

GJB5000A四级军用软件组织过程绩效基线和模型的建立

叶赫镠, 费玮莹, 程春姬, 马 晋

(中国航空无线电电子研究所, 上海 200233)

摘要: 实施GJB5000A四级定量项目管理能大大提升软件的质量, 而实施定量项目管理需要组织过程绩效提供过程绩效数据、基线和模型。本文从数据采集、数据预处理、基线建立、模型建立等几个方面描述了建立GJB5000A四级软件组织过程绩效基线和模型的方法, 并用统计学的假设检验方法验证了绩效基线和模型的正确性。

关键词: GJB5000A-2008; 组织过程绩效; 基线和模型

中图分类号: TP311.5 **文献标识码:** A

The Building of GJB5000A ML4 Organizational Process Performance Baseline and Model for Military Software

YE Chengliu, FEI Weiyang, CHENG Chunji, MA Jin

(China National Aeronautical Radio Electronics Research Institute, Shanghai 200233, China)

Abstract: The implementation of GJB5000A ML4 quantitative project management can greatly improve the quality of software, for which organizational process performance is needed to provide process performance data, baselines and models. This paper introduces the building method of GJB5000A ML4 software organizational process performance baselines and models, and verifies the baselines and models by hypothesis testing.

Keywords: GJB5000A-2008; organizational process performance; baselines and models

1 引言(Introduction)

近年来, 我国武器装备正处于快速发展时期, 军用软件的需求日趋增多, 软件的规模、应用范围、功能和性能都飞速增长, 软件逐步成为武器装备的灵魂与核心。例如, 一个飞机的航电系统非常复杂, 由多个系统构成, 例如通信系统、显示系统、飞控系统等等, 并且综合到航电系统的功能越来越复杂, 任务更加多样化, 系统高度综合化, 航电软件已成为大规模甚至超大规模的复杂软件系统。同时, 航电软件的交付质量、周期、成本对飞机研制影响越来越大, 也越来越受到重视。

采用科学的方法对软件研制过程进行有效管理, 是实现提升质量和风险成本控制的必然途径。总装备部先后出台了GJB5000-2003《军用软件能力成熟度模型》和GJB5000A-2008《军用软件研制能力成熟度模型》, 对软件研制组织开展软件研制能力的评价。2013年, 总装备部颁布了GJB8000-2013《军用软件研制能力等级要求》, 对军用软件研制单位应达到的军用软件研制能力等级提出了明确的要求。

在GJB5000A军用软件研制能力成熟度等级4中, 有两个过程域^[1]: 组织过程绩效和定量项目管理。一个组织整体运作

的关键因素在于组织过程绩效^[2]，建立和维护组织过程绩效基线和模型是非常关键的，是提升组织能力和进行定量管理的基础。

2 GJB5000A组织过程绩效概述(An overview of GJB5000A organizational process performance)

过程绩效是遵循某个过程所达到的实际结果的测量，包括过程测量和产品测量。组织过程绩效的目的，在于建立并维护对组织的标准过程绩效的定量了解，并提供过程绩效数据、基线和模型，以用于定量管理组织项目^[1]。

组织过程绩效收集以往项目的历史数据，选择过程，确定哪些过程适于测量，然后建立过程绩效测量项，确定哪些测量项有助于判定过程绩效，之后建立质量和过程绩效目标，确定所选择过程的质量和过程绩效目标。最后组织利用各种统计技术，建立和维护组织过程绩效基线和模型。

过程绩效基线是某一单个度量指标的预期表现或某些度量指标集合的预期表现。基线用于比较实际过程绩效和预期过程绩效的基准，为项目管理和实施活动提供参考点。

过程绩效模型是对过程属性和过程工作产品之间关系的描述，基于历史过程绩效数据而建立，并使用项目中收集的过程测量和产品测量进行校准，最终用于预测遵循过程将能达到的结果。

3 建立军用软件组织过程绩效基线和模型(Building organizational process performance baseline and model)

建立组织过程绩效的前提是组织已积累了大量的以往项目的历史数据。本文假设组织的软件质量和过程绩效目标已确定。为了建立软件组织过程绩效基线和模型，本文从数据采集、数据预处理、基线建立、模型建立等几个方面进行描述。

3.1 数据采集

由于组织过程绩效基线和模型是基于组织的大量项目历史数据，所以首先需要收集大量软件项目历史数据。数据收集范围按照过程绩效测量项的要求，通常涵盖软件项目的各阶段数据，包括了规模、工作量、质量、进度、风险、测试、生产率、过程活动的有效性和人力资源等。

数据采集途径可以通过项目测量与分析记录表、项目管理工具，软件工程组的其他数据记录或组织资产库等。

数据采集的原则：正确性、精确性、真实性、连贯性、完整性。

3.2 数据预处理

对于采集到的项目历史数据，需要进行预处理和有效性分析，确保数据的准确性和完整性。

对于某些数据，需要进行合并、拆分、去重、分组、转换、派生、建立子集、缺失值处理等，通常可通过一些软件工具或编写一些简单程序完成，如果历史数据量少，也可手工完成。

处理缺失值通常包含以下步骤^[3]：

- (1)识别缺失数据；
- (2)检查导致数据缺失的原因；
- (3)删除包含缺失值的实例或用合理的数值代替(插补)缺失值。

采集的项目历史数据，还需要判断其中是否存在异常点，这可以通过统计图形“箱线图”(图1)确认。箱线图可以粗略地看出数据是否具有对称性、中心位置、分布范围和异常点等信息。如果有异常点，需要进行原因分析。结合实际项目情况，这些异常数据要么进行修正，要么被剔除。如果是特殊原因引起的，该项目数据可删除。

图1中，项目数据中有两个超过图中须线上下限的观测值，这些点有异常的嫌疑。分析发现，此两个项目和其他数据项目不是同一软件类型，需要归类到另一软件类型中分析。因此，该软件类型数据组中应删除此两个项目数据。

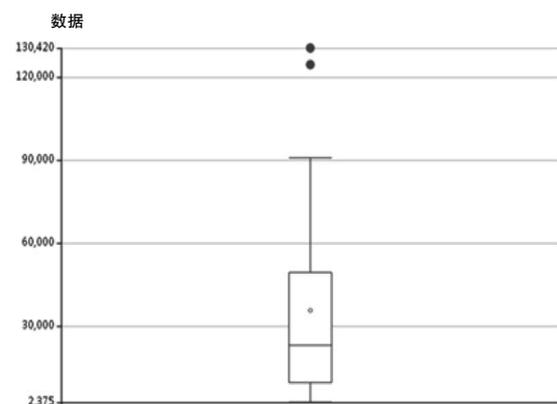


图1 箱线图

Fig.1 BoxPlot

3.3 组织过程绩效基线的建立

基线代表了量化的过程能力，为每个过程元素创建一个

单独的基线，可以为项目提供一些通常的指导。基线数据的适宜性与产品的应用领域、裁剪过程和组合顺序、级别、团队经验等与产品相关的属性有关，组织可以依据不同属性，利用统计方法建立多种基线。

3.3.1 随机变量及其分布

用来表示随机现象结果的变量称为随机变量，分为离散型随机变量和连续型随机变量。

离散型随机变量的常见分布有二项分布、泊松分布(Poisson Distribution)、超几何分布，其他还有0—1分布、负二项分布、几何分布、整数均匀分布、任意离散分布等。

连续型随机变量的常见分布有正态分布、均匀分布、对数正态分布、指数分布、Weibull分布，其他还有三角分布、Beta分布、Gamma分布、Laplace分布、logistic分布、对数Logistic分布、最大极值分布、最小极值分布等。

质量管理中最常遇到的连续分布是正态分布。如果某项指标受到很多项随机因素的干扰，而每项干扰都很小的话，则所有干扰影响的综合结果将导致此项指标的分布为正态分布^[4]。

当我们不能确定某个实数该用哪种形式的分布时，采用正态分布是默认的比较好的选择，有两个原因。第一，我们想要建模的很多分布真实情况比较接近正态分布。中心极限定理(Central Limit Theorem)说明很多独立随机变量的和近似服从正态分布。这意味着在实际中，很多复杂系统都可以被成功建模成正态分布的噪声，即使系统可以被分解成具有更多结构化行为的各个部分。第二，在具有相同方差的所有可能的概率分布中，正态分布具有最高的熵(平均信息量)。

3.3.2 了解数据分布

利用统计图形“直方图”可以了解数据的位置状况、离散程度和分布情况，并且可与要求的分布进行比较。直方图将样本值划分为许多称为区间的间隔。条形表示落于每个区间内的观测值的数量(频率)。通常使用包含拟合概率密度函数(pdf)的直方图,如图2所示。图2中，计算出该组样本数据的正态分布参数均值和标准差，然后画出了概率密度函数(pdf)拟合分布线，由图中看出，数据比较拟合正态分布。

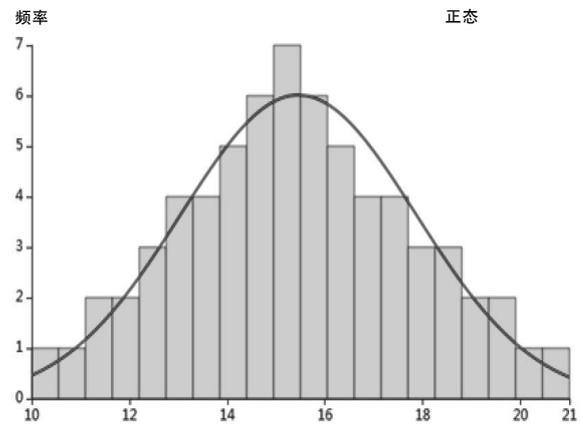


图2 直方图

Fig.2 Histogram

3.3.3 数据正态性检验

使用统计图形“概率图”有助于确定特定分布是否拟合数据，或者比较不同的样本分布。

概率图使用最小二乘估计法对参数进行估计，计算累积分布函数(cdf)和关联的置信区间。概率图使用了AD假设检验来验证数据是否服从指定分布。

概率图中，如果分布与您的数据拟合，则标绘点将大致形成一条直线，标绘点将近似于拟合分布线，AD统计量会很小，而关联的p值将大于所选的 α 显著性水平(通常为0.05或0.10)。

概率图如图3所示，中心的红色直线是拟合分布线，弯曲的蓝线内是拟合分布的大约95%的置信区间。图3同时计算出AD值为0.472,P值为0.2226。从概率图中看，数据基本在95%置信区间以内，且P值大于 α 显著性水平0.05，数据可接受服从正态分布的原假设。概率图同时计算出该组数据的正态分布参数均值和标准差。

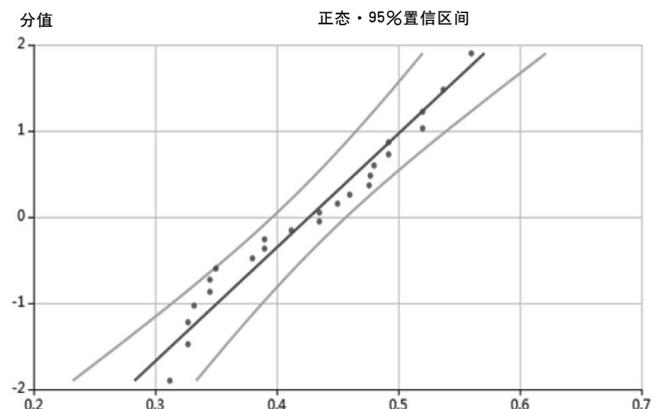


图3 概率图

Fig.3 Probability plot

3.3.4 数据稳定性分析

通过以上方法，计算出了数据的分布参数，现在需要分析数据的稳定性，以发现并消除变差的特殊原因。

在统计过程控制(SPC)中，变差的来源分为普通原因和特殊原因。随着时间的推移，普通原因会产生一个稳定的且可重复的分布，称为“处于统计上受控制的状态”。如果一个过程只存在变差的普通原因且不改变时，该过程的输出是可以预测的。

如果存在变差的特殊原因，随着时间的推移，过程的输出将不稳定，这种不稳定的过程的表现是不可预测的。

区分过程的普通原因和特殊原因的简单有力的工具是控制图^[5]。控制图跟踪一段时间内的过程统计量并检测是否存在特殊原因。控制图结构如图4所示。控制图计算该组数据的平均值和标准差 σ ，默认情况下在统计量的平均值处绘制中心线CL，在中心线以上 3σ 处绘制控制上限UCL，在中心线以下 3σ 处绘制控制下限LCL。

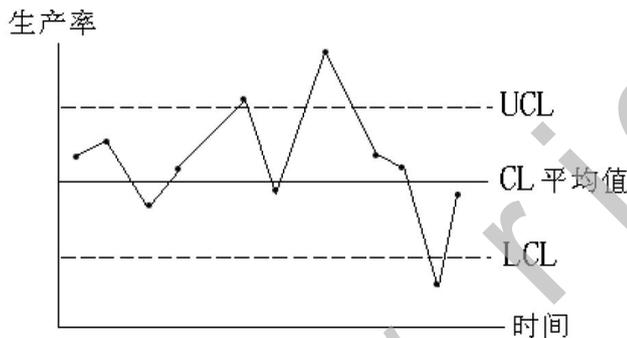


图4 控制图结构

Fig.4 The structure of control chart

在SPC中，基线(即过程能力)定义为平均值的上下 3σ 的区间范围。通常，大约68.26%的数据落在平均值两侧各一倍 σ 区间范围内，大约95.44%的数据落在平均值两侧各 2σ 区间范围内，大约99.73%的数据落在平均值两侧各 3σ 区间范围内。

所以，不管分布如何，使用 3σ 作为控制限都可以保证非常少的误报，即使基础的分布明显地不是正态分布；同时，落到 3σ 控制限外的数据点极有可能是特殊原因造成的。因此， 3σ 控制限提供了一个稳健的标准，对所有的基础分布都能有效地工作。

在控制图中，需要识别特殊原因，有几种识别特殊原因

的准则，详见表1。确定使用哪个或几个准则组合取决于被研究/控制的过程。

表1 典型特殊原因识别准则

Tab.1 Typical special causes identification criteria

典型特殊原因识别准则	K值
1个点距离中心线大于K个标准差	3
连续K点在中心线同一侧	9
连续K个点，全部递增或全部递减	6
连续K个点交替上下变化	14
2/3个点，距离中心线(同侧)大于2个标准差	
4/5个点，距离中心线(同侧)大于1个标准差	
连续K个点，距离中心线(任一侧)1个标准差以内	15
连续K个点，距离中心线(任一侧)大于1个标准差	8

有一个或多个数据点落在UCL和LCL的控制限之外，需要根据实际情况判定是否出现特殊原因，确是特殊原因者可以删除该组数据，未发现异常者必须保留此组数据且增大观测数据。

控制图的类型很多，常用的控制图按数据类型分为计量型和计数型控制图，计量型主要针对连续数据(如工作量、进度、规模等)。计数型主要是针对离散型数据(如缺陷类型、项目复杂度、软件类型、项目类型等)。根据度量指标特性，应选择合适的控制图，各类控制图的选择如图5所示。

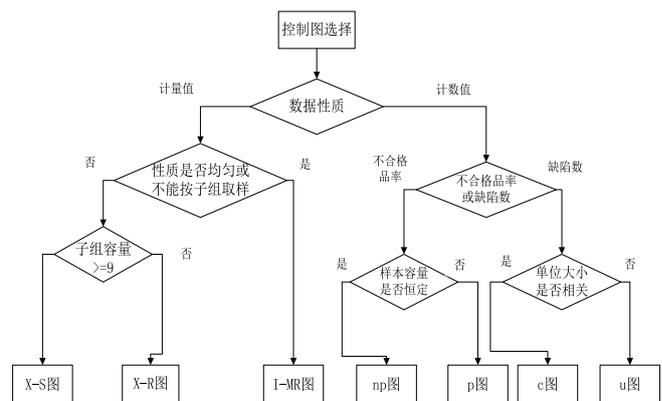


图5 控制图选择

Fig.5 Selection of control charts

I-MR控制图的I图如图6所示，MR图如图7所示。从图中可看出，所有数据都落在UCL和LCL的控制限之内，未出现特殊原因，该过程稳定可控。

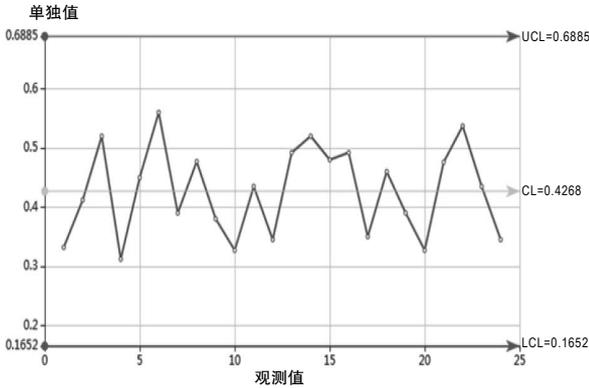


图6 I-MR控制图的I图

Fig.6 I chart of I-MR control chart

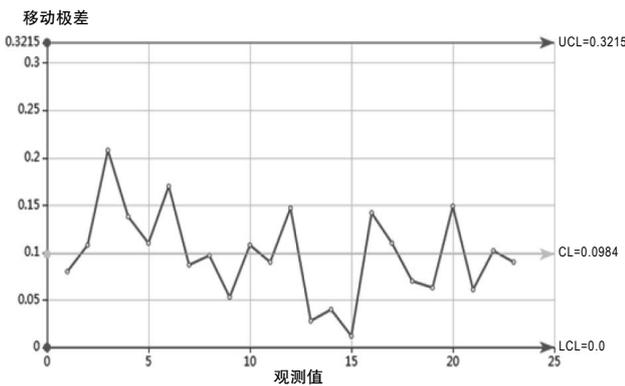


图7 I-MR控制图的MR图

Fig.7 MR chart of I-MR control chart

3.3.5 基线的获取

如果所有数据点都落在控制图的UCL和LCL的控制限之内,说明该控制限是合理的。可以把该度量指标的基线数据值定义为(LCL~CL~UCL)。

当运用控制图时,不断地有大部分数据出现在控制限外,说明该过程是不稳定的、失控的。对一个失控的过程不应该建立基线,首先要做的是判定失控的原因,采取直接、迅速的措施来稳定过程,其次才考虑进行数据采集和建立基线。

如果数据量较多,标准差σ较大时,应考虑数据分组或分层建立基线。

3.4 组织过程绩效模型的建立

3.4.1 因变量(Y)和自变量(X)相关性分析

基于年度软件过程质量和绩效目标(质量、工作量、进度等),可确定因变量(Y)和自变量(X)。确定了某过程绩效测量项作为因变量(Y),可采用散点图和相关性分析,找出过程中影响因变量(Y)的有关数据因子,即因变量(Y)对应的自变

量(X)。

散点图用来图示说明两个变量之间的关系,把每一对(x_i, y_i)看成直角坐标系中的某个点,在图中标出n个点。散点图如图8所示。

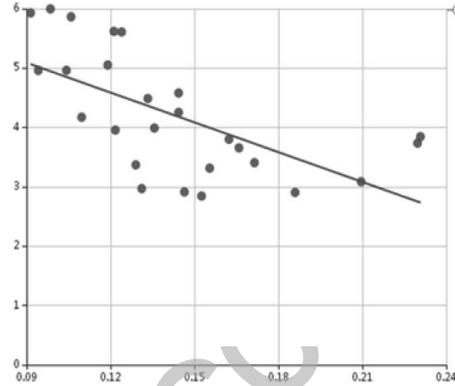


图8 散点图

Fig.8 Scatterplot

如果散点图呈现如图8的形状,即N个点基本在一条直线附近但又不完全在这条直线上,可以用相关系数r这个统计量来衡量其密切程度。

相关系数是协方差除以每个变量的标准差的乘积,公式为:

$$r = \frac{\text{Cov}(X, Y)}{\sigma_x \sigma_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{L_{xy}}{\sqrt{L_{xx}L_{yy}}}$$

相关系数特点是:

(1)当r=±1时, n个点完全在一条直线上,这时称两个变量完全线性相关。

(2)当r=0时, 称两个变量不相关,这时散布图上n个点可能毫无规律,不过也可能两个变量间存在非线性相关。

(3)当r>0时,称这两个变量正相关,此时x增大, y也增大。

(4)当r<0时, 称这两个变量负相关,此时x增大, y减小。

(5)|r|愈大, 线性相关就愈强; 接近1表示存在很强的线性关系。

图8中相关系数r为-0.6295,说明两者间存在负相关, X轴变量是Y轴因变量的一个自变量(X)。

因变量(Y)有多个自变量(X), 应分析自变量(X)间的相关性。如果自变量(X)间有相关性, 则不能存在同一个数学模型公式中。

3.4.2 选择适用的统计分析技术

统计分析技术有很多种,例如回归分析、变异数分析、方差分析(ANOVA)、t-检验、配对检验、卡方检验、Logit分对数等。因变量(Y)和自变量(X)分为连续型和离散型数据,应根据因变量(Y)和自变量(X)不同的类型组合,选择不同的分析技术,详见表2。

表2 分析技术选择

Tab.2 Selection of analytical techniques

自变量X \ 因变量Y	连续数据	离散数据
	连续数据	线性回归
离散数据	方差分析、T检验、配对T检验	卡放检验

例如如果因变量(Y)和自变量(X)都是连续数据,可采用线性回归分析建立相应的模型。回归分析用来分析因变量(Y)与一个或多个自变量(X)之间关系并为之建模,分为简单回归、多元回归、Logistic回归和泊松回归等。普通最小二乘法(OLS)^[3]回归是最常见的统计分析方法,在OLS回归中,估计方程可通过确定将样本的数据点与由方程预测的值之间的距离平方和最小化的方程计算得出。

回归方程拟合出来后需要进行显著性检验^[3],从总体上判定回归方程是否有效。检验方法有相关系数、方差分析、F检验、T检验等。当回归方程效果显著时,进行各个回归系数的显著性检验,判断回归方程中哪些自变量(X)是显著的。将效果不显著的自变量(X)删除,以优化模型。

如果在软件项目中,系统测试缺陷密度和系统测试单位规模用例数、系统测试单位规模工作量、系统测试人员平均技能系数有关。使用多元线性回归分析方法,得到系统测试缺陷密度的回归模型方程为:系统测试缺陷密度 $=-0.0264+0.0502 \times$ 系统测试人员平均技能系数 $+0.202 \times$ 系统测试单位规模用例数 $+0.0265 \times$ 系统测试单位规模工作量。用ANOVA方差来分析回归方程的显著性检验,F值(服从包含项自由度和误自由度的F分布)较大,P值 <0.05 ,说明回归方程总效果是显著的。各个自变量的P值均小于0.05,说明每个

自变量都有显著意义。 $R-Sq=94.5\%$, $R-Sq(\text{调整})=93.8\%$,说明模型拟合,可接受此模型。将来,如果组织积累了更多的项目历史数据,可以基于更多的数据重新建立基线和模型。

后续软件项目,可结合建立的组织过程绩效基线和模型,来进行定量项目管理。

4 结论(Conclusion)

GJB5000A军用软件研制能力成熟度模型等级4中,组织过程绩效基线和模型可以帮助组织科学地确定项目的目标,进行定量项目管理,通过预测确定达到目标的确性概率,确保目标的最终实现。

本文结合实际例子,采用统计过程控制技术和统计算法描述了建立GJB5000A四级组织过程绩效基线和模型的方法,并描述了用假设检验方法验证绩效基线和模型的方法。本文对于建立军用软件组织过程绩效具有很好的参考价值。

参考文献(References)

- [1] 石柱.军用软件研制能力成熟度模型及其应用[M].北京:中国标准出版社,2009:205-206.
- [2] 刘文红,马贤颖,董锐,等.基于CMMI的软件工程实施:高级指南[M].北京:清华大学出版社,2015:365-366.
- [3] 方匡南,朱建平,姜叶飞.R数据分析:方法与案例详解[M/CD].北京:电子工业出版社,2015:87-89;206-217.
- [4] 薛毅,陈立萍.统计建模与R软件[M].北京:清华大学出版社,2007:31-32.
- [5] 蒋家东,冯允成.统计过程控制[M].北京:中国质检出版社,2011.

作者简介:

叶赫镠(1972-),男,本科,工程师.研究领域:软件工程.
 费玮莹(1970-),女,本科,工程师.研究领域:软件工程.
 程春姬(1974-),女,硕士,高级工程师.研究领域:软件工程.
 马晋(1979-),男,硕士,高级工程师.研究领域:机载嵌入式软件,软件工程.