

小样本飞机生产质量偏差数据分析与预测方法研究

王陆航, 张冬冬, 卢鹤, 李汝鹏, 葛小丽

引用本文

王陆航, 张冬冬, 卢鹤, 李汝鹏, 葛小丽. [小样本飞机生产质量偏差数据分析与预测方法研究](#)[J]. 计算机科学, 2024, 51(11A): 240300123-8.

WANG Luhang, ZHANG Dongdong, LU Hu, LI Rupeng, GE Xiaoli. [Study on Analysis and Prediction Method of Small Sample Aircraft Production Quality Deviation Data](#) [J]. Computer Science, 2024, 51(11A): 240300123-8.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于GM-LSTM模型的南京市老年人口预测研究](#)

Research on Elderly Population Prediction Based on GM-LSTM Model in Nanjing City
计算机科学, 2021, 48(6A): 231-234. <https://doi.org/10.11896/jsjx.200900142>

[基于聚类与特征融合的蛋白质亚细胞定位预测](#)

Prediction of Protein Subcellular Localization Based on Clustering and Feature Fusion
计算机科学, 2021, 48(3): 206-213. <https://doi.org/10.11896/jsjx.200200081>

[基于行为特征分析的微博恶意用户识别](#)

Microblogging Malicious User Identification Based on Behavior Characteristic Analysis
计算机科学, 2018, 45(12): 111-116. <https://doi.org/10.11896/j.issn.1002-137X.2018.12.017>

[基于主成分分析法的人工免疫识别软件缺陷预测模型研究](#)

Research on Software Defect Prediction Based on AIRS Using PCA
计算机科学, 2017, 44(Z6): 483-485. <https://doi.org/10.11896/j.issn.1002-137X.2017.6A.107>

[本体语义相似度自适应综合加权算法研究](#)

Adaptive Ontology Semantic Similarity Comprehensive Weighted Algorithm
计算机科学, 2016, 43(10): 242-247. <https://doi.org/10.11896/j.issn.1002-137X.2016.10.046>

小样本飞机生产质量偏差数据分析与预测方法研究

王陆航¹ 张冬冬² 卢 鹄³ 李汝鹏³ 葛小丽³

1 同济大学数学科学学院 上海 200092

2 同济大学电子与信息工程学院 上海 201804

3 上海飞机制造有限公司航空制造技术研究所 上海 201324

(2050032@tongji.edu.cn)

摘 要 随着现代工业水平和对飞机精度要求的不断提升,对飞机生产质量的分析和管控方法越来越受到各大航空企业的重视。当前阶段,针对飞机装配偏差存在可参考样本数据少、不确定性大、非线性、多层级装配等固有特征,传统的分析方法难以准确地构建飞机生产偏差分析模型。因此,以飞机生产过程的偏差为研究目标,对飞机生产质量偏差数据分析与预测方法展开系统研究。首先分析各个零件之间的偏差关系,基于主成分分析法识别对总偏差影响最大的关键零件,找到重点预测的目标;接着从实际生产的类正态数据出发,重点关注关键零件,实现了基于正态云模型的偏差数据预测、生成与验证,得到更多样本的飞机生产质量偏差数据及其隶属度,一定程度上缓解了“小样本”的问题,并基于 k -折交叉验证对预测模型进行了评估;最后构建了基于改进的灰色预测模型的多源数据融合的装配偏差波动区间协同预测模型,“小样本”问题的缓解使得区间预测更加精细、科学,在公差数据的参考下,预测飞机生产质量偏差所在的区间范围,为实际生产和制定公差修正机制提供指导。

关键词: 小样本数据;飞机生产质量管控;数学建模;主成分分析法;正态云模型; k -折交叉验证;灰色预测模型

中图分类号 TP399

Study on Analysis and Prediction Method of Small Sample Aircraft Production Quality Deviation Data

WANG Luhang¹, ZHANG Dongdong², LU Hu³, LI Rupeng³ and GE Xiaoli³

1 School of Mathematical Sciences, Tongji University, Shanghai 200092, China

2 College of Electronic and Information Engineering, Tongji University, Shanghai 201804, China

3 Aviation Manufacturing Technology Research Institute, Shanghai Aircraft Manufacturing Co., Ltd., Shanghai 201324, China

Abstract With the development of modern industrial capabilities and the demand for increased precision in aircraft, the analysis and control of aircraft production quality have become a focal point for major aerospace enterprises. At the current stage, traditional analytical methods face challenges in accurately constructing deviation analysis models for aircraft production due to inherent features such as limited reference sample data, significant uncertainties, non-linearity, and multi-level assembly deviations. Therefore, this paper focuses on the deviations in the aircraft production process and systematically explores methods for analyzing and predicting aircraft production quality discrepancies. Firstly, this paper analyzes the deviation relationships among various components, identifying the key component with the greatest impact on total deviation based on principal component analysis, which pinpoints the focus for predictive targeting. Subsequently, starting from actual production data resembling a normal distribution, this paper pays special attention to key components, enabling the prediction, generation, and validation of deviation data based on the normal cloud model, which yields aircraft production quality deviation data and their memberships for a greater variety of samples, alleviating the issue of “small sample” and evaluating the predictive model through k -fold cross-validation. Finally, a cooperative predictive model for assembly deviation fluctuation intervals based on the improved grey forecasting model and multi-source data fusion is established. The alleviation of the “small sample” issue enhances the precision and scientific nature of interval prediction. With reference to tolerance data, predict the interval range where aircraft production quality deviations are located, providing guidance for actual production and the formulation of tolerance correction mechanisms.

Keywords Small sample data, Aircraft production quality control, Mathematical modeling, Principal component analysis, Normal cloud model, K -fold cross validation, Grey forecasting model

基金项目:国家重点研发计划(2021YFB3301901)

This work was supported by the National Key R&D Program of China(2021YFB3301901).

通信作者:张冬冬(ddzhang@tongji.edu.cn)

1 引言

在当前竞争高度激烈的航天制造环境中,航天制造公司关注的重点是如何以更高的效率制造技术更强、品质更好的飞机产品^[1]。而提升飞机产品品质的关键在于对生产零件的质量进行有效管控。事实上,在飞机制造生产过程中,尺寸偏差通常导致约70%的质量问题^[2]。零件制造偏差、工装设备的稳定性、环境等多种因素都会影响飞机结构件的形位公差,最终影响飞机尺寸精度^[3]。因此,在复杂的飞机生产实际当中,建立飞机生产质量偏差分析与预测模型显得十分必要。

飞机的实际装配过程非常复杂,每个关键控制点通过多种零件组装完成,各零件之间表现出多层次、强耦合的特性。因此,需要对各零件数据进行分析,揭示各个数据维度之间的关系,找到引起偏差的关键零件。

在飞机装配公差设计阶段,需要进行装配精度预测,以及及时发现不合理的公差设计并进行调整,保证装配质量。飞机装配偏差预测建模过程存在干扰多、采样样本少、数据不完备等特征;此外,不同数据链含有的信息量及其重要性也不同。因此,需要构建基于小样本的飞机装配偏差预测和生成模型,为飞机装配公差设计提供科学的参考,对提升飞机装配质量具有重要意义。

鉴于上述情况,本文以可参考样本数据少、不确定性大、非线性、多层级装配的飞机装配偏差为主要研究对象,构建基于数学建模方法的小样本飞机生产质量偏差数据分析与预测模型。引入主成分分析法和加权主成分分析法,对较高的样本维度进行降维,识别引起偏差的关键零件类型,为生产提供关注重点;引入正态云模型,模拟和预测更多的仿真数据,缓解“小样本”问题;引入灰色预测模型,预测飞机生产偏差所在区间,为生产提供指导。通过充分利用和挖掘检测数据的潜在信息,实现基于数据挖掘的复杂飞机结构件装配偏差的优化设计。本文研究有助于缩短飞机研制周期,提高装配效率,为飞机制造质量管控提供技术支持。

2 相关工作

飞机制造领域的质量管控一直以来都是极为关键的,任何制造缺陷都可能对飞机的安全性和性能产生重大影响。而对生产质量偏差数据的分析与预测,为未来生产提供指导,是质量管控的重要环节。国内外先进的航空制造公司以及相关学者在飞机生产质量偏差分析、预测方面进行了大量研究。与此同时,汽车生产领域的许多学术研究也对飞机生产质量偏差分析、预测具有一定的指导意义。

2.1 飞机质量偏差研究相关工作

随着航空工业的不断发展和飞行技术的日新月异,飞机质量偏差研究成为航空领域备受关注的课题。在追求更高的安全性、效率和性能的背景下,科研人员不断致力于深入了解飞机质量偏差的起因、影响以及可能的解决方案。飞机质量偏差主要基于过往的历史数据,存在过程复杂、难以建模等问题。因此,为了确保飞机在运行过程中的安全可靠,飞机质量偏差研究成为一项迫切的任务,许多学者对此展开了研究。

Cheng^[4]主要研究数字化装配技术,通过挖掘零件之间

的空间几何关系,建立了以变形为主导的单个站位的飞机数字化装配偏差模型,实现对飞机数字化装配偏差的预测。Sun^[5]以飞机壁板件结构为研究对象,建立了壁板件装配偏差传递模型,并基于模型对壁板件的装配偏差进行预测。Zhao^[6]将运动装配偏差源分为零件几何位置偏差、几何形状偏差和装配位置偏差3类,基于几何关系分析和极值法得到运动装配精度预测值。Yang^[7]研究飞机中的广义预测控制理论,针对复杂的飞机非线性模型实现了预测和控制。Hugues等^[8]构建了一个故障和维护场景模型,在飞机设计阶段预测飞机的操作可靠性,并通过年度统计样本对预测概率进行验证。

许多与飞机相关的研究需要用到软件仿真或有限元仿真等方法来应对飞机质量偏差数据普遍存在的“小样本”问题。由于软件仿真要求工程实践方面的基础知识,而有限元仿真则需要大量的计算资源和时间,因此一些学者也针对“小样本”做了研究。

Wang^[9]根据现场实际使用通用飞机的情况,收集通用飞机起落架结构减摆器的故障数据,针对故障数据样本少的特点,假设了故障数据的数理统计模型,通过最小二乘法对故障数据的参数进行估计。Zhou^[10]对货舱门周围蒙皮损伤进行小样本统计和分析。Li等^[11]引入偏最小二乘回归算法,应对观测数量有限、缺失数据较多的情况,估算飞机成本关系。Deng^[12]以飞机结构件为研究对象,针对小样本数据驱动的装配偏差传递、预测与控制展开系统研究。

2.2 偏差分析与预测方法研究的相关工作

针对多维指标的多组数据,要对其进行特征提取和数据降维,主成分分析法是一种常用的方法。Chu等^[13]针对车门质量控制任务,构建偏差关系树,并用主成分分析法进行偏差诊断,找到产生偏差的主要影响因素。Hu等^[14]利用主成分分析法对轿车装配尺寸偏差进行误差识别,准确识别误差来源。

根据中心极限定理,若一个随机变量结果是由大量微小、独立因素加和构成的,则其可以近似地视为正态分布^[15]。因此可以基于历史的测量偏差数据,将偏差数据分布近似为正态分布,对未来偏差数据进行预测。Li等^[16]将地表沉陷近似为正态分布,并基于正态分布时间函数对其进行预测。Wang等^[17]引入了正态云模型,将正态分布扩展为类正态分布,更完整地描述了实际过程与正态分布的关系,并预测工厂生产的某零件直径。Li等^[18]在云模型中引入了隶属度的概念,用于衡量某一数据在某一分布中的符合程度,通过隶属度在多点的分布构成隶属云,分析了隶属云发生器的数学模型。Guo等^[19]使用基于组合逻辑的逻辑回归算法对广告点击率进行预测与分析,并用 k -折交叉验证等方法分析预测算法的性能。Cao等^[20]基于神经网络对价格指数进行预测,并通过RMSE等指标对预测模型的偏差进行评估。

实际生产是在制定的公差指导下进行的,因此基于公差也可以对偏差所处的区间进行预测。Liu等^[21]融合模型与理论的偏差数据,建立数据驱动模型,用于车身装配偏差预测。Xue等^[22]引入灰色系统理论,并基于灰色预测模型和精度检验对水电工程工期偏差进行预测。Han等^[23]基于历史质量偏差累积,对卷烟生产未来偏差进行预测,并提出调整方案。

3 模型假设与模型构建

3.1 基于主成分分析法的特征提取和关键零件识别

在对“小样本”的飞机质量偏差数据进行分析时,为了挖掘各零部件之间的数据关系,需要寻找能够有效表征大部分信息的数据。与此同时,关键零部件的准确定位也是实现偏差分析与控制的重要步骤。在此背景下,引入主成分分析法(Principal Component Analysis, PCA)和加权主成分分析法(Weighted Principal Component Analysis, WPCA)对偏差数据进行特征提取和关键零件识别。

PCA将原来具有一定相关性的指标(例如 P 个指标)进行重新组合,形成一组新的彼此之间互不相关的指标,这些新指标是原来 P 个指标的线性组合(或称综合性指标),用这组新的指标来代替原来的指标进行后续分析^[14]。

对案例分析部分介绍的数据进行主成分分析法过程的描述:在数据集 X 中有 $n=80$ 个样本,每个样本观测 $P=3$ 个变量,即:

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ \vdots & \vdots & \vdots \\ x_{80,1} & x_{80,2} & x_{80,3} \end{bmatrix} \quad (1)$$

定义 $x_i = (x_{i1}, x_{i2}, \dots, x_{i80,i})^T, i=1, 2, 3$ 。PCA将原来的3个观测变量 x_1, x_2, x_3 进行组合,形成3个新变量(组合变量),如式(2)所示:

$$\begin{cases} F_1 = \sum_{i=1}^{80} w_{11} x_{i1} \\ F_2 = \sum_{i=1}^{80} w_{12} x_{i2} \\ F_3 = \sum_{i=1}^{80} w_{13} x_{i3} \end{cases} \quad (2)$$

定义 $F = [F_1, F_2, F_3]$ 。利用主成分的方差贡献率度量变换后的 F 从原始数据集 X 中提取了多少信息。方差贡献率计算式如式(3)所示:

$$\alpha_i = \frac{\lambda_i}{\sum_{j=1}^3 \lambda_j}, i=1, 2, 3 \quad (3)$$

某一主成分对应方差贡献率越高,意味着该主成分综合原指标信息的能力越强。在解决实际问题时,一般选取前 d ($d < 3$)个主成分,使其累计方差贡献率满足达到80%以上的要求,用选取的前 d 个主成分代替原来的3个变量进行分析,可以实现数据降维的目的。

而WPCA根据不同样本和特征在识别过程中所起的作用不同,对其设定相应权重,不同程度地加强各样本和特征,从而对样本进行有效的特征提取,消除离群点的影响,以此来提高样本数据的识别率。此处我们对每个样本赋以合适的权重来凸显其在整个样本组中的不同重要性^[24]。采用“和法”计算权重公式如式(4)所示:

$$w_i = \frac{\sum_{j=1}^3 x_{ij}}{\sum_{k=1}^3 \sum_{l=1}^{80} x_{kl}}, i=1, 2, \dots, 80 \quad (4)$$

并对加权后的矩阵使用主成分分析法,得到WPCA分析结果。

若通过PCA和WPCA得到的结果显示某些主成分的累积方差解释比例较高,则这些主成分可以有效地表征原数据的大部分信息。此外,主成分中的系数能够揭示每个数据

维度在相应主成分中的权重,从而提供关于关键维度、关键零件的信息。

3.2 基于历史数据的装配偏差预测分析

在预测未来的飞机生产偏差数据时,基于人工智能的算法在“小样本”情境下性能表现相对较差,而基于概率论的计算方法更能有效地应对“小样本”的挑战,为未来的偏差数据提供更可靠的预测。因此,引入概率论的方法对偏差数据进行预测和生成。

在现实生活中,概率论的用处极为广泛。而正态分布在各种概率分布中居于首要地位,广泛存在于自然现象、社会现象、科学技术以及生产活动中。中心极限定理从理论上阐述了产生正态分布的条件:若决定某一随机变量结果的是大量微小的、独立的随机因素之和,并且每一因素的单独作用相对均匀且较小,没有一种因素可起到压倒一切的主导作用,那么这个随机变量一般近似于正态分布。

在飞机生产过程中,由于人员、机器、材料、方法、环境、测量等众多微小因素的共同作用,且这些因素相互间作用相对较小,因此可以近似于正态分布来考虑。为了更加严谨地衡量其分布的复杂性,用云模型来描述各种因素共同作用的随机性,将正态分布扩展为类正态分布,用一个新的独立参数——超熵,来衡量偏离正态分布的程度^[18]。

3.2.1 基于云模型的飞机生产偏差数据预测方法

在统计学和模糊数学的基础上,云模型把模糊性和随机性完全集成在一起,研究人类社会中不确定性的普遍规律。

云的数字特征用期望值 Ex 、熵 En 、超熵 He 3个数值来表征^[17],它完成定性和定量相互间的映射。其中,期望值 Ex 是概念在论域中的中心值,是最能代表这个定性概念的值,通常是云重心对应的 x 值,它应该100%地隶属于这个定性概念;熵 En 是定性概念模糊度的度量,反映了在论域中可被这个概念所接收的数值范围;超熵可以理解熵 En 的熵,反映了云滴的离散程度。云模型的示意图如图1所示。

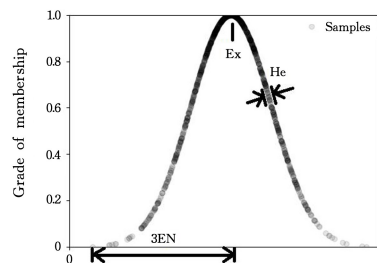


图1 云模型示意图

Fig.1 Cloud model diagram

正态云是最重要的云模型,因为社会和自然科学的各个分支都已证明了正态分布的普适性。对于正态云来说,由期望和熵可确定云的期望曲线方程如式(5)所示:

$$y = e^{-\frac{(x-Ex)^2}{2En^2}} \quad (5)$$

首先引入映射算法,根据云的统计特性,利用类正态数据 X 值,输出数据对应的云模型的期望、熵和超熵。在本文数据中, $n=80$ 。

算法1 映射算法

输入: $X = \{x_1, x_2, \dots, x_i, \dots, x_{80}\}$

输出:预测云模型的期望 Ex 、熵 En 和超熵 He

$$1. Ex = \text{mean}(x_i) = \frac{1}{80} \sum_{i=1}^{80} x_i$$

2. 一阶绝对中心矩 $M_1 = \frac{1}{n} \sum_{i=1}^n |x_i - Ex| = \frac{1}{80} \sum_{i=1}^{80} |x_i - Ex|$
3. 二阶平方矩 $M_2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - Ex)^2 = \frac{1}{79} \sum_{i=1}^{80} (x_i - Ex)^2$
4. $En = \sqrt{\frac{\pi}{2}} * M_1$
5. $He = \sqrt{M_2 - En^2}$

利用映射算法计算得到的云模型的 3 个数字特征值, 通过预测算法产生预测数据, 预测得到的数据和生成该云模型的 3 个数字特征的原始数据具有相同的分布规律, 代表的是同一个类正态分布。

算法 2 预测算法

输入: 预测云模型的期望 Ex , 熵 En 和超熵 He

输出: 预测值 $X = \{x_1, x_2, \dots, x_i, \dots, x_{80}\}$

1. 产生一个期望值为 En , 均方差为 He 的正态随机数 $\langle En \rangle'$;
2. 产生一个期望值为 Ex , 均方差为 $\langle En \rangle'$ 的正态随机数 x ;
3. 计算 $y = e^{-\frac{(x-Ex)^2}{2En^2}}$;
4. 令 (x, y) 为一个预测数据;
5. 重复步骤 1-步骤 4, 直到产生需求数量的预测值。

最后构建生成正态随机数算法, 通过输入的云模型期望和熵, 生成符合该云模型的若干模拟数据, 缓解“小样本”问题。

算法 3 生成正态随机数算法

输入: Ex, En

输出: 符合正态分布 $N(Ex, En)$ 的随机数 x

1. 产生 24 个 $[0, 1]$ 间的均匀随机数之和 $sum = \sum_{i=1}^{24} randi$, $randi$ 为 $[0, 1]$ 区间的均匀随机数;
2. 根据中心极限定理, 产生符合标准正态分布 $N(0, 1)$ 的随机数 $x = sum - 6$;
3. 将 x 平移变换为符合正态分布 $N(Ex, En)$ 的随机数 $x = Ex + x * En$;
4. 返回 $x^{[17]}$ 。

3.2.2 基于 k -折交叉验证和 RMSE 指标的预测结果验证

k -折交叉验证是一种用于评估预测模型性能的统计方法。在本文数据中应用取 $k=5$ 。 k -折交叉验证充分利用了整个数据集的信息, 避免了数据的浪费。

均方根误差(Root Mean Squared Error, RMSE)的计算式如式(6)所示:

$$RMSE = \sqrt{\frac{\sum_{i=1}^s (y_i - \hat{y}_i)^2}{s}} = \sqrt{\frac{\sum_{i=1}^{16} (y_i - \hat{y}_i)^2}{16}} \quad (6)$$

其中, $s=16$ 为样本个数, y_i 为第 i 个样本的真实值, \hat{y}_i 为第 i 个样本的预测值。 $RMSE$ 越小, 说明预测值和真实值越接近, 模型的拟合效果越好。

3.3 基于多源数据融合的装配偏差预测分析

“小样本”问题得以缓解后, 单次生产的偏差数据最小值和最大值更加精确、可靠。因此引入灰色预测模型和多源数据融合, 基于偏差数据最值和公差范围, 预测生产偏差数据所在区间, 并为公差修正机制的制定奠定基础。

飞机制造过程中涉及众多零部件和系统的组装, 这些数据来自不同装配单元和产品层, 具有不同的特性和准确性。装配偏差的预测需要充分考虑不同数据源的信息, 以便更准确地了解装配偏差的来源和影响^[12]。图 2 为基于多源数据融合的装配偏差波动区间预测示意图。

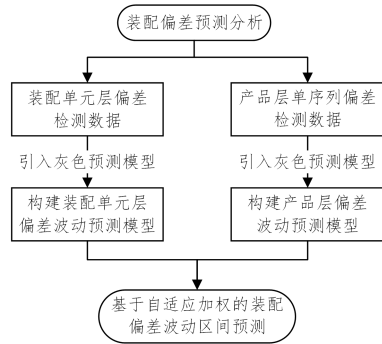


图 2 基于多源数据融合的装配偏差预测分析示意图

Fig. 2 Schematic diagram of assembly deviation prediction analysis based on multi-source data fusion

3.3.1 基于灰色预测模型的飞机生产偏差区间预测模型

灰色系统理论(Grey System Theory)是一种用于分析和建模具有不完全信息或数据不充分的系统的数学方法。其中, GM(1, N)模型是一种灰色预测模型, 适用于复杂的时间序列数据建模和预测。

在本文中设置 $N=2$, 即构建 GM(1, 2)模型。对于飞机壁板灰色系统^[25], 构建系统的特征数据序列 $X_1^{(0)} = (x_1^{(0)}(1), x_1^{(0)}(2), x_1^{(0)}(3))$ 和系统的相关因素数据序列 $X_2^{(0)} = (x_2^{(0)}(1), x_2^{(0)}(2), x_2^{(0)}(3))$, 计算 $X_i^{(1)} = (x_i^{(1)}(1), x_i^{(1)}(2), x_i^{(1)}(3)) (i=1, 2)$ 为 $X_i^{(0)} (i=1, 2)$ 的一阶累加生成序列, 其中, $x_i^{(1)}(k) = \sum_{j=1}^k x_i^{(0)}(j), k=1, 2, 3$ 。而 $Z_1^{(1)} = (z_1^{(1)}(2), z_1^{(1)}(3))$ 为 $X_1^{(1)}$ 的紧邻均值生成序列, 即 $z_1^{(1)}(k) = \frac{1}{2}(x_1^{(1)}(k-1) + x_1^{(1)}(k)), k=2, 3$ 。构建 GM(1, 2)模型 $x_1^{(0)}(k) + a z_1^{(1)}(k) = b_2 x_2^{(1)}(k)$ 。定义 $\hat{a} = (a, b_2, b_3)^T$, 则 GM(1, 2)模型针对参数 \hat{a} 的最小二乘估计值满足如下公式:

- 1) 当 $n < N+1$ 时, 且 $|B^T B| \neq 0, \hat{a} = B^T (B^T B)^{-1} Y$
- 2) 当 $n = N+1$ 时, 且 $|B| \neq 0, \hat{a} = B^{-1} Y$
- 3) 当 $n > N+1$ 时, 且 $|B^T B| \neq 0, \hat{a} = (B^T B)^{-1} B^T Y$

其中, $B = \begin{bmatrix} -z_1^{(1)}(2) & x_2^{(1)}(2) \\ -z_1^{(1)}(3) & x_2^{(1)}(3) \end{bmatrix}, Y = \begin{bmatrix} x_1^{(0)}(2) \\ x_1^{(0)}(3) \end{bmatrix}$ 。

定义 $\frac{dx_1^{(1)}(t)}{dt} + a x_1^{(1)}(t) = b_2 x_2^{(1)}(t)$ 为 GM(1, 2)模型的白化微分方程。则 GM(1, 2)模型的时间响应式如式(7)所示:

$$\hat{x}_1^{(1)}(k) = \left[x_1^{(0)}(1) - \frac{1}{a} b_2 x_2^{(1)}(1) \right] e^{-a(k-1)} + \frac{1}{a} b_2 x_2^{(1)}(k) \quad (7)$$

还原式为式(8)、式(9):

$$\hat{x}_1^{(0)}(1) = x_1^{(1)}(1) \quad (8)$$

$$\hat{x}_1^{(0)}(k) = \hat{x}_1^{(1)}(k) - \hat{x}_1^{(1)}(k-1), k=2, 3 \quad (9)$$

实际上, GM(1, 2)模型更加注重对关联因素之间的线性关系的挖掘, 而实际飞机生产过程中各个因素可能存在着复杂的非线性关系。对 GM(1, N)模型进行非线性优化, 可以加强预测区间的可靠性。GM(1, N, $x_1^{(0)}$)如式(10)所示:

$$x_1^{(0)}(k) + a^1 x_1^{(1)}(k-1) = b_2^1 x_2^{(1)}(k) \quad (10)$$

GM(1, N, $x_1^{(1)}$)如式(11)所示:

$$x_1^{(1)}(k) + a^2 x_1^{(1)}(k-1) = b_2^2 x_2^{(1)}(k) \quad (11)$$

由各个累加序列的定义可知, GM(1, N, $x_1^{(0)}$)模型、

GM(1, N, $x_1^{(1)}$)模型与 GM(1, N)模型之间可以相互转化,仅存在参数上的差别,通过 $a^1 = \frac{a}{1+0.5a}$, $b_1^1 = \frac{b_2}{1+0.5a}$, $a^2 = \frac{-1+0.5a}{1+0.5a}$, $b_2^1 = \frac{b_2}{1+0.5a}$ 这4个公式可以实现不同模型参数之间的转化。通过系数上的变化,将线性的模型进行非线性优化。

其中,GM(1, N, $x_1^{(0)}$)模型以原始值作为目标值,一般在数据增长较慢时选取;而 GM(1, N, $x_1^{(1)}$)模型以累加值作为目标值,当数据增长相对较快或数据波动频繁、需要充分降低数据的波动性和灰性时选取^[26-27]。

3.3.2 装配单元层检测数据的装配偏差波动预测

分析的飞机部件有 $n=3$ 个装配单元、 $N=2$ 个装配层级,其中第1个层级为特征数据序列,第2个层级为相关因素数据序列。特征数据序列输入为 $n=3$ 个装配单元的偏差数据最小值,应用3.3.1相关公式即可计算得到最小预测累加序列;同样地,输入偏差数据最大值,即可得到最大预测累加序列。公式如式(12)~式(15)所示:

$$\hat{X}_i^{(0)}(1)_{\min} = \hat{X}_i^{(1)}(1)_{\min} \quad (12)$$

$$\hat{X}_i^{(0)}(k)_{\min} = \hat{X}_i^{(1)}(k)_{\min} - \hat{X}_i^{(1)}(k-1)_{\min}, k=2,3 \quad (13)$$

$$\hat{X}_i^{(0)}(1)_{\max} = \hat{X}_i^{(1)}(1)_{\max} \quad (14)$$

$$\hat{X}_i^{(0)}(k)_{\max} = \hat{X}_i^{(1)}(k)_{\max} - \hat{X}_i^{(1)}(k-1)_{\max}, k=2,3 \quad (15)$$

因此可以得出装配单元层的装配偏差波动预测区间,如式(16)所示:

$$\{\hat{X}_1^{(0)}\}_{\text{assembly}} = \{\min x_1^{(0)}, \max x_1^{(0)}\} \quad (16)$$

3.3.3 产品层单序列检测数据的装配偏差波动预测

分析的飞机部件装配偏差 C_p 有 $\omega=3$ 个不同的偏差指标,记为 $C_p = (C_p^1, C_p^2, C_p^3)$,其中 $C_p^i = [C_{p\min}^i, C_{p\max}^i]$, $i=1,2,3$ 。分别构建装配偏差指标区间最小化序列 $C_{p\min} = \{C_{p\min}^1, C_{p\min}^2, C_{p\min}^3\}$ 和装配偏差指标区间最大化序列 $C_{p\max} = \{C_{p\max}^1, C_{p\max}^2, C_{p\max}^3\}$ 。运用同样的公式即可求解产品层的装配偏差波动预测区间,如式(17)所示:

$$\{\hat{C}_p\}_{\text{product}} = \{\min \hat{C}_p, \max \hat{C}_p\} \quad (17)$$

3.3.4 基于自适应加权的装配偏差波动区间预测

在本文的分析当中,由于仅考虑装配单元层和产品层两个链路的数据融合,因此 $o=2$ 。融合上面两个部分的计算结果,将装配单元层装配偏差波动预测区间和产品层装配偏差波动预测区间相结合,共同构建飞机部件偏差检测数据预测区间 D_p ,如式(18)所示:

$$D_p = \{\{\hat{X}_1^{(0)}\}_{\text{assembly}}, \{\hat{C}_p\}_{\text{product}}\} \quad (18)$$

定义不同链路偏差检测数据融合的权重因子为 w_i ($1 \leq i \leq 2$),并满足式(19)的要求。

$$\sum_{i=1}^2 w_i = 1 \quad (19)$$

其中, w_i 越大,代表对应的链路偏差检测数据对装配偏差的影响越大。装配偏差均方误差的计算式如式(20)所示:

$$\sigma^2 = \sum_{i=1}^2 w_i \sigma_i^2 \quad (20)$$

其中, σ_i^2 值越小,表明数据融合后该链路检测数据精度越高。基于二次函数最小值极限理论,需要得到最小均方误差 σ^2 ,则最优权重因子的计算式如(21)所示^[28]:

$$w_i' = \frac{\sigma_i^2}{\sum_{j=1}^2 \sigma_j^2} \quad (21)$$

根据自适应最优加权偏差检测数据融合,可得检测数据的最小融合值 H_{\min} 和最大融合值 H_{\max} :

$$H_{\min} = w_1' \{\hat{X}_1^{(0)}\}_{\text{assembly min}} + w_2' \{\hat{C}_p\}_{\text{product min}} \quad (22)$$

$$H_{\max} = w_1' \{\hat{X}_1^{(0)}\}_{\text{assembly max}} + w_2' \{\hat{C}_p\}_{\text{product max}} \quad (23)$$

据此构建最优加权数据融合后的飞机部件装配偏差预测波动区间 D_p' ,如式(24)所示:

$$D_p' = [H_{\min}, H_{\max}] \quad (24)$$

4 实验验证

4.1 数据来源与假设

本文选择以国内某民机公司生产的某型号飞机壁板的检测报告作为数据来源,一共考查80个关键控制点,每个点有3种零件的数据,分别是蒙皮、长桁、隔框的数据。即壁板由3种零件在80个关键控制点的组装而生产完成,生产质量偏差也主要来自这80个关键控制点。每个零件在关键控制点的装配都受到各自公差的影响,公差是每个零件在每个关键控制点理论上的最大误差范围,是控制精度计划的基准^[29]。

由于篇幅的问题,只选取其中20个点的数据进行展示,如表1所列。

表1 各个零件的偏差数据

Table 1 Deviation data of each component

关键控制点编号	蒙皮	长桁	隔框
1	0.072	0.113	0.168
2	0.055	0.142	0.200
3	0.025	0.210	0.267
4	0.084	0.090	0.148
5	0.061	0.088	0.138
6	0.047	0.102	0.154
7	0.048	0.189	0.239
8	0.040	0.145	0.200
9	0.104	0.102	0.153
10	0.102	0.154	0.210
11-70
71	0.078	0.139	0.195
72	0.069	0.149	0.203
73	0.102	0.092	0.144
74	0.086	0.141	0.195
75	0.072	0.130	0.192
76	0.062	0.083	0.138
77	0.064	0.219	0.283
78	0.067	0.147	0.206
79	0.085	0.065	0.115
80	0.077	0.141	0.209

其中,蒙皮装配公差设定为 ± 0.1 mm,长桁装配公差设定为 ± 0.2 mm,隔框装配公差设定为 ± 0.25 mm;在产品层,以装配完成后的蒙皮外形轮廓度、表面平滑度、波纹度作为装配质量指标,其设定的公差分别为 ± 0.3 mm, ± 0.8 mm, ± 0.5 mm。

为更便捷地用模型对数据进行分析,下面做一些假设:

- 1)在此处暂不考虑测量因素,即每个关键控制点的三维坐标值都是真实、准确的;
- 2)不考虑除关键控制点外其他位置的装配情况对飞机部件的影响;
- 3)此处飞机装配以蒙皮铆接为主;
- 4)假设整个工序合格,则针对测量数据的分析是有实际意义的;
- 5)同一个点3个零件的数据有一定的相关性;

6)生产偏差定义为实际生产的三维点与理论设定的三维点的欧氏距离;

7)生产偏差由众多微小因素共同作用构成,任一干扰都不是决定性的。

4.2 关键零件分析

基于主成分分析法和加权主成分分析法,将蒙皮、长桁、隔框3个零件的偏差数据视为每个关键控制点偏差数据的3个维度,消除样本个数维度大、分析难度大的问题,通过严谨的数理计算识别和分析关键零件。

通过主成分分析法计算可得主成分为:

$$\begin{bmatrix} -0.08838982 & 0.66238690 & 0.74392932 \\ -0.99571800 & -0.03845834 & -0.08406315 \\ -0.02707204 & -0.74817415 & 0.66294988 \end{bmatrix}$$

方差解释比例为:

$$[0.92721273 \quad 0.06840875 \quad 0.00437852]$$

通过加权主成分分析法计算可得主成分为:

$$\begin{bmatrix} 0.10757439 & 0.61867396 & 0.77824821 \\ -0.99240147 & 0.11384602 & 0.04667331 \\ 0.05972490 & 0.77735552 & -0.62621987 \end{bmatrix}$$

方差解释比例为:

$$[0.96814573 \quad 0.03035283 \quad 0.00150144]$$

由WPCA的结果分析可知,第一个主成分的方差解释比例接近97%,说明第一个主成分涵盖了原数据80个关键控制点的大部分信息。且第一个主成分对应的系数矩阵为 $[0.10757439, 0.61867396, 0.77824821]$,即在表征大部分信息的第一个主成分中,隔框偏差数据对应的权重最大,因此推断隔框为该部件当中的关键零件。

若不采用PCA或WPCA而直接对数据进行观察或分析,则无法精确地识别关键零件。其他常用的数据降维算法如线性判别分析是一种有监督的学习算法,岭回归算法侧重改善多重共线性问题,更适用于较高维的数据集。因此,针对本文的飞机壁板生产偏差数据,WPCA能够取得相对理想的降维效果。

4.3 类正态模型验证与预测

为了证明预测算法的合理性、正确性,采用5-折交叉验证,即 $k=5$ 。对于有80个关键控制点的数据,由于隔框为该部件中的关键零件,因此暂时只对隔框进行预测模型验证。进行5-折交叉验证,每次5-折交叉验证的过程都比较RMSE指标。结果如表2所列。

表2 基于k-折交叉验证和RMSE指标的模型验证

Table 2 Model validation based on k-fold cross-validation and

RMSE metric	
k-折交叉验证步骤序号	RMSE
1	0.022503984463698001
2	0.018770883926128832
3	0.006478228827434013
4	0.008866382692469019
5	0.024818012890879135

表2中的RMSE指标相对于隔框偏差数据和公差都较低,证明预测的效果较好。

基于正态云模型,利用实际生产的历史数据,预测20个来自于对应正态分布的数据并且展示它们各自的隶属度。同样地,主要展示对关键零件——隔框的生产质量偏差数据预测,如表3所列。

表3 对隔框的偏差数据预测结果

Table 3 Prediction results of deviation data for frames

预测点编号	预测偏差数据	对应隶属度
1	0.25642116	0.28653744
2	0.16924080	0.97591252
3	0.18352768	0.99723152
4	0.09434735	0.20924061
5	0.15456409	0.87163971
6	0.18509251	0.99431256
7	0.11029442	0.35501410
8	0.20031386	0.91503530
9	0.15302750	0.85681712
10	0.16459814	0.95106185
11	0.15295172	0.85607036
12	0.26292600	0.22958282
13	0.17924692	0.99990179
14	0.14427304	0.76224098
15	0.20883873	0.83646861
16	0.12612295	0.53886768
17	0.18840771	0.98474797
18	0.08038733	0.12048310
19	0.12491475	0.52394641
20	0.18731704	0.98839654

类似地,可以对蒙皮、长桁的类正态数据进行预测和生成。对实际生产数据和预测数据进行均值和标准差的比较,如表4所列。

表4 实际数据与预测数据的均值、标准差对比

Table 4 Comparison of mean and standard deviation between actual and predicted data

比较指标	实际生产偏差数据	预测偏差数据	预测偏差百分比/%
均值	0.17992499999999997	0.1663406851379428	7.55
标准差	0.048679763505999084	0.04592571742798609	5.66

实际数据与预测数据的均值和标准差的偏差百分比都处于相对较低的水平,预测效果较好。

4.4 偏差区间预测

将灰色预测模型应用到数据中,即可确定 $n=3, N=2$ 。由于数据之间不存在量纲的差异和绝对值上的大差距,因此省去数据净化的过程。

4.4.1 基于GM(1, N)模型的区间分析

特征数据序列为各个零件偏差数据的最小值和最大值构成的上下界区间。

$$\begin{aligned} X_1^{(0)} &= (x_1^{(0)}(1), x_1^{(0)}(2), x_1^{(0)}(3)) \\ &= ([0.022, 0.114], [0.029, 0.22], [0.083, 0.3]) \end{aligned}$$

相关因素数据序列为各个零件设定的公差值:

$$X_2^{(0)} = (x_2^{(0)}(1), x_2^{(0)}(2), x_2^{(0)}(3)) = (0.1, 0.2, 0.25)$$

计算累加指标:

$$X_1^{(1)} = ([0.022, 0.114], [0.051, 0.334], [0.134, 0.634])$$

$$X_2^{(1)} = (0.1, 0.3, 0.55)$$

$$Z_1^{(1)} = ([0.0365, 0.224], [0.0925, 0.484])$$

建立白化微分方程:

$$x_1^{(0)}(k) + a z_1^{(1)}(k) = b_2 x_2^{(1)}(k), k=2, 3$$

首先计算最小预测值。构造矩阵:

$$B = \begin{bmatrix} -0.0365 & 0.3 \\ -0.0925 & 0.55 \end{bmatrix}, Y = \begin{bmatrix} 0.029 \\ 0.083 \end{bmatrix}.$$

求解得到:

$$\hat{a} = \begin{bmatrix} -1.166 \\ -0.045 \end{bmatrix}$$

即 $a = -1.166, b_2 = -0.045$ 。

构建时间响应式:

$$\hat{x}_1^{(1)}(k) = \left[x_1^{(0)}(1) - \frac{1}{a} b_2 x_2^{(1)}(1) \right] e^{-a(k-1)} + \frac{1}{a} b_2 x_2^{(1)}(k)$$

计算得到:

$$\hat{x}_1^{(1)} = (0.022, 0.0698, 0.208)$$

基于还原式计算可得:

$$\hat{x}_1^{(0)} = (0.022, 0.0478, 0.1382)$$

同理可以计算最大预测值。最终得到装配单元层偏差范围在 $[-0.319, -0.022] \cup [0.022, 0.319]$ 区间内。同样地,计算产品层偏差范围在 $[-0.485, -0.022] \cup [0.022, 0.485]$ 区间内。

融合装配单元层和产品层数据:

$$D_p = \{[-0.319, -0.022] \cup [0.022, 0.319], [-0.485, -0.022] \cup [0.022, 0.485]\}$$

加权的最小预测值 $H_{\min} = 0.022$ 。对于加权的最大预测值,计算:

$$w_1' = \frac{0.319^2}{0.319^2 + 0.485^2} = 0.302$$

$$w_2' = \frac{0.485^2}{0.319^2 + 0.485^2} = 0.698$$

加权计算得到最大预测值 $H_{\max} = 0.435$ 。

综上, $(D_p)' = [-0.435, -0.022] \cup [0.022, 0.435]$ 。

4.4.2 基于改进的非线性优化 GM(1, N) 模型的区间分析

基于 GM(1, N, $x_1^{(0)}$) 模型求解: 首先计算最小预测值。

由最小二乘计算得到 $a^1 = -0.7366, b_2^1 = -0.028$; 根据时间响应式计算得到 $\hat{x}_1^{(1)} = (0.022, 0.049, 0.100)$; 基于还原式计算得到 $\hat{x}_1^{(0)} = (0.022, 0.027, 0.051)$ 。同理计算得到最大预测值对应的序列为 $\hat{x}_1^{(0)} = (0.114, 0.266, 0.317)$ 。

综上可知由 GM(1, N, $x_1^{(0)}$) 模型求解得到装配单元层偏差范围在 $[-0.317, -0.022] \cup [0.022, 0.317]$ 区间内。

基于 GM(1, N, $x_1^{(1)}$) 模型求解: 首先计算最小预测值。由最小二乘计算得到 $a^2 = -0.151, b_2^2 = -0.042$; 根据时间响应式计算得到 $\hat{x}_1^{(1)} = (0.022, 0.077, 0.145)$; 基于还原式计算得到 $\hat{x}_1^{(0)} = (0.022, 0.055, 0.068)$ 。同理计算得到最大预测值对应的序列为 $\hat{x}_1^{(1)} = (0.114, 0.5211, 0.680689484)$ 。

以上在装配单元层的结果表明 GM(1, N, $x_1^{(1)}$) 模型不适合在本数据集中使用。本数据集中数据波动相对平稳, 因此适用 GM(1, N, $x_1^{(0)}$) 模型进行非线性优化。

基于 GM(1, N, $x_1^{(0)}$) 模型同理计算产品层偏差范围在 $[-0.402, -0.022] \cup [0.022, 0.402]$ 区间内。融合装配单元层和产品层数据得到偏差所在区间为 $[-0.369, -0.022] \cup [0.022, 0.369]$ 。

4.4.3 与其他预测算法的对比分析

将基于 GM(1, N, $x_1^{(0)}$) 预测的区间范围与其他预测算法得到的结果进行对比, 如表 5 所列。

表 5 不同预测算法的结果对比

Table 5 Results comparison of different prediction algorithms

预测算法	装配单元层预测区间	产品层预测区间	加权预测区间
GM(1, N)	$[-0.319, -0.022] \cup [0.022, 0.319]$	$[-0.485, -0.022] \cup [0.022, 0.485]$	$[-0.435, -0.022] \cup [0.022, 0.435]$
GM(1, N, $x_1^{(1)}$)	$[-680684273, -0.022] \cup [0.022, 680684273]$
神经网络 ^[30]	$[-0.289, -0.018] \cup [0.018, 0.289]$	$[-0.647, -0.063] \cup [0.063, 0.647]$	$[-0.587, -0.059] \cup [0.059, 0.587]$
多元线性回归 ^[31]	$[-0.292, -0.015] \cup [0.015, 0.292]$	$[-0.171, -0.044] \cup [0.044, 0.171]$	$[-0.261, -0.041] \cup [0.041, 0.261]$
GM(1, N, $x_1^{(0)}$)	$[-0.317, -0.022] \cup [0.022, 0.317]$	$[-0.402, -0.022] \cup [0.022, 0.402]$	$[-0.369, -0.022] \cup [0.022, 0.369]$

本文最终选用的预测方法是 GM(1, N, $x_1^{(0)}$) 模型。GM(1, N) 模型、GM(1, N, $x_1^{(1)}$) 模型、神经网络得到的偏差区间范围较大, 不够精确, 无法为飞机生产过程提供较为有效的指导; 而多元线性回归预测的区间上界更加精确, 但部分实际生产的数据超出预测的区间上界, 与实际生产不符; GM(1, N, $x_1^{(0)}$) 模型得到的结果更加精确, 且符合实际生产, 能够进一步降低飞机生产过程中的超差率。

“小样本”问题的缓解, 也使得区间预测更加准确。参考公差预测得到的偏差区间范围, 为后续的实际生产和公差修正机制提供指导。

4.5 结果分析

在基于主成分分析法和加权主成分分析法的特征提取部分, 基于加权主成分分析法对 3 个零件维度数据进行分析, 得到第一个主成分的方差解释比例接近 97%, 数据的大部分信息都被集中在第一个主成分, 各个零件数据之间可能存在很强的线性关系。第一个主成分可以简化分析, 在某些情况下可以用第一个主成分来代表原始数据。根据第一个主成分的系数矩阵, 可知隔框为关键零件。

在类正态模型的验证和预测部分, 将实际生产的数据视为类正态数据, 首先利用 k -折交叉验证和 RMSE 指标评估了预测模型, 每一折的 RMSE 指标都处于较低的水平, 证明预测方法的性能较好。而后基于正态云模型对偏差数据进行预测, 重点关注关键零件, 生成了更大规模的模拟实际生产的偏差数据, 缓解了“小样本”的问题。

在基于灰色预测模型和多源数据融合的偏差区间预测部分, 首先计算装配单元层和产品层单序列两个数据链路的偏差预测区间, 并通过自适应加权, 将两个数据链路的区间加权求和得到飞机在该部件上的装配偏差波动区间的范围, 为生产提供一定的指导。非线性优化的灰色预测模型能够得到更准确的区间范围。其中“小样本”问题的缓解, 令区间预测更加精准、科学。

结束语 本文对小样本飞机生产质量偏差数据分析与预测方法进行了系统研究。通过基于主成分分析法和加权主成分分析法的特征提取模型, 成功挖掘了飞机各个零件维度特征之间的关系, 识别了飞机生产过程中的关键零件, 获取了飞机质量偏差预测方法的重点研究对象。考虑到飞机生产过程中的“小样本”问题, 采用了基于正态云模型的数据预测和生成模型, 有效模拟了飞机生产过程中多种微小因素的影响, 为飞机实际生产提供了可靠的数据支持。为了验证所采用的预测算法的性能, 引入了基于 k -折交叉验证和 RMSE 指标的验证算法, 从而确保了预测算法的可靠性和泛化能力。在指导飞机生产实践方面, 不仅仅局限于模拟数据, 更需要考察飞机生产偏差波动区间范围。通过基于非线性优化的灰色预测模型的偏差波动区间预测模型, 全面地了解飞机各个生产层的偏差情况, 为飞机实际生产提供更具指导性的信息; 并基于多源数据融合技术, 得到总的飞机偏差波动区间范围, 为公差修正机制的制定提供了重要依据。

参考文献

- [1] YU H D, LAI X M, LIN Z Q. Prediction and Control Method of Dimensional Accuracy in Assembly Manufacturing of Large Thin-walled Structures in Aerospace[J]. Shanghai Aerospace, 2020, 37(3):1-10.
- [2] SU C, WANG Z, XU H. Analysis of Assembly Deviation of Flexible Parts Based on Finite Element Method[J]. Machinery Manufacturing & Automation, 2019(4):79-83.
- [3] ZHU P, YU J, ZHENG X, et al. Modeling of Deviation Transmission Network and Error Tracing in Mechanical Assembly Process[J]. Journal of Zhejiang University: Engineering Science Edition, 2019, 53(8):12.
- [4] CHENG L. Research on Modeling and Key Technologies of Digital Assembly Deviation for Large Aircraft Based on Key Characteristics[D]. Hangzhou: Zhejiang University, 2014.
- [5] SUN H. Analysis and Prediction of Assembly Deviation for Aircraft Panel-like Flexible Components[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2015.
- [6] ZHAO D P, TIAN X Y, GENG J H. Prediction of Motion Assembly Accuracy Based on Clearance Connectors and Multi-dimensional Vector Rings[C]//The third National Modern Manufacturing Integration Technology Academic Conference. 2014: 1-12.
- [7] YANG Z J. Research on Predictive Control Theory of Nonlinear Systems in Aircraft[D]. Wuhan: Wuhan University of Technology, 2006.
- [8] HUGUES E, CCR E, CHARPENTIER E, et al. Application of Markov processes to predict aircraft operational reliability[J/OL]. 2002. <https://www.semanticscholar.org/paper/Application-of-Markov-processes-to-predict-aircraft-Hugues-Ccr/bd5be5dea7b4f4beca1206ab9042eb250d5e3118>.
- [9] WANG B. Study on the Fault Trend of General Aircraft Structure under the Condition of Small Sample Data[D]. Deyang: Civil Aviation Flight University of China, 2013.
- [10] ZHOU F W. Small Sample Analysis of Skin Damage Around B737NG Cargo Door of [J]. Aviation Maintenance & Engineering, 2021(1):57-59.
- [11] LI S A, SONG B F, ZHANG H X. Method for Multivariate Analysis with Small Sample in Aircraft Cost Estimation[J]. Journal of Aircraft, 2007, 44(3):1042-1045.
- [12] DENG B. Prediction and Control of Assembly Deviation Transmission for Aircraft Structural Components Driven by Small Sample Data[D]. Nanchang: Nanchang Hangkong University, 2021.
- [13] CHU G, ZHANG S, JIN S. Research on Deviation Control of Car Body Doors in Sedans[J]. Machinery Design & Manufacture, 2003(5):3.
- [14] HU M, LAI X, LIN Z. Application Research of Principal Component Analysis Method in Dimensional Deviation of Car Assembly[J]. China Mechanical Engineering, 2002, 13(6):3.
- [15] LI F Q, GAO W S. Probability Theory and Mathematical Statistics[M]. Jilin: Jilin University Press, 2004.
- [16] LI C Y, GAO Y G, CUI X M. Research on Prediction of Surface Dynamic Settlement Based on Normal Distribution Time Function[J]. Rock and Soil Mechanics, 2016(S1):9.
- [17] WANG Z N, XIAO M Q, LI Y, et al. Prediction Algorithm of Cloud Model for Log-normal Distribution Data[J]. Computer Applications and Software, 2009, 26(9):3.
- [18] LI D Y, MENG H J, SHI X M. Membership Cloud and Its Generator[J]. Journal of Computer Research and Development, 1995, 32(6):6.
- [19] GUO S Z, LIAO X F, XIAN K Y. Logistic Regression Click Prediction Algorithm Based on Composite Structure[J]. Computer Science, 2024, 51(2):73-78.
- [20] CAO L, SHANG W, XIE S Y, et al. Research on Price Index Prediction Based on AGNN Public Opinion Index Network[J]. Chinese Journal of Management, 2023, 20(3):411.
- [21] LIU Y H, ZHANG S M, CHU G P. Combination modeling of auto body assembly dimension propagation considering multi-source information for variation reduction[J]. Assembly Automation, 2019, 39(4):514-522.
- [22] XUE Z H, ZHANG K. Analysis of Duration Deviation Prediction Based on Grey System Theory[J]. Shanxi Architecture, 2009, 35(8):215-216.
- [23] HAN L, LIU X, WANG Z, et al. Integrated Control of Cigarette Rolling Quality Combining Deviation Prediction with Parameter Optimization[J]. Packaging Engineering, 2023, 44(15):217-222.
- [24] CHEN P. Research on Principal Component Analysis Method and Its Application in Feature Extraction[D]. Xi'an: Shaanxi Normal University, 2014.
- [25] ZENG L. Improved Grey Multivariate GM(1, N) Model and Its Application [J]. Journal of Southwest University: Natural Science Edition, 2019, 41(9):9.
- [26] ZHOU W, FANG Z G. Research on Nonlinear Optimization GM(1, N) Model and Its Application. Systems Engineering and Electronics, 2010(2):5.
- [27] FU Z W, YANG Y K, WANG T Y. Prediction of Urban Domestic Water Demand in Haiyan County Based on Improved Nonlinear Optimization GM(1, N) Model[J]. Water Resources and Power, 2019, 37(10):4.
- [28] THORDARSON F O, MADSEN H, NIELSEN H A, et al. Conditional weighted combination of wind power forecasts[J]. Wind Energy, 2010, 13(8):751-763.
- [29] XU Y Q. Precision Control in Shipbuilding; Mathematical Methods of Tolerance Theory[J]. Journal of Zhenjiang Shipbuilding Institute, 1988(Z1):49-58.
- [30] MCCULLOCH W S, PITTSW. A Logical Calculus of the Ideas Immanent in Nervous Activity[J]. Journal of Symbolic Logic, 1943, 9(2):49-50.
- [31] WACHSMUTH A, WILKINSON L, DALLAL G E. Galton's Bend; An Undiscovered Nonlinearity in Galton's Family Stature Regression Data and a Likely Explanation Based on Pearson and Lee's Stature Data[J/OL]. 2003. <https://www.semanticscholar.org/paper/Galton%27s-Bend%3A-An-Undiscovered-Nonlinearity-in-Data-Wachsmuth-Wilkinson/flbdf6b19c6cac497147274cef8a14d448d23b4?pdf>.



WANG Luhang, born in 2002, undergraduate. His main research interest is aircraft production quality control.



ZHANG Dongdong, born in 1977, Ph.D., professor, Ph.D supervisor. Her main research interests include image processing and deep learning.