows indicating the number of failures remaining to be discovered. If a model indicates that the amount of time needed to discover a failure is infinite, you will see a value of "NA" in column 6, as is shown for the Delayed S-Shape model in this example.

Finally, you can save the results of the query as either a CSV or PDF file by clicking the "Save Model Predictions" button. Clicking this button will bring up a dialog box in which you specify the name of the file you want to save and the folder in which you want to save it. The type of file the query results are saved at is controlled by the radio buttons just above the "Save Model Predictions" button. The default file type is PDF. You can also save the detailed results as a CSV file.

EVALUATING MODEL PERFORMANCE AND APPLICABILITY

After applying one or more models to a set of failure data, you will want to know which model or models produce better predictions. This version of The SRT includes two ways of evaluating the performance of the models you have run to identify those that will give you better predictions – the Akaike Information Criterion (AIC) and Predictive Sum of Squared Error (PSSE). These are shown in tabular format in Figure 22.



Evaluate Model goodness of fit and Applicability

Choose one or more models for which the results will be evaluated.

Choose one or more sets of model results: Delayed S-Shape, Geometric, Goel-Okumoto, Jelinski-Moranda, Weibull

Specify the Percent Data for PSSE: 0.9

Save model evaluations as PDF or CSV?

○ CSV ● PDF

Save Model Evaluations

Show 10 entries

Search:

	Model	AIC	PSSE
	All	All	All
1	Delayed S-Shape	2075.146	296.34925
2	Geometric	1937.034	84.32708
3	Goel-Okumoto	1953.613	23.07129
4	Jelinski-Moranda	1950.541	24.39945
5	Weibull	1938.161	74.94496

Showing 1 to 5 of 5 entries

Previous 1 Next

Figure 22 - Evaluating Model Performance

The controls on the left side of Figure 22 allows you to pick the models you want to compare, set up the PSSE analysis, and save the model evaluations as a CSV or PDF file. The pull-down menu labeled "Choose one or more models for which the results will be evaluated" identifies those models that completed successfully. You can choose one or more models to compare from this menu. To perform the PSSE evaluation, you will need to choose a subset of the data to create an evaluation model, which is then applied to the remainder of the data. The square of the differences between the model's predictions for the remainder of the data are then computed and summed for each of the models you have chosen. A lower PSSE value indicates less distance between the model's predictions and the actual data, which in turn suggests a better-performing model.

The AIC is an information theoretic measure of the relative likelihood of the models being compared. A lower AIC value for a model indicates that it is more likely that the model characterizes the data well. The idea is to identify the model that minimizes the information lost. A model that minimizes the information loss is more likely to perform better and therefore be a more appropriate model. The AIC tells you how much more likely it is that model 2 minimizes the information loss than model 1. Note that AIC does not tell you anything about how well the model results fit the data in an absolute sense. If none of the model results fit the data well, AIC will not be able to warn you of this.

As with detailed model queries, you can save the results of model evaluations as either a CSV or PDF file by clicking the "Save Model Evaluations" button. Clicking this button will bring up a dialog box in which you specify the name of the file you want to save and the folder in which you want to save it. The type of file the evaluation results are saved in is controlled by the radio buttons just above the "Save Model Evaluations" button. The default file type is PDF; you can also save the model evaluations as a CSV file.

CONCLUSION

An open source software reliability tool (SRT) has been implemented and the source code made available from https://github.com/lfiondella/SRT. A web-based instance of the tool is accessible at http://www.sasdlc.org/. This report provided instructions for users. The contributors guide is available from: https://github.com/lfiondella/SRT/blob/master/documentation/contributors guide/SRT contributors guide.pdf. Statistical algorithms based on the Expectation Maximization (EM) and Expectation conditional maximization (ECM) algorithms that are both stable and fast were developed to enhance the robustness of software reliability model fitting. The details of these algorithms are reported in the technical publications that resulted from this research and given in Appendix A: List of Resulting Publications. The PI will visit the Technical Lead Thierry Wandji at NAVAIR during the summer of 2016 through a summer fellowship from the Office of Naval Research. Together, they will apply the tool to software failure data from DoD systems collected during testing. This will allow them to refine the tool and further enhance its suitability for use to ensure the reliability of software integrated into DoD systems.

APPENDIX A: LIST OF RESULTING PUBLICATIONS

V. Nagaraju, *Algorithms for Nonhomogeneous Poisson Process (NHPP) Software Reliability Models and Software Rejuvenation considering Correlated Failures*, Master of Science in Computer Engineering, University of Massachusetts, Dartmouth, December 2015.

V. Nagaraju, L. Fiondella, C. Jayasinghe, P. Zeephongsekul, and T. Wandji, Performance Optimized Expectation Conditional Maximization Algorithm for Non-homogeneous Poisson Process Software Reliability Models, In preparation for the *IEEE Transactions on Reliability*.

V. Nagaraju, A. Krishna Murthy, and L. Fiondella, P. Zeephongsekul, and T. Wandji, Expectation Conditional Maximization Algorithms for Failure count Non-homogeneous Poisson Process Software Reliability Models, Submitted to the *International Conference on Reliability and Quality in Design*, Aug 2016.

V. Nagaraju, K. Katipally, R. Muri, T. Wandji, and L. Fiondella, An Open Source Software Reliability Tool: A Guide for Users, Submitted to the 2016 International Conference on Reliability and Quality in Design, Aug 2016.

V. Nagaraju, L. Fiondella, P. Zeephongsekul, and T. Wandji, An Adaptive EM Algorithm for the Maximum Likelihood Estimation of Non-homogeneous Poisson Process Software Reliability Growth Models. Submitted to *Reliability Engineering and System Safety*.

P. Zeephongsekul, C. Jayasinghe, L. Fiondella, and V. Nagaraju, Maximum Likelihood Estimation of Parameters of NHPP Software Reliability Models Using EM and ECM Algorithms, Submitted to *IEEE Transactions on Reliability*.

V. Nagaraju, T. Wandji, and L. Fiondella, Maximum Likelihood Estimation of a Non-homogeneous Poisson Process Software Reliability Model with the Expectation Conditional Maximization Algorithm, *Conference on Applied Statistics in Defense*, Fairfax, VA, October 2015.

V. Nagaraju and L. Fiondella, Introduction to Mathematical Software Reliability Models, In Proc. of *Annual Reliability and Maintainability Symposium*, Tucson, AZ, Jan 2016.

APPENDIX B: CITED AND RELATED REFERENCES

Musa, J. D. (2004). Software reliability engineering: more reliable software, faster and cheaper. Tata McGraw-Hill Education.

Pham, H. (1999). Software Reliability. John Wiley & Sons, Inc.

Lyu, M. R. (1996). Handbook of Software Reliability Engineering. CA: IEEE Computer Society Press. Available for download at http://www.cse.cuhk.edu.hk/~lyu/book/reliability/ (verified 2/28/2016)

Musa, J. D., Iannino, A., & Okumoto, K. (1987). Software Reliability: Measurement, Prediction, Application. McGraw-Hill, Inc.Agarwal, S., & Dagli, C. H. (2013).