

Projektvorhersage - die nichtlineare Zukunft

Markus Manleitner

Product Release
Dräger medical
Moislinger Allee 53-55
23542 Lübeck
Markus.Manleitner@draeger.com

1 Problem

In jedem Projekt gibt es Kennzahlen. Oft entsteht der Wunsch aus den jüngsten Trends eine Prognose für die Zukunft abzuleiten und sehr schnell wird durch die letzten Werte eine Gerade gelegt und weit in die Zukunft verlängert. Doch die Zukunft ist nichtlinear und deshalb führen viele leichtfertig erstellte Vorhersagen mit Regressionsgraden in die Irre. Zum einen, weil die Kennzahl gar nicht geeignet ist, zum anderen weil nicht alle Kennzahlen linear verteilt sind. Neben der Güte des Modells ist für den Anwender die Güte der Vorhersage viel entscheidender, aber kennen Sie diese Werte für Ihre Prognosen?

2. Modellierung

In rund 20 Projekten wurden die Fehlerfinderaten (Abb.1 links) und die kumulierten Summen pro Monat (Abb.1 rechts) aufgetragen, sowie eine Auswahl von möglichen Modellfunktionen, die den s-förmigen Verlauf der Daten unterstützen.

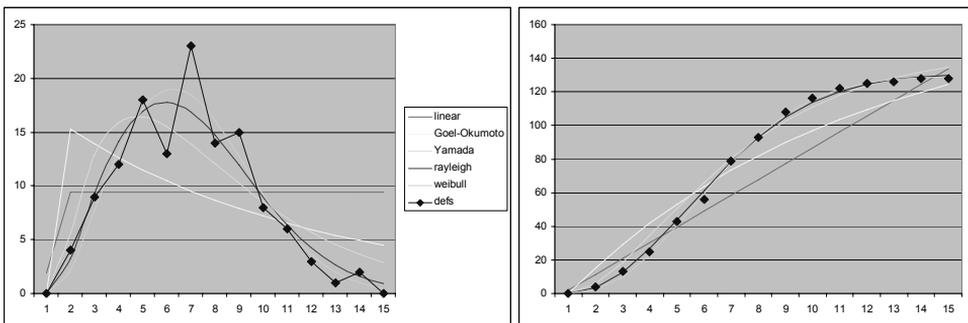


Abbildung 1 Fehler pro Monat und kumulierte Werte

Die Weibullverteilung erreicht durch drei Formparameter die beste Anpassung an die realen Datenpunkte (siehe Abb.2), dieser zusätzliche Freiheitsgrad reduziert jedoch in gleichem Maß die Güte der Vorhersage bei unvollständigen Datenreihen. Da der Formfaktor ohnehin bei allen Projekten nahe 2.0 lag haben wir uns daher für die Modellierung mit der Rayleigh-Verteilung entschieden und mit nur zwei Freiheitsraden eine gute Modellierung erreicht.

	linear		Goel-Okumoto		Yamada		Rayleigh		Weibull
corr	0,9648	corr	0,9837	corr	0,9953	corr	0,9992	corr	0,9996
error	4290	error	2462	error	430	error	56	error	41

Abbildung 2 Güte der Modelle

Die horizontale Skalierung der Modellfunktion erfolgt über $a=1/(2 \cdot T_{max} \cdot T_{max})$, dem Zeitpunkt, an dem die Verteilungsfunktion $v(t) = 2adte^{-at^2}$ ihr absolutes Maximum erreicht. Die Güte der Abschätzung erfolgt über den Pearson-Korrelationsfaktor. Er verändert sich nicht mit der vertikalen Skalierung, womit eine Unabhängigkeit der beiden Parameter erreicht wird. Die vertikale Skalierung erfolgt über den Parameter d, der dem Maximalwert für V(t) für t -> unendlich entspricht. Die Güte der Abschätzung wird über die Summe der Fehlerquadrate ermittelt.

Für den Verwender ist weniger die Güte der Modellierung von Interesse, als die Güte der Vorhersage. Nach dem Wendepunkt können Genauigkeiten zwischen 0,1 und 5% für eine 6-Monatsprognose erreicht werden. Der Trend der beiden Formparameter a und d stellt ein wichtiges Indiz für die Güte der Vorhersage dar. Verlässliche Prognosen können also nur aus einem kontinuierlichen Metrikprogramm heraus erstellt und beurteilt werden.

3. Ausblick

Die S-Kurven lassen sich neben der Vorhersage auch für eine Reifegradbestimmung des Produktes verwenden. Durch eine Normierung der Achsen lassen sich Projekte miteinander vergleichen, wenn die Prozesse dafür stabil und gleichartig genug sind.

Rayleigh-Verteilungen lassen sich auch auf Projektbudgets anwenden. [DL02] Die Anwendung erstreckt sich auf Plausibilitätsprüfung von Budgetplanungen, Vorhersage von zu erwartenden Ausgaben und Budgetüberschreitungen.

Literaturverzeichnis

- [Kan02] Stephen H. Kan: Metrics and models in software quality engineering
- [Lyu96] Michael R. Lyu: Handbook of software reliability engineering
- [DL02] David Lee: Norden-Rayleigh Analysis: A useful tool for EVM in development projects