The Poverty of Reliability Growth Models

Timm Grams

University of Applied Sciences

Marquardstr. 35, D-36039 Fulda, Germany

Timm.Grams@et.fh-fulda.de

http://www.fh-fulda.de/~fd9006/index.htm

Introduction. Seemingly the field of reliability growth modelling faces the same basic, i. e. philosophical, problems as the so called historicism. According to Popper long-term prophecies cannot be applied to social systems we are all caught in. And "it is wise to combat the most urgent and real social evils one by one, here and now" (Popper, 1963).

The following criticism is located on a methodological level. We should be aware of the applicability of the methodological rules to be chosen. Thomas S. Kuhn and Paul Feyerabend ("Anything goes") are two well known opponents to Popper's approach. But even Feyerabend, the proponent of an anarchistic methodology, admits that there may "come a time when it will be necessary to give reason a temporary advantage and when it will be wise to defend its rules to the exclusion of everything else" (1975).

The science of reliability growth models shows some of the features of what Thomas S. Kuhn (1962) calls "normal science": "Though science surely grows in depth, it may not grow in breadth as well". (p. 170). These are the phases of "normal science", when "the research worker is a solver of puzzles, not a tester of paradigms" (p. 144). "Scientific training" - within the context of normal science - "is not well designed to produce the man who will easily discover a fresh approach" (p. 166).

This short paper reports on my efforts of assessing reliability growth models in respect to their prognostic strength and their ability of supporting the *learning from errors* under the light of Popper's criteria.

Reliability growth models under the light of Popper´s Criteria. Reliability growth models (RGM) are meant to predict the future behaviour of software on the basis of past experience. The predictions cannot be corroborated by means of experiments.

Knowledge is *synthetic* (or *empirical*) insofar as certain assumptions are involved which are not valid *a priori* and which are to be corroborated through experience. It is *analytic* in those parts which are solely based on logical deductions and mathematics. The criteria Popper (1980) puts on the synthetic or empirical content of prediction methods - the predictive strength of the methods - are:

Falsifiability: A prediction method of any value cannot be true under all circumstances. There must be the possibility of failure. A statement like "It will rain or not rain here tomorrow" are not regarded as empirical, simply because it cannot be refuted.

Corroboration: A (falsifiable) prediction method is said to be corroborated if its predictive value has been demonstrated under many different conditions.

Objectivity: The objectivity of predictions and propositions lies in the fact that they can be inter-subjectively tested.

Simplicity: A prediction method may not depend on too many adaptable parameters. Because otherwise its elusion of falsification is too easy and its predictive value is too small.

With respect to objectivity the RGMs are not worse than other prediction methods. But in my opinion all the other of Popper's criteria are missed by RGMs.

- Reliability growth models cannot be falsified in the sense of Popper's *criterion of falsifiability.*

- This is mainly due to the fact that there is an abundance of models and parameters such that it is nearly impossible of not finding a model fitting with some given experimental or field data.

- Thus the models don't fulfil the *criterion of simplicity*. In this context "simple" doesn't mean "easily understandable". It is meant in the sense that the fitting of a *simple* straight line with some points in the plane is more convincing and possesses more empirical power than the fact that the points may be approximated by some higher order curve (not simple).

- Consequently these models cannot be *corroborated* (in Popper's sense). In all demonstrations of the models I know so far, the selection of a model and its fitting to the data is done in retrospect. I don't know any falsifiable and non-trivial *prediction scheme* for software reliability on the basis of these models.

RGM and the Trivial Reliability Prediction (TRP) under the light of statistical criteria. There is a method known from the weather forecast, the "trivial forecast": Tomorrow the weather will be the same as today. This trivial forecast has some predictive accuracy.

I wondered whether the trivial reliability growth model would compare well with the highly sophisticated models under the light of the following statistical criteria: 1. the *u-Plot* for detecting consistent bias, and 2. the *Prequential Likelihood Ratio* for detecting warranted or unwarranted noisiness of the prediction (see Bev Littlewood, 1990). The experiments showed that the trivial model compared well with the sophisticated reliability growth models. For details refer to my web page http://www.fh-fulda.de/ ~fd9006/RGM/RGMcriticized.html.

The preceding paragraph suggests that the TRP may possibly have some empirical content: The model is simple and, because its goal is so modest, could possibly be corroborated for software with a lot of small faults, i. e. for *unreliable software in a later phase of debugging*.

Conclusions. RGMs make strong statements for weak and basically unaltered software, or - at least - stable production processes. But the engineer should ask for strong software, and in the case of weak software he is forced to make revolutionary changes in the software production process. Such changes make the reliability growth models completely worthless.

We should not give up the principles of if-then-sciences and the principles of empiricism. That is because we want to learn something about our work and we want to predict the behaviour of our artefacts.

The engineers effort seems to be badly allocated in the calculation of RGMs. He should better be involved in activities aiming at better software. These are the *methods of fault-intolerance.* (Fault tolerance does not belong to the engineer's virtues.)

References

Feyerabend, P.: Against Method. Verso, London 1975

Kuhn, T. S.: The structure of scientific revolutions. The University of Chicago Press, Chicago 1962

Littlewood, B.: Modelling Growth in Software Reliability. In: Software Reliability Handbook. Elsevier, London, New York 1990 (Edt: Paul Rook), pp. 137-153, 401-412

Popper, K. R.: Prediction and Prophecy in Social Sciences. In: Conjectures and Refutations. The Growth of Scientific Knowledge. Routledge, London 1963

Popper, K. R.: The Logic of Scientific Discovery. Unwin Hyman, London 1980