Software Reliability Growth Models: Systematic Descriptions and Implementations

I. Corro Ramos[1]

E. Brandt[2], A. Di Bucchianico[1] and R. Henzen[2]

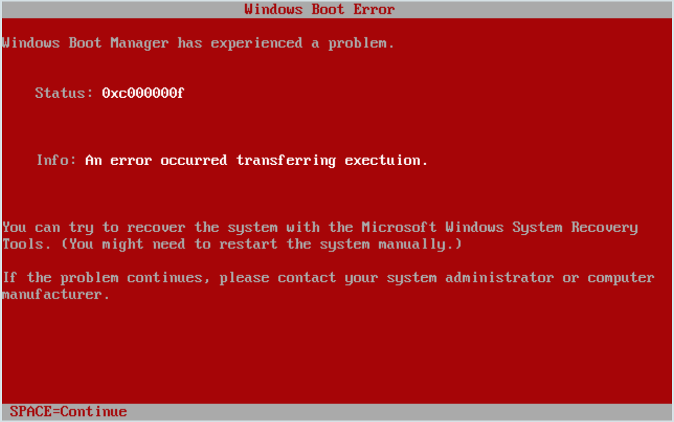
ENBIS7 Dortmund, Germany, September 26th, 2007

[1] Eindhoven University of Technology, [2] Refis System Reliability Engineering

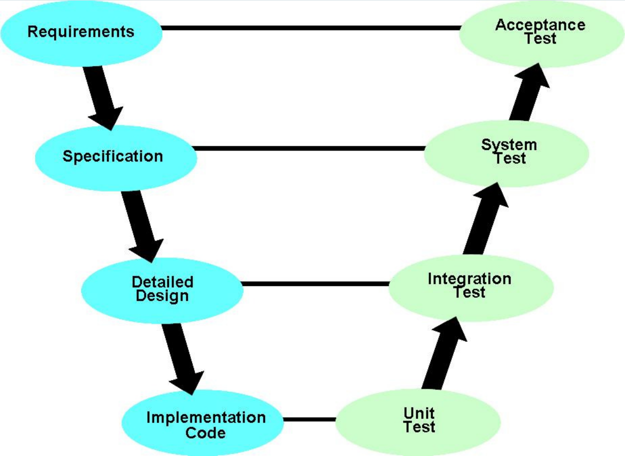
Outline

- Introduction
- Systematic description of software reliability growth models
- Systematic implementations: software reliability tool
- Conclusions
- Questions

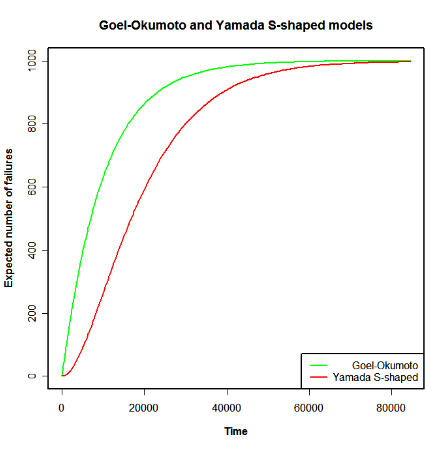
Red Screen of Death



Software testing



Reliability growth data



Software reliability growth models (SRGM)

Statistics is needed exhaustive testing not always feasible

SRGM used to:

- Resource planning
- Release decisions
- Certification
- Find one or two more (QE paper)

Typical statistical quantities of interest:

- Probability of no failure in a given time period
- Expected time until next failure
- Number of errors left

Problems

- Insufficient documentation of the models
- More than 200 models known
- Out-dated algorithms

What is needed:

- Systematic description of SRGM (general methodology + algorithms)
- Support pre-selection of models
- Apply specific state-of-the-art algorithms for the models (convergence issues)
- Tool to perform analyses

Systematic description of software reliability growth models

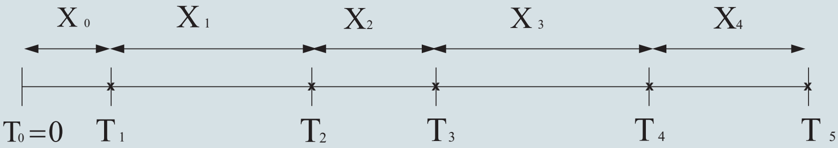
1. Probability model
2. Trend analysis
3. Model estimation
4. Model validation
5. Prediction

Systematic description of software reliability growth models (Cont.)

1. Probability model
   1.1 Joint distribution of $\{N(t)\}_{t>0}$ and/or $T_1, T_2, \ldots, T_n$ and/or $X_1, X_2, \ldots, X_n$ (including (in)dependence structure)
   1.2 Model assumptions
   1.3 Interpretation of model parameters
2. Trend analysis
   2.1 $\ldots$

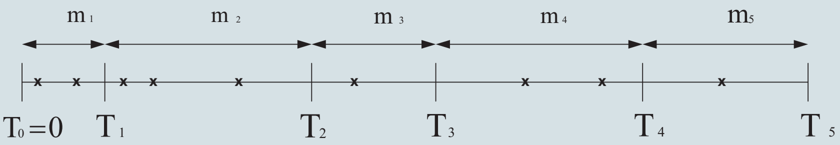
Data type (I)

Ungrouped or exact data



Data type (II)

Grouped or interval count data



SRGM description

- **GOS** models: $T_1, T_2, \ldots, T_n$ order statistics of a sample $Z_1, Z_2, \ldots, Z_N$ with c.d.f. F_θ.
  - Jelinski-Moranda or EOS: $X_i \sim Exp(\lambda(N - i + 1))$ i.i.d.
- **NHPP** models: $N(t_i) - N(t_{i-1}) \sim Poisson(\lambda(t))$ independent with mean

$$\mathbb{P}\left[N(t_i) - N(t_{i-1}) = k\right] = e^{-(\Lambda(t_i)-\Lambda(t_{i-1}))} \frac{(\Lambda(t_i) - \Lambda(t_{i-1}))^k}{k}$$

  - Goel-Okumoto: $\Lambda(t) = a(1 - e^{-bt})$
  - Yamada S-shaped: $\Lambda(t) = a(1 - (1 + bt)e^{-bt})$
  - Duane: $\Lambda(t) = \left(\frac{t}{\alpha}\right)^\beta$

Systematic description of software reliability growth models (Cont.)

3. Parameter estimation
   - 3.1 Point estimation procedures for parameters (Maximum Likelihood and/or Least Squares)
     - 3.1.1 Data requirements
     - 3.1.2 Existence results for parameter estimates
     - 3.1.3 Performance of parameter estimators (bias, efficiency)
     - 3.1.4 Algorithms to compute parameter estimates
   - 3.2 Confidence interval procedures for parameters (Maximum Likelihood and/or Least Squares)
     - 3.2.1 Distributional description of underlying estimators
     - 3.2.2 Algorithms to compute confidence intervals

Model estimation

- The unknown number of errors N is also a parameter (standard asymptotic theory is not applicable)
- ML equations usually do not have closed-form solution (use specific algorithm for each model, see Knafl and Morgan for some NHPP models)
- Direct maximization may cause numerical problems (parameters of different order of magnitude (Yin, Trivedi paper))

Model (pre-)selection

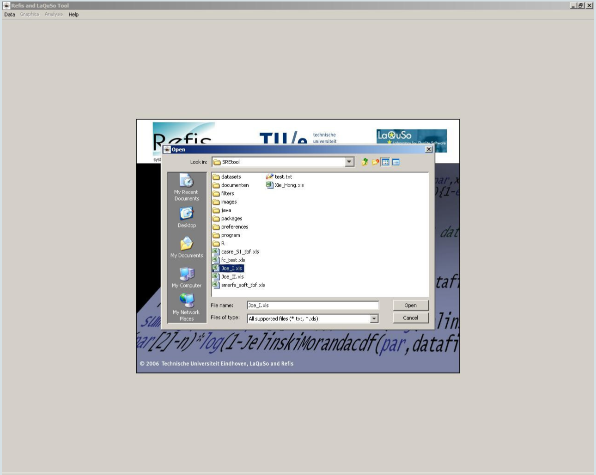
Data Requirements and Assumptions	Relative Importance	Geometric	Jelinski-Moranda	Littlewood-Verrall	Musa basic	Musa-Okumoto	Goel-Okumoto	Shick-Wolverton	Schneidewind	Yamada S-shaped	Duane
Data may be exact failure times (ungrouped data)	2	x	x	x	x	x	x	x	x	x	x
Data may be grouped failure times (interval count data)	2	x	x	x	x	x	x	x	x	x	x
Testing intervals may be of different length	3	x	x	x	x	x	x	x		x	x
Failures need not occur equally likely	2	x		x		x	x		x	x	x
Detection of faults may be dependent of each other	2			x	x	x	x		x	x	x
Failures need not be of the same severity	1			x	x	x	x		x	x	x
Detection rate depends on time (testing effort)	3			x	x	x	x		x	x	x
Detection rate depends on number of remaining defects	3	x	x					x			
Failures need not be repaired instantaneously	3		x		x		x	x	x	x	x
Imperfect repair of defects allowed	2										
Infinite number of errors allowed	2					x					x

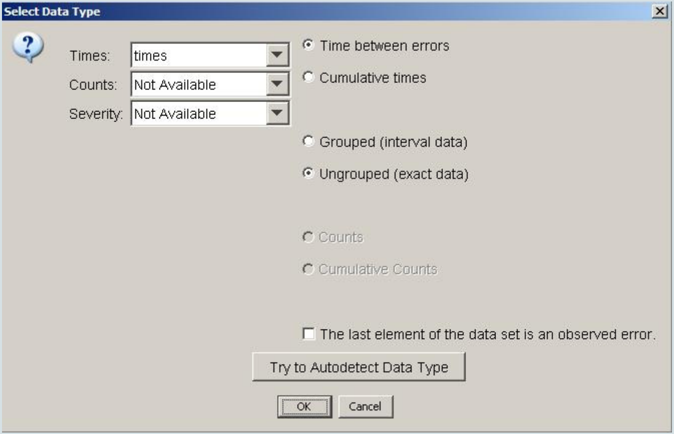
Systematic implementations: software reliability tool

- Existing packages do not make full use of state-of-the art statistical methodology
- Interface in Java (platform independent)
- Statistical computations in R (open-source free software)
- Communication: JRI and JavaGD libraries (Computational Statistics group, University of Augsburg)
- Financially supported by a grant of the Dutch Innovation Platform

Example


Refis and LaQuSo Tool
Data Graphics Analysis Help
Refis
system reliability engineering
TU/e technische universiteit eindhoven
LaQuSo
Laboratory for Quality Software
SRE Tool
© 2006 Technische Universiteit Eindhoven, LaQuSo and Refis

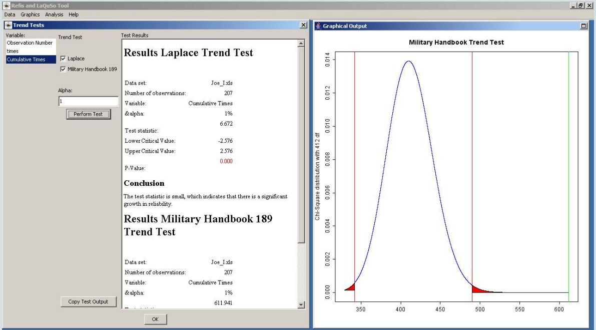

Refis and LaQuSo Tool
Data Graphics Analysis Help
Open
Look in: SREtool
My Recent Documents
Desktop
My Documents
My Computer
My Network Places
datasets
documenten
filters
images
java
packages
preferences
program
R
casre_S1_tbf.xls
fc_test.xls
Joe_I.xls
Joe_II.xls
smerfs_soft_tbf.xls
test.txt
Xie_Hong.xls
File name: Joe_I.xls
Files of type: All supported files (*.txt, *.xls)
Open
Cancel
© 2006 Technische Universiteit Eindhoven, LaQuSo and Refis

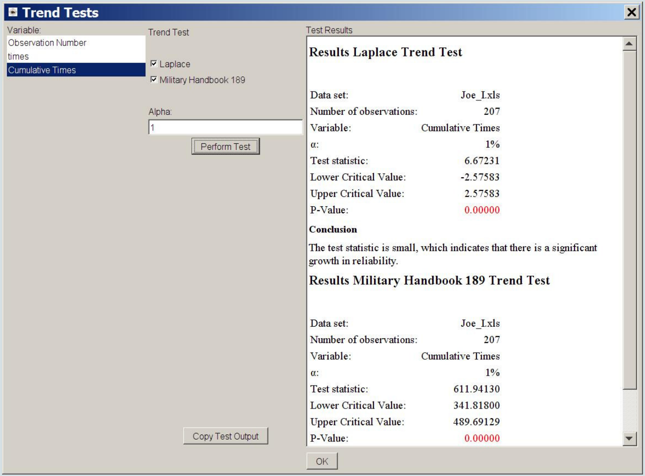

Select Data Type
Times: times
Counts: Not Available
Severity: Not Available
Time between errors
Cumulative times
Grouped (interval data)
Ungrouped (exact data)
Counts
Cumulative Counts
The last element of the data set is an observed error.
Try to Autodetect Data Type
OK
Cancel

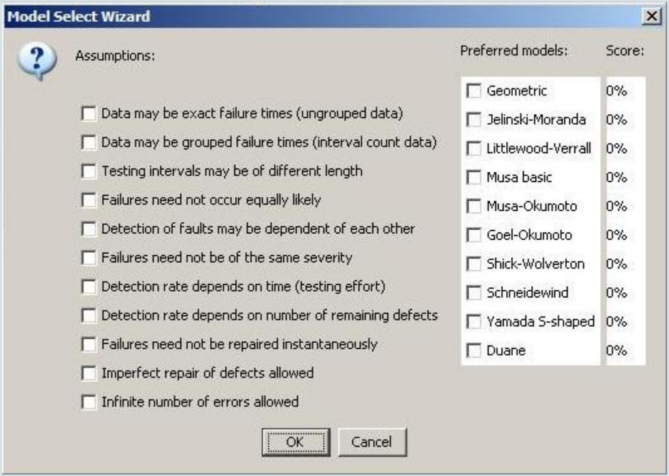
Joe_I.xls

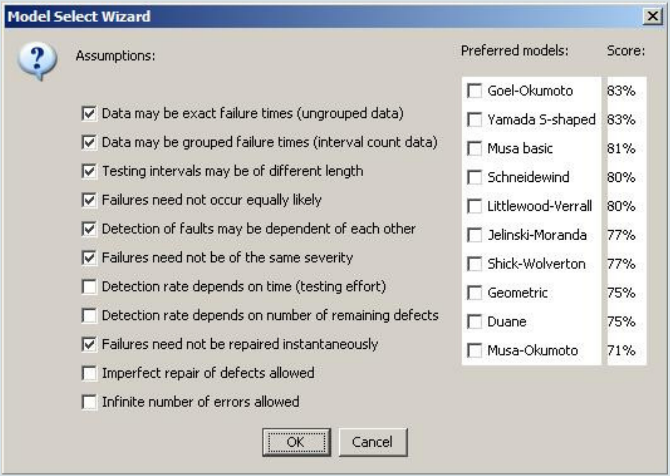
	Observation Number	times
1	1	39
2	2	10
3	3	4
4	4	36
5	5	4
6	6	5
7	7	4
8	8	91
9	9	49
10	10	1
11	11	25
12	12	1
13	13	4
14	14	30
15	15	42
16	16	9
17	17	49
18	18	44
19	19	32
20	20	3
21	21	78
22	22	1
23	23	30
24	24	205
25	25	5

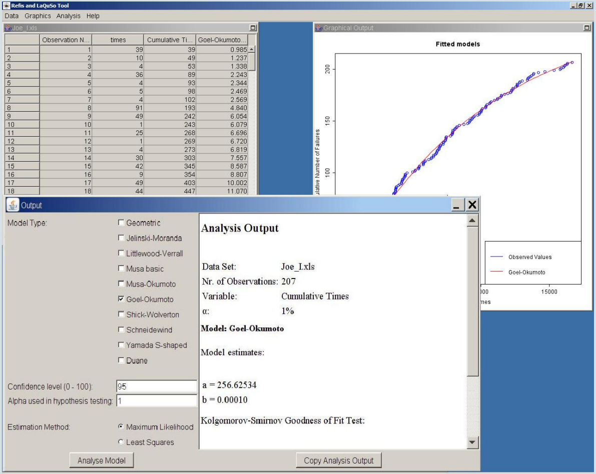
Data | Filtered Data | Model Analysis Results

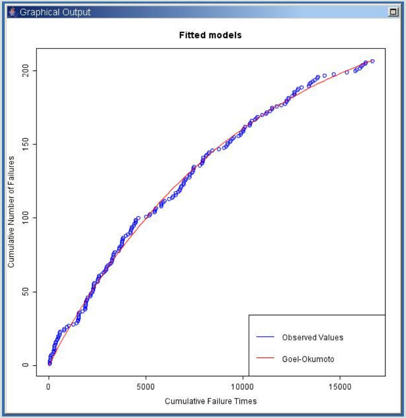

Refis and LaQuSo Tool
Data Graphics Analysis Help
Trend Tests
Variable:
Observation Number
times
Cumulative Times
Trend Test
Laplace
Military Handbook 189
Alpha:
1
Perform Test
Copy Test Output
Test Results
Results Laplace Trend Test
Data set: Joe_I.xls
Number of observations: 207
Variable: Cumulative Times
&alpha: 1%
6.672
Test statistic:
Lower Critical Value: -2.576
Upper Critical Value: 2.576
0.000
P-Value:
Conclusion
The test statistic is small, which indicates that there is a significant growth in reliability.
Results Military Handbook 189 Trend Test
Data set: Joe_I.xls
Number of observations: 207
Variable: Cumulative Times
&alpha: 1%
611.941
OK
Graphical Output
Military Handbook Trend Test
Chi-Square distribution with 412 df
0.000
0.002
0.004
0.006
0.008
0.010
0.012
0.014
350
400
450
500
550
600

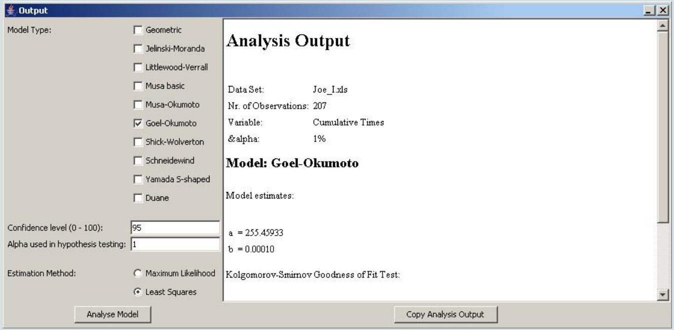

Trend Tests
Variable:
Observation Number
times
Cumulative Times
Trend Test
Laplace
Military Handbook 189
Alpha:
1
Perform Test
Copy Test Output
Test Results
Results Laplace Trend Test
Data set: Joe_I.xls
Number of observations: 207
Variable: Cumulative Times
α: 1%
Test statistic: 6.67231
Lower Critical Value: -2.57583
Upper Critical Value: 2.57583
P-Value: 0.00000
Conclusion
The test statistic is small, which indicates that there is a significant growth in reliability.
Results Military Handbook 189 Trend Test
Data set: Joe_I.xls
Number of observations: 207
Variable: Cumulative Times
α: 1%
Test statistic: 611.94130
Lower Critical Value: 341.81800
Upper Critical Value: 489.69129
P-Value: 0.00000
OK


Model Select Wizard
Assumptions:
Data may be exact failure times (ungrouped data)
Data may be grouped failure times (interval count data)
Testing intervals may be of different length
Failures need not occur equally likely
Detection of faults may be dependent of each other
Failures need not be of the same severity
Detection rate depends on time (testing effort)
Detection rate depends on number of remaining defects
Failures need not be repaired instantaneously
Imperfect repair of defects allowed
Infinite number of errors allowed
Preferred models:
Score:
Geometric 0%
Jelinski-Moranda 0%
Littlewood-Verrall 0%
Musa basic 0%
Musa-Okumoto 0%
Goel-Okumoto 0%
Shick-Wolverton 0%
Schneidewind 0%
Yamada S-shaped 0%
Duane 0%
OK
Cancel


Model Select Wizard
Assumptions:
Data may be exact failure times (ungrouped data)
Data may be grouped failure times (interval count data)
Testing intervals may be of different length
Failures need not occur equally likely
Detection of faults may be dependent of each other
Failures need not be of the same severity
Detection rate depends on time (testing effort)
Detection rate depends on number of remaining defects
Failures need not be repaired instantaneously
Imperfect repair of defects allowed
Infinite number of errors allowed
Preferred models:
Score:
Goel-Okumoto 83%
Yamada S-shaped 83%
Musa basic 81%
Schneidewind 80%
Littlewood-Verrall 80%
Jelinski-Moranda 77%
Shick-Wolverton 77%
Geometric 75%
Duane 75%
Musa-Okumoto 71%
OK
Cancel


Refis and LaQuSo Tool
Data Graphics Analysis Help
Joe_I.xls
Observation N... times Cumulative Ti... Goel-Okumoto...
1 1 39 39 0.985
2 2 10 49 1.237
3 3 4 53 1.338
4 4 36 89 2.243
5 5 4 93 2.344
6 6 5 98 2.469
7 7 4 102 2.569
8 8 91 193 4.840
9 9 49 242 6.054
10 10 1 243 6.079
11 11 25 268 6.696
12 12 1 269 6.720
13 13 4 273 6.819
14 14 30 303 7.557
15 15 42 345 8.587
16 16 9 354 8.807
17 17 49 403 10.002
18 18 44 447 11.070
Graphical Output
Fitted models
Observed Values
Goel-Okumoto
15000
Output
Model Type:
Geometric
Jelinski-Moranda
Littlewood-Verrall
Musa basic
Musa-Okumoto
Goel-Okumoto
Shick-Wolverton
Schneidewind
Yamada S-shaped
Duane
Confidence level (0 - 100): 95
Alpha used in hypothesis testing: 1
Estimation Method:
Maximum Likelihood
Least Squares
Analyse Model
Analysis Output
Data Set: Joe_I.xls
Nr. of Observations: 207
Variable: Cumulative Times
α: 1%
Model: Goel-Okumoto
Model estimates:
a = 256.62534
b = 0.00010
Kolgomorov-Smirnov Goodness of Fit Test:
Copy Analysis Output


Graphical Output
Fitted models
Cumulative Number of Failures
0
50
100
150
200
0
5000
10000
15000
Cumulative Failure Times
Observed Values
Goel-Okumoto


Output
Model Type:
Geometric
Jelinski-Moranda
Littlewood-Verrall
Musa basic
Musa-Okumoto
Goel-Okumoto
Shick-Wolverton
Schneidewind
Yamada S-shaped
Duane
Confidence level (0 - 100):
95
Alpha used in hypothesis testing:
1
Estimation Method:
Maximum Likelihood
Least Squares
Analyse Model
Analysis Output
Data Set: Joe_I.xls
Nr. of Observations: 207
Variable: Cumulative Times
&alpha: 1%
Model: Goel-Okumoto
Model estimates:
a = 255.45933
b = 0.00010
Kolgomorov-Smirnov Goodness of Fit Test:
Copy Analysis Output

Conclusions

- Systematic description of software reliability growth models needed
- Model assumptions
- Numerical problems in model estimation
- Results for both grouped and ungrouped data
- Software reliability tool
- Programmed in Java (interface) and R (statistical computations)

Questions

Advanced statistical issues

- Model validation
  - Standard goodness-of-fit test result inappropriate in general
  - Alternatively: conditional goodness-of-fit tests
  - Unconditional goodness-of-fit tests for diverse subclasses of models, only for ungrouped data!
- Model prediction
  - One-sided confidence intervals
  - Asymptotic confidence intervals ($N \to \infty$ or $t \to \infty$)
  - Number of remaining errors, reliability (time to next error) of the system, predicted intensity function, etc ...