Berichte aus der Wirtschaftsinformatik

**Michael Grottke**

# Modeling Software Failures during Systematic Testing

## The Influence of Environmental Factors

D 29 (Diss. Universität Erlangen-Nürnberg)

Shaker Verlag
Aachen 2003

**Bibliographic information published by Die Deutsche Bibliothek**
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the internet at http://dnb.ddb.de.

Zugl.: Erlangen-Nürnberg, Univ., Diss., 2003

**n2**



Printed in Germany.

ISBN 3-8322-2035-6
ISSN 1438-8081

Shaker Verlag GmbH • P.O. BOX 101818 • D-52018 Aachen
Phone: 0049/2407/9596-0 • Telefax: 0049/2407/9596-9
Internet: www.shaker.de • eMail: info@shaker.de

# Acknowledgments

This work is based on the research results I attained while I was employed at the University of Erlangen-Nuremberg for working on the project PETS (Prediction of software Error rates based on Test and Software maturity results), partially funded by the IST (information society technologies) program of the European Union.

I am grateful to Professor Dr. Ingo Klein of the Chair of Statistics and Econometrics for letting me devote my time to this tightly-scheduled project; otherwise, meeting the milestones would not have been possible. He also gave me the academic freedom to focus on those approaches and techniques that I deemed most promising.

Furthermore, my thanks go to Professor Dr. Michael Amberg of the Chair of Information Systems III for accepting the task of second reviewer of this dissertation.

The European Union did not only finance my position as a research assistant, but it also enabled the acquisition of the latest literature as well as my participation at various conferences on software quality around the world. Through its support preliminary results of the PETS project could be discussed with international researchers.

I am indebted to the employees of the industrial partners Acron Communications (Prague), AS Aprote (Tartu), imbus AG (Möhrendorf) and Procedimientos-Uno, SL (Málaga) for their willingness to endure my questions, my presentations and the additional burden of data collection. Their input ensured the practical relevance of the research carried out.

Moreover, I would like to extend my thanks to my colleagues at the Chair of Statistics and Econometrics, especially for their understanding concerning my intermediate position between research and practice, or between project work and affiliation with the Chair.

Special appreciation is offered to PD Dr. Susanne Rässler of the Chair of Statistics and Econometrics and to Dipl.-Inf. Thomas Rossner of imbus AG for their valuable comments on chapters of my manuscript.

Finally and most importantly, I wish to thank my parents and my brother for their encouragement over the past few years. Without their support this work could not have been written.

Michael Grottke

*September 2003*

# Contents

# List of figures

# List of tables

# Frequently used notation

## Functions and random variables

| | |
|---|---|
| $\Gamma(x)$ | gamma function |
| $\Delta^n f(x)$ | $n^{th}$ forward difference of $f(x)$ |
| $\exp(x)$ | exponential function |
| $E(X)$ | expected value of random variable $X$ |
| $F(x)$ | distribution function |
| $I_A(x)$ | indicator function for $x \in A$ |
| $\ln(x)$ | natural logarithm of $x$ |
| $\mathcal{L}$ | likelihood function |
| $P(H_1)$ | probability of the event $H_1$ |
| $P(H_1 \mid H_2)$ | probability of the event $H_1$ conditional on the event $H_2$ |
| $P_{[m]}$ | probability of the occurrence of exactly $m$ events |
| $S_m$ | sum of the probabilities of the simultaneous occurrence of at least $m$ events |

## Goodness of fit measures and other metrics

| | |
|---|---|
| $\text{AIC}_j$ | Akaike's information criterion calculated based on the first $j$ observations |
| $\alpha$ | Cronbach's $\alpha$ |
| $C$ | condition number |
| $d_{(p)}$ | $p$-quantile of the observations of $d$ |
| $d_{(0.5)}$ | 0.5-quantile of the observations of $d$ (= median of $d$) |
| $\kappa$ | Cohen's $\kappa$ |
| $\lambda_{\min}, \lambda_{\max}$ | smallest and largest eigenvalue of a matrix |
| $R^2$ | coefficient of determination |
| $R^2_{adj}$ | adjusted coefficient of determination |
| $R^2_{AN}$ | Aldrich's and Nelson's pseudo $R^2$ measure |
| $R^2_{MF}$ | McFadden's pseudo $R^2$ measure |
| $R^2_{MZ}$ | McKelvey's and Zavoina's pseudo $R^2$ measure |
| $\rho$ | empirical correlation coefficient |

## Miscellaneous

| | |
|---|---|
| $\mathbb{N}$ | integer numbers |
| $\mathbb{N}_0$ | integer numbers including zero |
| $\mathbb{R}_0^+$ | positive real numbers including zero |
| $\lfloor x \rfloor$ | largest integer less than or equal to $x$ |
| $\binom{n}{k}$ | binomial coefficient $= \frac{n!}{k!(n-k)!}$ |
| $\infty$ | infinity |
| $\propto$ | proportional to |

## Abbreviations

| | |
|---|---|
| 4 GL | fourth generation programming language |
| CAF | CMM Appraisal Framework |
| CASE | computer aided software engineering |
| CBA IPI | CMM-Based Appraisal Framework for Internal Process Improvement |
| CM | configuration management |
| CMM | Capability Maturity Model |
| CPU | central processing unit |
| CUS | customer-supplier process category |
| DDIF | development difficulty |
| DEFF | development effort |
| DEPI | development effort performance index |
| DMSK | development team manager's skill level |
| DRPI | development runtime performance index |
| DRUN | development runtime |
| DTSI | size of the development team |
| ENG | engineering process category |
| FDEN | fault density |
| IEC | International Electrotechnical Commission |
| ISO | International Organization for Standardization |
| LS-Cum | Least squares estimation based on the cumulative number of failure occurrences |
| LS-Delta | Least squares estimation based on the number of failure occurrences per test case |
| MAN | management process category |
| MARE | mean absolute relative error |
| MIS | management information system |

| | |
|---|---|
| ML-NHPP | Maximum likelihood estimation based on the likelihood following from interpreting the model as a non-homogenous Poisson process model |
| ML-SetupC | Maximum likelihood estimation based on the likelihood implied by the model setup |
| ORG | organization process category |
| PAUT | proportion of automated test cases |
| PETS | Prediction of software Error rates based on Test and Software maturity results |
| PGSK | programmers' general skill level |
| PNDT | proportion of new members in the development team |
| PNTT | proportion of new members in the test team |
| PRCH | proportion of requirements changed after the specification phase |
| PRCO | proportion of reused code |
| PREJ | proportion of rejected failure messages |
| PSSK | programmers' specific skill level |
| PTED | proportion of testers with a special education as test engineers |
| QA | quality assurance |
| SARE | short term absolute relative error |
| SCAP | selective capability rating |
| SCE | Software Capability Evaluation |
| SEI | Software Engineering Institute |
| SICC | size of the compiled code |
| SPICE | Software Process Improvement and Capability dEtermination |
| SSE | error sum of squares |
| SST | total sum of squares |
| SUP | support process category |
| SW-CMM | Capability Maturity Model for Software |
| TCAP | testing capability rating |
| TDIF | testing difficulty |
| TEFF | testing effort |
| TEPI | testing effort performance index |
| TGSK | testers' general skill level |
| TMSK | test team manager's skill level |
| TRPI | testing runtime performance index |
| TRUN | testing runtime |
| TSSK | testers' specific skill level |
| TTSI | size of the test team |

## Software reliability models / software failure models

| | |
|---|---|
| $c(\tilde{t})$ | deterministic relative code coverage function |
| $E$ | code construct state "eliminated" |
| $EC$ | code construct state "eliminated and correct" |
| $EF$ | code construct state "eliminated and faulty" |
| $g(\tilde{t})$ | testing efficiency function |
| $g_{\tilde{i}}$ | testing efficiency at the $\tilde{i}^{th}$ stage of testing |
| $G$ | total number of code constructs |
| $G_{A,i}$ | number of code constructs that are located in state $A$ after execution of the $i^{th}$ test case |
| $G_{A,p,i}$ | number of code constructs that are located in state $A$ before the $i^{th}$ test case execution and that are exercised by this test case |
| $G_{A \to B,i}$ | number of code constructs residing in state $A$ before the $i^{th}$ test case execution and in state $B$ afterwards |
| $G_{TF \to T,p,i}$ | number of already tested, faulty code constructs that are tested and replaced during the $i^{th}$ test case |
| $G_{TF \to TF,p,i}$ | number of already tested, faulty code constructs that are tested and replaced without activating the fault during the $i^{th}$ test case |
| $i$ | number of test cases executed |
| $i_j$ | test case at which the $j^{th}$ measurement was taken |
| $i_t$ | total number of test cases in the test plan |
| $\tilde{i}$ | generic discrete measure of testing progress |
| $K$ | fault exposure ratio |
| $\kappa(\tilde{t})$ | expected relative code coverage function |
| $\lambda(\tilde{t})$ | failure intensity function |
| $\lambda_{\tilde{i}}$ | failure intensity at the $\tilde{i}^{th}$ stage of testing |
| $m_i$ | number of failures experienced / faults detected during the first $i$ test cases |
| $\Delta m_i$ | number of failures experienced / faults detected during the $i^{th}$ test case |
| $\Delta m_j^*$ | number of failures experienced / faults detected during the $j^{th}$ observation period, i.e. between the $(i_{j-1}+1)^{th}$ and the $i_j^{th}$ test case |
| $M(\tilde{t})$ | random variable denoting the cumulative number of failures experienced by time $\tilde{t}$ |
| $M_{\tilde{i}}$ | random variable denoting the cumulative number of failures experienced during the first $\tilde{i}$ stages of testing |
| $\mu(\tilde{t})$ | mean value function, $E(M(\tilde{t}))$ |
| $N$ | expected number of inherent faults |
| $\nu_d$ | expected number of inherent detectable faults |

| | |
|---|---|
| $p$ | number of code constructs executed per test case |
| $q_i$ | number of code constructs exercised during the first $i$ test cases |
| $\Delta q_i$ | number of code constructs exercised during the $i^{th}$ test case |
| $\Delta q_j^*$ | number of code constructs exercised during the $j^{th}$ observation period, i.e. between the $(i_{j-1}+1)^{th}$ and the $i_j^{th}$ test case |
| $Q(\tilde{t})$ | random variable denoting the cumulative number of code constructs exercised by time $\tilde{t}$ |
| $Q_{\tilde{i}}$ | random variable denoting the cumulative number of code constructs exercised during the first $\tilde{i}$ stages of testing |
| $r$ | redundancy level |
| $s$ | fault activation probability |
| $t$ | testing effort |
| $t^*$ | calendar time |
| $\tilde{t}$ | generic continuous measure of testing progress |
| $T$ | code construct state "tested" |
| $TC$ | code construct state "tested and correct" |
| $TF$ | code construct state "tested and faulty" |
| $\tau$ | CPU execution time |
| $u_0$ | number of inherent faults |
| $U$ | code construct state "untested" |
| $UC$ | code construct state "untested and correct" |
| $UF$ | code construct state "untested and faulty" |
| $w(t^*)$ | instantaneous testing effort at calendar time $t^*$ |
| $W(t^*)$ | cumulative testing effort until calendar time $t^*$ |
| $X_{I,i}$ | random variable denoting the number of faulty code constructs exercised at least once during the first $i$ test cases |
| $\Delta X_{I,i}$ | random variable denoting the number of faulty code constructs exercised for the first time by the $i^{th}$ test case |
| $\Delta X_{II,i}$ | random variable denoting the number of faulty code constructs exercised by the $i^{th}$ test case |
| $\Xi_i$ | random variable denoting the number of faulty constructs either corrected or eliminated during the first $i$ test cases |
| $z_i$ | probability with which a certain previously not eliminated code construct is executed by the $i^{th}$ test case |
| $z_a(\tilde{t})$ | per-fault hazard rate at time $\tilde{t}$ |
| $z_{a,\tilde{i}}$ | per-fault detection rate at the $\tilde{i}^{th}$ stage |
| $z(\Delta\tilde{t} \mid \tilde{t}_{n-1})$ | hazard rate of the application after the $(n-1)^{th}$ failure occurrence |