

doi: 10.7690/bgzdh.2022.10.009

基于优化 VARIMA 的无人机发动机飞行参数预测模型

王凤芹¹, 耿宝亮², 杜晶¹

(1. 海军航空大学航空基础学院, 烟台 264001; 2. 海军航空大学作战勤务学院, 烟台 264001)

摘要: 为了预测无人机发动机飞行参数值, 建立一种多参数关联的向量求和自回归移动平均模型(vector auto regression integrated moving average, VARIMA)。模型以无人机飞行数据作为输入, 利用模拟退火算法对 VARIMA 模型参数进行优化, 构建多源飞行参数关联的数据模式, 利用构建好的数据模式实现状态参数的预测。选取无人机多次飞行的飞参数据进行实验。实验结果表明: 优化后的 VARIMA 预测模型预测性能好, 比优化前的预测用时节省了 0.23 s。

关键词: 飞行参数预测; 无人机健康管理; 求和自回归移动平均模型; 向量自回归模型; 时间序列分析算法
中图分类号: V247.5 **文献标志码:** A

Flight Parameter Prediction Model of UAV Engine Based on Optimized VARIMA

Wang Fengqin¹, Geng Baoliang², Du Jing¹

(1. College of Basic Sciences for Aviation, Naval Aviation University, Yantai 264001, China;
2. College of Combat Service, Naval Aviation University, Yantai 264001, China)

Abstract: In order to predict the flight parameters of UAV engine, a vector auto regression integrated moving average (VARIMA) model with multi-parameter correlation was established. The model takes the UAV flight data as the input, uses the simulated annealing algorithm to optimize the VARIMA model parameters, constructs the data pattern of multi-source flight parameters association, and uses the constructed data pattern to realize the prediction of state parameters. The flight data of multiple flights of UAV are selected for the experiment. The experimental results show that the optimized VARIMA prediction model has good prediction performance, and the prediction time is saved by 0.23 s compared with that before optimization.

Keywords: flight parameter prediction; UAV health management; ARMA model; VAR model; time series analysis algorithm

0 引言

随着无人机的广泛应用, 其飞行安全成为无人机领域重点关注的问题。发动机作为无人机的核心分系统, 其工作状态是影响无人机飞行安全的关键因素。国内外学者对发动机状态参数的监测进行了大量研究与尝试, 方法大致可以分为 2 类: 1) 基于部件物理模型的方法, 例如 Borguet 等^[1]通过建立发动机各部件性能参数的回归残差模型, 实现发动机各性能参数的异常检测; Tsoutsani 等^[2]通过构建部件健康状况的动态模型实现发动机剩余使用寿命预测。2) 基于数据驱动的方法, 例如 Sina 等^[3]和 Fast 等^[4]对发动机构建神经网络监测模型, 用于发动机的状态监测; 潘鹏飞^[5]利用航空发动机在试飞过程中产生的数据, 建立基于遗传算法改进优化的神经网络的参数预测模型; 吕超等^[6]设计一种基于贝叶斯信息准则的密度峰值聚类分析算法和共享邻域算法相结合的算法对航空发动机数据进行准确

类与标记, 为航空发动机状态监测提供数据支持; 杨学森等^[7]针对传感器得到的缺失数据和误差数据以及传感器失效的情况下, 提出一种动态数据融合模型, 该模型对涡扇发动机的物理参数进行数据融合并预测。基于物理模型的方法构建模型复杂, 预测结果受限于模型准确度的限制^[8-9]。文献[3]—[7]提供基于航空发动机的试验数据或仿真数据的监测模型, 但未针对实际飞行过程发动机产生的飞参数据进行分析。

随着传感器技术的发展, 飞机上安装了成百上千个传感器, 将传感器获得的飞机各工作部件的状态参数、飞行姿态参数和环境信息等与飞行相关的参数, 采集并存储到快速存取记录器(quick access recorder, QAR)中^[10]。如何充分利用这些数据, 对其进行有效地分析处理, 为飞行质量评估、飞行安全监控提供辅助决策是当前飞行领域中的一大研究热点。

收稿日期: 2022-06-18; 修回日期: 2022-07-10

基金项目: 大学自主立项科研课题(2021HDZZLX012)

作者简介: 王凤芹(1982—), 女, 山东人, 博士, 副教授, 从事装备保障信息化、大数据分析与应用研究。

E-mail: nudtwangfengqin@163.com。

笔者针对无人机历史飞行过程中产生的飞参数数据，首先采用数据预处理技术对无人机飞行大数据进行处理，然后构建改进向量求和自回归移动平均模型 (VARIMA) 实现发动机状态参数预测。模型基于模拟退火算法对 VARIMA 的参数进行优化，既考虑发动机各监控参数之间的内在逻辑关系，又考虑了参数自身的时间和误差扰动的影响。

1 无人机发动机飞行参数预测建模

1.1 问题分析

在飞行过程中，无人机借助传感器采集得到了大量的 QAR 飞参数数据，其中发动机采集得到的主要状态参数及其意义如表 1 所示。无人机操控人员通过监控这些状态参数的取值来监测发动机的工作状态。当状态参数的取值超出其上下限，或者数据出现异常波动情况，操控人员要进行相应的故障检测与处理，防止飞行事故的发生。无人机操控人员希望通过分析无人机历史 QAR 飞参数数据，预测发动机状态参数变化趋势，在下次起飞前事先了解发动机的状态参数未来一段时间的情况，为故障检测与处理留有更充分的时间。

表 1 无人机发动机主要状态参数及其意义

序号	状态参数	参数意义
1	发动机转速	属于快速变化的信号，表明发动机的瞬时工作状态，实时反映发动机的工况是否正常
2	节流门位移	反映了发动机的功率状态。在 100% 节流门开度时，发动机可以发挥出最大连续功率；节流门开度减小，发动机的功率逐渐降低
3	滑油压力	滑油压力过低或者因滑油管堵塞而造成的滑油压力过高，发动机磨损会较大，磨损太大会损坏发动机
4	排气温度	属于快速变化的信号，表明发动机的瞬时工作状态，实时反映发动机的工况是否正常

由于 QAR 飞参数数据是时间序列数据，各参数的相邻取值具有依赖性，可以通过时间序列分析算法，分析按照时刻点采集得到的各个参数所反映出来的变化规律与发展趋势，从而预测参数未来可能的取值。美国学者 Box 和英国学者 Jenkins 针对基于时间序列的参数，对参数进行平稳性分析，非平稳序列转化为平稳序列后选择合适的模型，计算模型参数，利用模型进行参数预测，并对预测结果进行评价的方法，称为 Box-Jenkins 方法^[10]。该方法为回归分析方法，也被称为求和自回归移动平均 (auto regression integrated moving average, ARIMA) 模型，该模型对一元参数的时间序列 $X_t = \{x_1, x_2, \dots, x_t\}$ ，参

数间的关系表达式为：

$$\phi(B)\nabla^d x_t = \Theta(B)\varepsilon_t \tag{1}$$

式中： $\phi(B) = \prod_{i=1}^p (1 - \lambda_i B)$ 为 p 阶自回归系数多项式； λ_i

为自回归方程的 p 个特征根； $\nabla^d x_t = \sum_{i=0}^d (-1)^i C_d^i x_{t-i}$ ，

$C_d^i = d! / i!(d-i)!$ 为对序列进行 d 阶差分运算；

$\Theta(B) = \prod_{k=1}^q (1 - \lambda_k B)$ 为 q 阶移动平均系数多项式；

$$\varepsilon_t = \frac{x_t}{\prod_{k=1}^q (1 - \lambda_k B)}$$

为随机干扰项。

ARIMA 模型不仅表示一元参数时间序列从过去到未来的内部相互依赖关系，同时表示现在时刻与过去时刻扰动之间的关系。在 QAR 记录的参数中，发动机状态的关联参数有多个，这些参数相互联系、相互影响，只采用针对一元指标建立的 ARIMA 模型，不能考虑到多元参数之间的影响。Sims 最早提出了向量自回归 (vector auto regressive, VAR) 模型。相对于 ARIMA 模型，VAR 模型考虑了多个参数间的滞后交叉影响。以二元参数序列 X_t 和 Y_t 为例，VAR 模型表示为：

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} w_x \\ w_y \end{bmatrix} + \begin{bmatrix} \alpha_{x1} & \beta_{x1} \\ \alpha_{y1} & \beta_{y1} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \dots + \begin{bmatrix} \alpha_{xp} & \beta_{xp} \\ \alpha_{yp} & \beta_{yp} \end{bmatrix} \begin{bmatrix} X_{t-p} \\ Y_{t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \end{bmatrix} \tag{2}$$

笔者采用将 ARIMA 模型和 VAR 模型组合的方式构建预测模型 VARIMA。另外，对于模型的参数 p 和 q 的设定，传统方法一般根据经验值设定初始值，然后在一段范围内采用逐个尝试的方法，最后从中取最优参数设定。这种方法收敛速度较慢，还有可能不是最优解。笔者对 VARIMA 模型进行改进，采用模拟退火算法对参数 p 和 q 进行优化设置，提高算法收敛速度。

1.2 基于优化 VARIMA 的飞行参数预测模型

将 ARIMA 模型和 VAR 模型组合，建立 VARIMA 模型，模型以无人机历史 QAR 飞参数数据作为输入，抽取发动机状态参数数据，通过 VARIMA 时间序列分析算法，利用模拟退火算法优化模型参数取值，实现发动机未来状态参数值的预测，模型工作流程如图 1 所示。

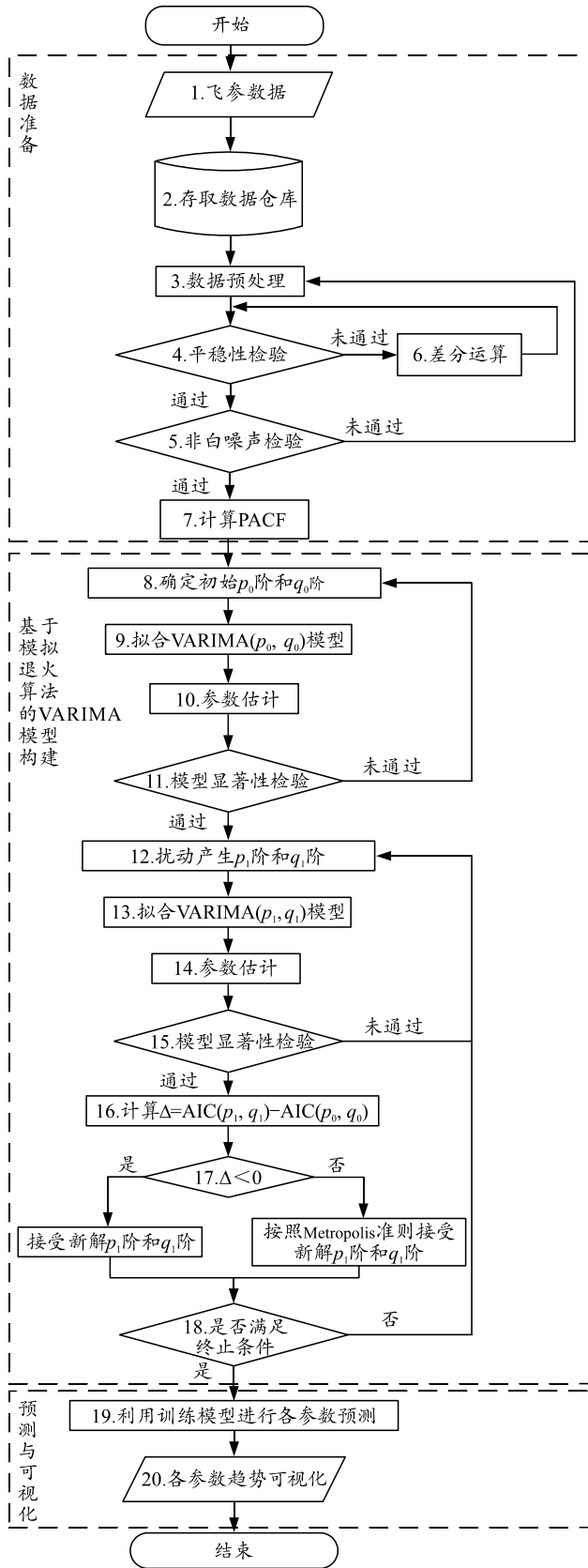


图 1 基于改进 VARIMA 状态参数预测模型

基于改进 VARIMA 发动机状态参数预测模型整个流程分为 3 部分：1) 数据准备；2) 基于模拟

退火算法的 VARIMA 模型构建；3) 预测与可视化。具体步骤包括：

第 1 步：利用数据抽取技术 (extract-transform-load, ETL) 将无人机各飞行架次历次飞行的 QAR 数据抽取到数据仓库。目的是将历史 QAR 数据整合到一起，为无人机的飞行决策提供分析依据。

第 2 步：将 QAR 数据存储到面向主题构建的无人机系统的数据仓库。为了分析无人机历次飞行产生的 QAR 数据，按照面向主题的方式构建数据仓库，以提高数据查询统计分析效率。

第 3 步：从数据仓库中读取数据，对数据进行预处理。QAR 中的数据来源于多个传感器，由于传感器工作状态可能出现不稳定、网络传输数据出现误差等问题造成采集的数据存在缺失值、异常值、冗余值等野值，首先需要对其进行数据预处理，才能进行数据分析。

数据预处理内容包括：

1) 数据二次采样。

QAR 参数采样频率较高，几十毫秒甚至几毫秒采集一次数据，如果对这些数据不加裁剪，直接进行分析，时间空间开销都比较大，为降低数据量级，提高分析效率，对数据样本进行二次采样，采样频率设置为每隔 1 s 截取数据。

2) 缺失值填补，异常值替换。

依据 QAR 飞参数据渐进变化的特点，对于缺失值、异常值，采用与其相邻数值的中位数进行填补、替换。

3) 数据标准化。

由于发动机转速、滑油温度、滑油压力和排气温度 4 个状态参数的单位和范围不同，而且它们的方差数量级差别很大，方差数量级较大的特征参数在目标函数值中所占比重较大，将导致方差数量级较小的特征参数不能有效影响模型，从而导致回归模型不准确，所以要对不同量级的特征参数进行标准化。Z-score、Min-max 和 MaxAbs 标准化方法为常见的 3 种数据标准化方法。由于传感器的不稳定性，一般会有离群数据存在，故选用 Z-score 数据标准化方法，即采用特征参数的测量值减去其均值，再除以其标准差的方法。

第 4 步：对预处理后的基于时间序列的各参数进行平稳性检验。由于向量自回归模型只适用于平稳的时间序列数据，所以需要对其进行平稳性检验之后才能构建模型。所谓平稳性检验，是通过判

别参数的自相关系数是否随着滞后阶数的增加而很快趋于 0，如果趋于 0 则序列是平稳的，否则序列不平稳。自相关系数 ρ_k 为：

$$\rho_k = \frac{\sum_{t=1}^{n-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2}, \quad \forall 0 < k < n. \quad (3)$$

如果 $\rho_k \rightarrow 0$ ，则序列平稳，转第 4 步，否则，转第 5 步。

第 5 步：对于平稳序列，再进行是否为白噪声序列，只有非白噪声序列才能构建 ARIMA 模型。白噪声的检验借助 Ljung-Box 检验。Ljung-Box 检验的原假设 H_0 ：数据总体的相关系数 $\rho_1 = \rho_2 = \dots = \rho_k$ 为 0，备择假设 H_1 ：原本的数据不独立，存在某个 $\rho_k \neq 0 (k \leq n)$ ，构造的统计量 $Q(k)$ ：

$$Q(k) = n(n+2) \sum_{i=1}^k (\rho_i / (n-i)). \quad (4)$$

式中：统计量 $Q(k)$ 服从自由度为 n 的卡方分布， n 为样本数目， k 为指定的数； ρ_i 为 i 阶自相关系数。白噪声序列是指对于给定的显著性水平 α ，其拒绝域 $Q(k) > \chi^2_{1-\alpha, n}$ 。若为白噪声序列，转第 2 步；非白噪声序列则转第 6 步。

第 6 步：对于非平稳序列，采用差分运算，对

$$\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = \begin{bmatrix} w_x \\ w_y \end{bmatrix} + \begin{bmatrix} \alpha_{x1} & \beta_{x1} \\ \alpha_{y1} & \beta_{y1} \end{bmatrix} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \dots + \begin{bmatrix} \alpha_{xp} & \beta_{xp} \\ \alpha_{yp} & \beta_{yp} \end{bmatrix} \begin{bmatrix} X_{t-p} \\ Y_{t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{xt} \\ \varepsilon_{yt} \end{bmatrix} - \begin{bmatrix} \theta_{x1} & \sigma_{x1} \\ \theta_{y1} & \sigma_{y1} \end{bmatrix} \begin{bmatrix} \varepsilon_{xt-1} \\ \varepsilon_{yt-1} \end{bmatrix} - \dots - \begin{bmatrix} \theta_{xq} & \sigma_{xq} \\ \theta_{yq} & \sigma_{yq} \end{bmatrix} \begin{bmatrix} \varepsilon_{xt-q} \\ \varepsilon_{yt-q} \end{bmatrix}. \quad (8)$$

该模型充分考虑了多元参数之间的相互影响，同时考虑了过去时刻的扰动对多元参数的影响。

第 10 步：对 VARIMA(p_0, q_0) 模型中的未知参数矩阵中的参数进行估计。矩估计、极大似然估计和最小二乘法为常见的 3 种估计方法。

第 11 步：针对构建的 VARIMA(p_0, q_0) 模型进行显著性检验，通过判断模型拟合后的残差序列是否为白噪声序列来说明拟合模型是否有效。若残差序列为白噪声序列，则模型拟合较好，相关信息都被拟合进模型中；若残差序列为非白噪声序列，则说明模型拟合不好，残差序列中还存在未被提取的相关信息，需转第 8 步重新调整模型初始阶数 p_0 和 q_0 重新进行拟合模型。

第 12 步：依据自相关系数 ρ_k 计算偏相关系数 PACF $_k$ 的性质，分别在阶数 p_0 和 q_0 附近产生新的 p_1 和 q_1 。

第 13 步：根据确定的新阶数 p_1 和 q_1 ，建立 VARIMA(p_1, q_1) 模型。

第 14 步：对于 VARIMA(p_1, q_1) 模型中的未知参数矩阵中的参数进行估计。

于波动较大的参数值用该值相邻项值差替代，从而降低数据之间的差异而趋于平稳，差分后转第 4 步。

第 7 步：根据自相关系数 ρ_k 计算偏相关系数 PACF $_k$ (partial correlation coefficient)，计算公式：

$$PACF_k = D_k / D, \quad \forall 0 < k < n; \quad (5)$$

$$D = \begin{vmatrix} 1 & \rho_1 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \dots & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \dots & 1 \end{vmatrix} \dots; \quad (6)$$

$$D_k = \begin{vmatrix} 1 & \rho_1 & \dots & \rho_1 \\ \rho_1 & 1 & \dots & \rho_2 \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \dots & 1 \end{vmatrix}. \quad (7)$$

第 8 步：确定模型初始自相关阶数 p_0 和移动平均阶数 q_0 ，这 2 个值的确定取决于自相关系数 ρ_k 和偏相关系数 PACF $_k$ 的性质。

第 9 步：根据确定的阶数，建立 VAR(p_0, q_0) 和 ARIMA(p_0, q_0) 相结合的 VARIMA(p_0, q_0) 模型。以二元参数序列 X_t 和 Y_t 为例，VARIMA(p, q) 模型表示为：

第 15 步：针对构建的 VARIMA(p_1, q_1) 模型进行显著性检验，检验模型是否有效，检验未通过转第 12 步重新调整模型阶数 p_1 和 q_1 进行拟合模型。

第 16 步：从所有通过显著性检验的 VARIMA(p_0, q_0) 和 VARIMA(p_1, q_1) 模型中选择最优模型。利用赤池信息准则 (akaike information criterion, AIC) 来判断模型的好坏。AIC 表达式为：

$$AIC = 2(p + q + 1) - 2 \ln(L). \quad (9)$$

式中 L 为模型的极大似然函数，该函数值越大，模型阶数 p 和 q 越低，则 AIC 值越小，说明模型越简洁，模型越优。

第 17 步：通过判断

$$\Delta = AIC(p_1, q_1) - AIC(p_0, q_0). \quad (10)$$

正负确定如何接受新的阶数 p_1 和 q_1 。若 $\Delta < 0$ ，则直接接受新解 p_1 和 q_1 ，否则按照 Metropolis 准则接受新的阶数 p_1 和 q_1 。此处为模拟退火算法思想的主要体现，借助 Δ 值进行阶数的选取，避免穷举阶数而造成的时间开销。

第 18 步：判断是否满足以下终止条件：

- 1) 连续若干个解都未被接受；

- 2) 温度超过设定的阈值;
- 3) 到达最大迭代次数。

若满足终止条件, 则结束最优模型选择过程, 否则, 转第 12 步继续寻找最优解。

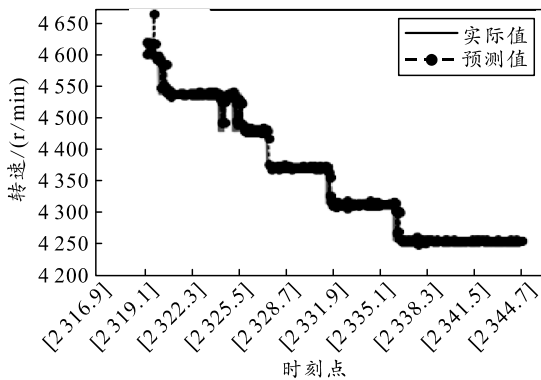
第 19 步: 利用构建的最优拟合模型 VARIMA(p, q) 进行参数预测, 监控发动机参数变化趋势。

第 20 步: 将各参数预测结果可视化输出。

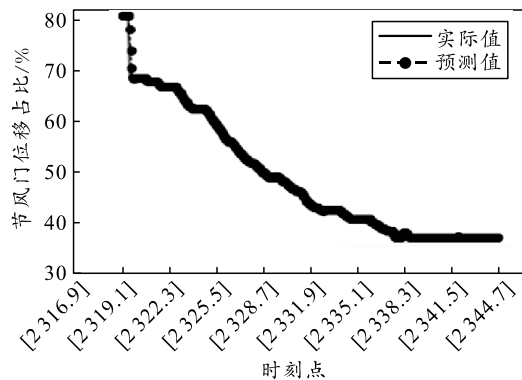
2 实验与结果分析

2.1 实验过程

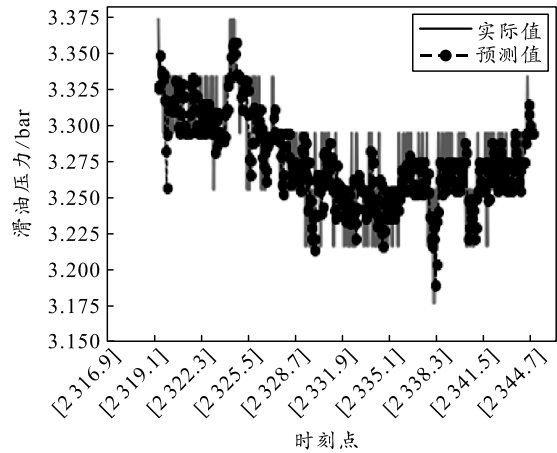
选取某型无人机 2020 年部分架次的活塞式发动机飞参数据作为实验数据, 取 80% 数据用来构建模型, 20% 数据用来预测。无人机飞行过程包括起飞、爬升、巡航、下降、进近、着陆 6 个阶段。由于不同阶段发动机的状态参数变化趋势各有不同, 所以将实验数据分成 6 部分 QAR 数据分别进行分析, 建立 VARIMA 模型, 预测参数取值。实验的编程语言为 Python, 开发环境为 Spyder, 用 numpy 库进行矩阵相关计算, 用 pandas 库进行数据读写分析, 用 statsmodels 库进行统计值的计算, 用 matplotlib 库绘制曲线图进行数据可视化。可视化输出如图 2 所示, 时间点为横轴。从图 2 中可以看出, 转速与节流门位移 2 个变量模型拟合的非常好, 对于滑油压力和排气温度参数存在一些误差。



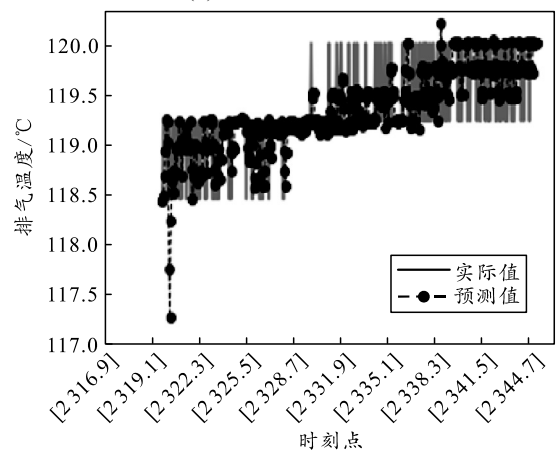
(a) 发动机转速预测



(b) 节流门位移预测



(c) 滑油压力预测



(d) 排气温度预测

图 2 参数预测结果可视化

2.2 实验结果分析

将 3 个模型 (VARIMA、ARIMA、VAR 模型) 分别对某架无人机单次完整飞行的飞行数据进行分析预测, 采用均方误差 (mean square error, MSE)、均方根误差 (root mean square error, RMSE) 和平均绝对误差 (mean absolute error, MAE) 这 3 个评价指标来衡量每个模型的预测效果, 即将预测值与实际值分别求解绝对误差、均方差和均方根误差, 对比结果如图 3 所示。

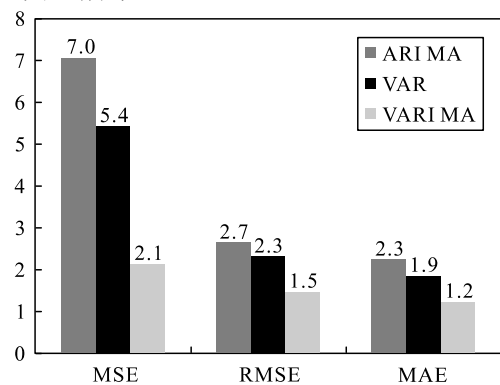


图 3 3 种算法预测性能指标对比

从上图中可以看出,优化 VARIMA 模型的误差明显低于 ARIMA 和 VAR,预测指标 MSE 低 4.9 和 3.3, RMSE 低 1.2 和 0.9, MAE 低 1.0 和 0.6,说明 VARIMA 模型的预测性能好。分析其原因,VARIMA 模型既考虑了参数之间的内在逻辑关系,又考虑了参数自身的时间影响和误差扰动的影响,将参数之间的相互作用考虑到预测模型,使得模型预测精度更高。

基于模拟退火算法优化后的 VARIMA 模型预测时间与优化前只是穷举参数方式的预测时间对照如表 2 所示。对于 2 000 条数据,优化后和优化前所用时间相差不大,随着数据量的增大,优化后模型用时比优化前都少,对于 20 000 条数据,VARIMA 模型优化后比优化前节省了 0.23 s。

表 2 VARIMA 算法优化前后预测时间对比 s

算法	数据量			
	2 000	5 000	10 000	20 000
优化后 VARIMA	0.08	0.22	0.51	0.84
VARIMA	0.08	0.26	0.62	1.07

3 结论

发动机作为无人机的核心分系统,其工作状态是影响无人机飞行安全的关键因素。为了监控无人机发动机的工作状态,将无人机系统采集得到的大量飞行过程中基于时间序列的飞参数据作为输入,提取飞参数据中影响发动机工作状态的关键参数,建立了基于多参数关联的向量求和自回归移动平均模型,利用模拟退火算法对模型参数进行了优化设置,提高了模型收敛速度。所建立的模型能有效预测发动机各关键参数的未来趋势,并将趋势可视化输出。针对无人机多个飞行架次的飞参数据进行实

验,实验结果表明:模型有效,预测精度高,无人机操控人员可以利用预测值及时发现发动机存在的安全隐患,为发动机的视情维修提供数据支持。

参考文献:

- [1] BORGUET S, LEONARD O, DEWALLEF P. Regression-Based Modeling of a Fleet of Gas Turbine for Performance Trending[J]. Journal of Engineering for Gas Turbines and Power, 2016, 138(2): 02120-1-021201-9.
- [2] TSOULSANIS E, MESKIN N, BENAMMAR M, et al. A Dynamic Prognosis Scheme for Flexible Operation of Gas Turbine [J]. Applied Energy, 2016, 164: 686-701.
- [3] SINA T S, SADOUGH V Z N, KHORASANI K. Dynamic Neural Network-Based Fault Diagnosis of Gas Turbine Engines[J]. Neurocomputing, 2014, 125(3): 153-165.
- [4] FAST M, ASSADI M. Development and Multi-Utility of an ANN Model for an Industrial Gas Turbine[J]. Applied Energy, 2009, 86(1): 29-37.
- [5] 潘鹏飞. 基于试飞数据的航空发动机状态监测与故障诊断[J/OL]. 推进技术. <https://doi.org/10.13675/j.cnki.tjjs.200707.2021-03-29>.
- [6] 吕超, 程弓, 刘云清. 基于 BDPCA 聚类算法的航空发动机故障数据标记[J]. 振动与冲击, 2020, 39(9): 35-41.
- [7] 杨学森, 杨桥, 张韦雅, 等. 某型涡扇发动机测量数据融合模型[J]. 航空动力学报, 2020, 35(3): 641-650.
- [8] GEORGE E, GWILYM M, GREGORY C. Time Series Analysis Forecasting and control[M]. Beijing: People's Posts and Telecommunications Press, 2005: 52-77.
- [9] WILMS I, BARBAGLIA L, CROVX C, et al. Multiclass vector auto-regressive models for multi-store sales data[J]. Journal of the royal statistical society series c-applied statistics, 2018, 67(2): 435-452.
- [10] 王凤芹, 徐廷学, 赵媛. 基于多源信息融合的飞行安全评估模型[J]. 兵工自动化, 2021(10): 52-55.