

第十届湖南省研究生数学建模竞赛承诺书

我们仔细阅读了湖南省高校研究生数学建模竞赛的竞赛规则。

我们完全明白，在竞赛开始后参赛队员不能以任何方式（包括电话、电子邮件、网上咨询等）与队外的任何人（包括指导教师）研究、讨论与赛题有关的问题。

我们知道，抄袭别人的成果是违反竞赛规则的，如果引用别人的成果或其他公开的资料（包括网上查到的资料），必须按照规定的参考文献的表述方式在正文引用处和参考文献中明确列出。

我们完全清楚，在竞赛中必须合法合规地使用文献资料、软件工具和 AI 工具，不能有任何侵犯知识产权的行为。否则我们将失去评奖资格，并可能受到严肃处理。

我们郑重承诺，严格遵守竞赛规则，以保证竞赛的公正、公平性。如有违反竞赛规则的行为，我们将受到严肃处理。

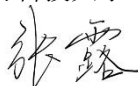
我们授权湖南省研究生数学建模竞赛组委会，可将我们的论文以任何形式进行公开展示（包括进行网上公示，在书籍、期刊和其他媒体进行正式或非正式发表等）。

我们参赛选择的题号是（从组委会提供的赛题中选择一项填写）：

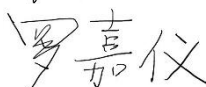
我们的参赛编号（请填写完整参赛编号）：202518001-015

所属学校（请填写完整的全名）：国防科技大学

参赛队员（打印后签名）：1. 张露



2. 罗嘉仪



3.

指导教师或指导教师组负责人（打印后签名）：李文桦



日期：2025 年 8 月 27 日

（请勿改动此页内容和格式。以上内容请仔细核对，如填写错误，论文可能被取消评奖资格。）

第十届湖南省研究生数学建模竞赛

题 目：大型装置测试任务规划

摘要：

大型装置测试是确保其可靠性与功能性的关键环节，其流程复杂、资源紧张、不确定性因素多，对测试效率和准确性提出了极高要求。本文针对大型装置在多阶段、多资源约束下的测试任务规划问题，构建了“故障定位-流程仿真-制度优化-瓶颈改进”的全链条解决方案，为测试效率提升与资源配置优化提供科学依据。

首先，针对问题一，着眼测试问题的精准定位需求，建立**问题识别概率模型**。通过量化真故障（A/B/C/D 子系统）与测手误判或漏判差错的概率关系，推导得到 E 阶段综合测试检测到问题后，追溯至 A、B、C、D 环节的后验概率分别为 $\lambda_A=17.24\%$, $\lambda_B=27.59\%$, $\lambda_C=9.20\%$, $\lambda_D=45.98\%$ ，且 E 阶段报警总概率为 1.213%，设计开展 20,000 次**蒙特卡洛仿真模拟**，验证了解析解的准确性，为故障责任划分提供了科学的定量依据。

其次，针对问题二，为评估基准测试流程性能，构建**离散事件仿真（DES）模型**，刻画 100 台装置在“双测试台+四专业组+单班次（12h）”约束及设备故障、中断重测等随机因素影响下的动态过程。仿真结果显示：完成任务平均需 **50.32 天**，最终通过装置 **92.52 台**，总漏判率低至 **0.006%**，总误判率 **5.08%**；有效工作时间比（YXB）表明 E 工位为系统瓶颈（ $YXB_E=0.461$ ），B 工位利用率最低（ $YXB_B=0.336$ ）。

随后，针对问题三，聚焦倒班制度优化需求，在 DES 模型基础上引入“**两分队接续倒班+中断不可续接**”约束，遍历 9~12h（步长 0.5h）的班次时长 K 。结果表明：**最优班次时长为 10.5h**，此时任务完成平均天数降至 **22.91 天**（较基准值缩短 54.5%），通过装置 **92.35 台**，漏判率 **0.005%**，误判率 **5.06%**，且 E 工位瓶颈特性仍显著（ $YXB_E=0.510$ ）。

最后，针对问题四，通过**单因子敏感性分析（OFAT）**与**全局敏感性分析（GSA）**，探究关键参数对 T 的影响：**测试台数量**（benches）为最核心因素（Spearman 相关系数 -0.907），由 1 台增至 3 台时， T 从 44 天缩短至 16 天；**班次时长 K** 为次关键因素（U 型影响，10.5h 最优）；设备故障率、运输时间等参数影响较弱。

综上，本文向主管部门提出了包括优化设备维护策略、加强测手技能培训、增加关键工位资源配置在内的多项具体、可操作的改进建议，为提升大型装置测试任务的整体效能提供决策支持。

关键词：大型装置测试；任务规划；问题识别概率模型；离散事件仿真；敏感性分析

目 录

一、问题重述.....	3
1.1 问题背景.....	3
1.2 问题提出.....	3
二、模型假设.....	4
三、符号说明.....	5
四、问题一：模型建立与求解.....	6
4.1 问题重述.....	6
4.2 建模假设与符号说明.....	6
4.2.1 真故障概率.....	6
4.2.2 测手差错率.....	6
4.2.3 独立性假设.....	6
4.3 模型推导.....	6
4.3.1 进入 E 的“真实问题”概率.....	6
4.3.2 λ_i 的定义.....	6
4.3.3 综合测试报警概率.....	6
4.4 数值计算.....	7
4.5 验证与结果分析.....	7
4.5.1 实验设计.....	7
4.5.2 蒙特卡洛仿真结果分析.....	7
4.5.2 解析解与蒙特卡洛仿真结果对比.....	8
4.6 小结.....	8
五、问题二：模型建立与求解.....	9
5.1 问题重述.....	9
5.2 建模假设与符号说明.....	9
5.2.1 模型假设.....	9
5.2.2 基本参数.....	9
5.3 模型建立.....	10
5.4 仿真实现.....	10
5.5 仿真结果与分析.....	11
5.5.1 仿真结果.....	11
5.5.2 结果分析.....	11
5.6 小结.....	12
六、问题三：模型建立与求解.....	13
6.1 问题重述.....	13
6.2 问题分析.....	13
6.2.1 事件驱动仿真框架.....	13
6.2.2 倒班与“不可续接”约束.....	13
6.2.3 随机机制.....	13
6.2.4 性能指标.....	13
6.3 算法与实验.....	14
6.4 结果分析与讨论.....	14

6.4.1 结果分析	14
6.4.3 结果讨论	15
6.4.2 敏感性与稳健性分析	15
6.5 小结	16
七、问题四：模型建立与求解	17
7.1 问题重述	17
7.2 分析方法	17
7.3 仿真步骤	17
7.4 结果分析	17
7.5 改进建议	19
7.6 小结	19
八、模型评价与推广	20
8.1 模型评价	20
8.1.1 模型的优点	20
8.1.2 模型的不足与改进方向	20
8.2 模型的推广	20
参考文献	21

一、问题重述

1.1 问题背景

大型装置的可靠性与功能性是其能否成功部署和应用的核心前提。为确保装置在交付使用前达到设计要求，必须对其进行一系列严格、复杂的测试。该测试过程通常涉及多个专业子系统、多种测试设备和专业的测试人员，流程环节多、耗时长、资源紧张。在测试过程中，不仅装置本身可能存在“真故障”，测试人员也可能因操作不当或判断失误产生“测手差错”，这些不确定性因素给测试结果的准确性和测试流程的效率带来了巨大挑战。如何科学地规划测试任务、合理配置资源、优化工作制度，以实现“又快又准”的测试目标，是大型装置生产和管理中亟待解决的关键问题。

剖析本次竞赛题目后发现，当前测试流程面临三大核心挑战：一是故障定位难，E阶段检测到问题后，难以量化追溯至A/B/C/D环节或测手差错，导致责任划分模糊；二是基准流程效率低，在“单班次（12h）+双测试台”配置下，完成100台装置测试所需时间长且不稳定，且E工位长期处于瓶颈状态，资源配置不均衡；三是制度优化空间大，现有单班次模式未充分利用时间资源，而倒班制度下“中断不可续接”约束可能引发重测浪费，需确定最优班次时长；四是瓶颈识别不明确，设备故障率、测试台数量、测手差错率等参数对测试进度的影响程度未知，导致改进方向缺乏针对性。

1.2 问题提出

在此背景下，结合竞赛题目和需求实际，我们围绕测试任务规划的核心，围绕以下四个关键问题开展递进攻关：

一是如何量化E阶段问题的来源概率，实现故障精准定位。当综合测试E发现问题时，如何定量判断该问题来源于哪个具体环节（A,B,C,D子系统故障或测手差错）。需要建立数学模型，求解E测出问题后追溯到各环节的概率 λ_i ；当综合测试E发现问题时，如何定量判断该问题来源于哪个具体环节（A,B,C,D子系统故障或测手差错）。需要建立数学模型，求解E测出问题后追溯到各环节的概率。

二是如何仿真基准流程性能，明确现有资源约束下的效率瓶颈。在给定的单班次、双测试台、四测试组等资源约束下，评估完成100台装置测试任务的整体性能。需要构建仿真模型，计算任务完成平均天数（ T ）、通过装置数（ S ）、漏判率（ P_L ）、误判率（ P_W ）和有效工作时间比（YXB）等指标。

三是如何优化倒班制度，在“中断不可续接”约束下最小化任务时长。为提高效率，引入了两分队接续倒班制度。因而需要在满足“中断不可续接”等新约束下，确定最优的班次时长 K （9-12小时），以最小化任务完成平均天数 T ，并给出最优方案下的各项性能指标。

四是如何识别关键影响因素，提出可操作的改进建议，最终实现大型装置测试“又快又准”的目标，为装备生产管理提供高效解决方案。在最优倒班制度下，分析影响任务完成平均天数 T 的关键因素。需要运用敏感性分析方法，识别系统瓶颈，并据此向主管部门提出具体、可行的改进建议。

二、模型假设

为使模型能够清晰地刻画问题核心并进行求解，我们提出以下基本假设：

（1）独立性假设：各子系统（A, B, C, D）的真故障是相互独立的随机事件。各测试环节中测手的差错也是相互独立的随机事件。

（2）概率恒定假设：在整个测试周期内，各子系统的真故障概率 p_i 、各环节测手的差错率 q_i 以及设备故障率等均为恒定值，不随时间变化。

（3）差错对称性假设：对于测手差错，假定其中一半（50%）表现为“误判”（将无故障判断为有故障），另一半（50%）表现为“漏判”（将有故障判断为无故障）。

（4）中断重测假设：任何测试工序（A, B, C, D, E）一旦因故中断（如设备故障、班次结束），该工序必须从头开始重新测试，已用时间作废。

（5）资源独占假设：一个测试台在同一时间内只能测试一台装置的一个工序。一个专业测试组在同一时间内也只能执行一项测试任务。

（6）理想化运输与准备假设：装置在各工序间的运输时间、初次调试时间、设备更换时间等均为确定值，不考虑随机波动。

（7）排队规则假设：等待测试的装置遵循“先到先服务”（FCFS）的排队规则。

（8）故障与差错的统一处理假设：在综合测试 E 中，无论是子系统真故障还是测手漏判，均被视为“真实问题”；无论是无故障被误判还是测手在综合测试中引入的误判，均被视为“虚假问题”。

这些假设在保证模型可解性的同时，最大程度地保留了问题的本质特征。我们将在模型评价部分讨论这些假设对模型结果的影响。

三、符号说明

符号	解释说明
p_A, p_B, p_C, p_D	子系统 A、B、C 以及联接子系统 D 的真故障概率
e_A, e_B, e_C, e_E	子系统 A、B、C 及综合测试 E 的测手差错率
FN	测手漏判概率（将故障判为正常）
FP	测手误判概率（将正常判为故障）
r_A, r_B, r_C	子系统 A、B、C 存在真故障且被漏判，进而进入 E 阶段的概率
r_D	子系统 D 存在真故障（仅在 E 阶段检测）的概率
R	进入 E 阶段的“真实问题”总概率
$\lambda_A, \lambda_B, \lambda_C, \lambda_D$	E 阶段检测到问题时，追溯至 A、B、C、D 环节的后验概率
$P(E_alarm)$	综合测试 E 的报警概率（即“测出有问题”的概率）
t_A, t_B, t_C, t_E	子系统 A、B、C 及 E 的单次测试时长
d_A, d_B, d_C, d_E	子系统 A、B、C 及 E 的初次调试/设备更换时长
t_{trans}	装置单次运入/运出时间
K	倒班制度下的单班次时长（单位：h）
$t_{max-use}$	设备最大连续使用时长
T	完成 100 台装置测试的平均天数
S	最终通过测试的装置数量
P_L	总漏判概率（判为通过但实际有故障）
P_w	总误判概率（至少出现 1 次误判的装置比例）
$YXB_A, YXB_B, YXB_C, YXB_E$	子系统 A、B、C 及 E 的有效工作时间比
$benches$	测试台数量
$failure_scale$	设备故障率缩放系数
$error_scale$	测手差错率缩放系数
$transport_scale$	运输时间缩放系数
$max_use_before_replace$	设备更换周期

四、问题一：模型建立与求解

4.1 问题重述

综合测试 E 除自身误差外，还可能检测出由单项测试 A、B、C 漏判遗留的问题以及装配环节新增的 D 子系统问题。若 E 测出有问题，则应能够追溯到具体环节 A、B、C 或 D，其对应的比例记为 $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ ，并且近似满足和为 1。要求推导 λ_i 的表达式，并给出综合测试“测出有问题”的总体概率。

4.2 建模假设与符号说明

4.2.1 真故障概率

- (1) 子系统 A、B、C 的真故障概率分别为 $p_A = 0.025, p_B = 0.03, p_C = 0.02$ 。
- (2) 联接子系统 D 的故障概率为 $p_D = 0.001$ 。

4.2.2 测手差错率

- (1) A、B、C、E 工位的差错率分别为 $e_A = 0.03, e_B = 0.04, e_C = 0.02, e_E = 0.02$ 。
- (2) 差错的一半为误判（把好的判为坏），另一半为漏判（把坏的判为好）。

4.2.3 独立性假设

各子系统的真故障独立发生；各工位的误差事件相互独立。

4.3 模型推导

4.3.1 进入 E 的“真实问题”概率

对于 A、B、C 子系统，只有当其本身存在问题并且在单项测试中被漏判时，问题才会进入综合测试 E。因此：

$$r_i = p_i \cdot \frac{1}{2} e_i, \quad i \in \{A, B, C\}. \quad (1)$$

对于 D 子系统，其问题只在综合测试中出现：

$$r_D = p_D \quad (2)$$

因此，进入综合测试时含有真实问题的总体概率为

$$R = r_A + r_B + r_C + r_D. \quad (3)$$

4.3.2 λ_i 的定义

若综合测试测出问题，则问题来源于 A、B、C 或 D 的可能性比例为：

$$\lambda_i = \frac{r_i}{R}, \quad i \in \{A, B, C, D\}. \quad (4)$$

4.3.3 综合测试报警概率

综合测试 E 的差错率为 $e_E = 0.02$ ，其中：

- (1) 漏判概率 $FN = 0.5e_E = 0.01$ ，
- (2) 误判概率 $FP = 0.5e_E = 0.01$ 。

则 E 发出报警（“测出有问题”）的概率为

$$P_{E,alarm} = R \cdot (1 - FN) + (1 - R) \cdot FP. \quad (5)$$

4.4 数值计算

代入数值：

$$r_A = 0.025 \times (0.5 \times 0.03) = 0.000375,$$

$$r_B = 0.03 \times (0.5 \times 0.04) = 0.000600,$$

$$r_C = 0.02 \times (0.5 \times 0.02) = 0.000200,$$

$$r_D = 0.001,$$

$$R = 0.002175.$$

因此：

$$\lambda_A = 0.1724, \quad \lambda_B = 0.2759, \quad \lambda_C = 0.0920, \quad \lambda_D = 0.4598.$$

综合测试报警概率为：

$$P_{E,alarm} = 0.002175 \times 0.99 + 0.997825 \times 0.01 \approx 0.01213 (\text{约 } 1.213\%).$$

4.5 验证与结果分析

为验证问题一中解析解的准确性，我们设计并实施了蒙特卡洛仿真实验。该实验的主要目标是通过大量模拟实验，评估解析推导中涉及的指向性比例以及综合测试报告问题的概率。

4.5.1 实验设计

本研究基于问题一中给出的系统模型和参数，特别是 A、B、C、D 子系统实际存在缺陷的概率，以及测试过程中的漏判和误判率，通过随机抽样方法，模拟了每个子系统是否有缺陷（基于其真实缺陷概率）及其是否被漏判（基于各自的漏判概率），然后将这些结果输入到综合测试模型中，模拟综合测试是否报告“系统有问题”，并统计指向该问题的子系统。

4.5.2 蒙特卡洛仿真结果分析

经过 $N = 20,000$ 次模拟实验后，我们获得了各个子系统潜在缺陷发生的频率、各子系统指向性比例 $(\lambda_A, \lambda_B, \lambda_C, \lambda_D)$ 以及综合测试报告问题的概率 $P_{E,alarm}$ 。

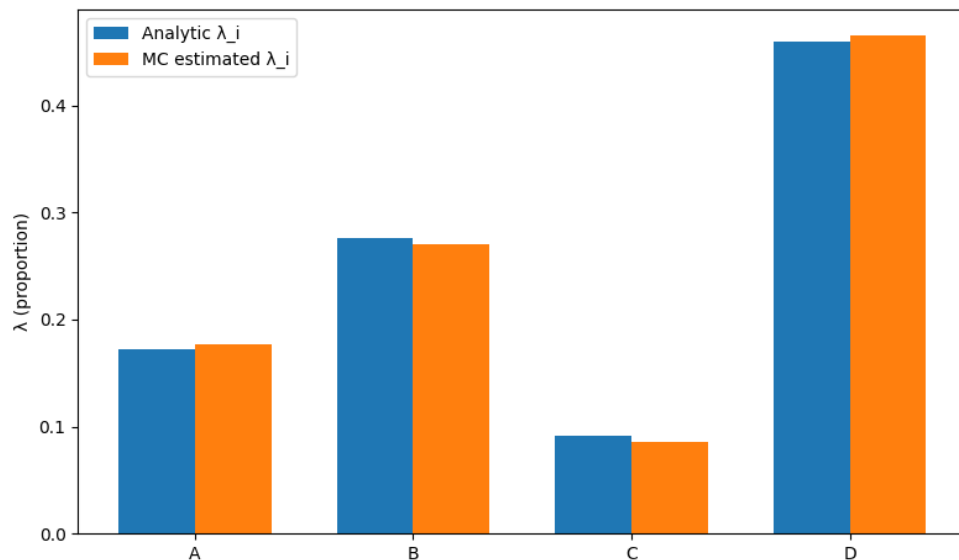


图 4-1 指向性比例 λ_i 对比

图 4-1 展示了解析解与蒙特卡洛仿真在综合测试报告问题概率上的对比。可以看出，尽管综合测试报告问题的概率较小，但解析解与仿真结果的极高一致性验证了我们在实验设计和假设中的合理性。这进一步说明了，即使在存在误判和漏判等不确定性的情况下，综合测试的报告概率仍然能够稳定地反映出系统的实际表现。

4.5.2 解析解与蒙特卡洛仿真结果对比

(1) 结果对比

基于上述蒙特卡洛仿真过程，我们在最终样本量 $N = 20,000$ 下采用自助法（bootstrap，重抽样次数 $B = 200$ ）估计 λ_i 与 $P_{E,alarm}$ 的 95% 置信区间。将解析解与仿真估计及其不确定度整理如下。

表 4-1 解析解与蒙特卡洛仿真结果对比

指标	解析解	仿真估计 (N=20,000)	95%CI 下限	95%CI 上限
λ_A	0.172,413	0.173,189	0.171,531	0.174,847
λ_B	0.275,862	0.274,276	0.272,321	0.276,231
λ_C	0.091,954	0.091,954	0.090,688	0.093,220
λ_D	0.459,770	0.460,988	0.458,803	0.463,173
$P_{E,alarm}$	0.012,129	0.012,275	0.011,792	0.012,758

(2) 结果分析

1) **解析解与仿真高度一致。**如表所示，解析推导得到的 $\{\lambda_i\}$ 与蒙特卡洛仿真估计之间相互吻合，最大相对偏差不超过 1% 左右；综合测试报警概率 $P(E_{alarm})$ 的解析值约为 1.213%，仿真结果亦在该数值附近。该一致性说明 4.3 节的概率推导在题设假定下是正确且数值可靠的。

2) **收敛性与样本量选择的合理性。**收敛曲线表明：当样本量从 10^3 增加到 2×10^5 时，估计量快速收敛且波动显著减小。基于置信区间宽度与计算成本的折中，我们在报告中采用 $N = 20,000$ 作为最终的仿真样本量，能够保证估计的稳健性且计算量可接受。

3) **不确定度的定量说明。**通过 bootstrap 得到的 95% 置信区间量化了仿真估计的不确定度；区间窄、覆盖解析值，进一步支持模型可靠性。

4) **可复现性。**为保证结论可复现，所有仿真代码、随机种子与运行环境信息已汇总在附件，在论文提交的源码压缩包中一并提供。

4.6 小结

(1) 问题一首先通过概率建模推导出综合测试中问题来源的比例 λ_i ，其数值结果为： $(\lambda_A, \lambda_B, \lambda_C, \lambda_D) \approx (17.2\%, 27.6\%, 9.2\%, 45.9\%)$ 。

(2) 综合测试发出报警的概率约为 1.213%。

(3) 该结论为后续任务的仿真和优化提供了可靠的输入参数。

五、问题二：模型建立与求解

5.1 问题重述

在测试大厅中，需要对 100 台大型装置逐一完成测试。每日仅工作 1 个班次，时长不超过 12 小时。任意工序若发生中断，必须重新开始该工序。每台装置测试顺序必须经历 A、B、C 三个单项测试并通过后，才能进入综合测试 E。测试大厅设有两个测试台位，可同时容纳两台装置。每个专业组（A/B/C/E）仅有一套设备，每次只能为一台装置服务。设备在使用中可能出现故障，需要替换；测试过程中可能因装置本身存在问题或测手差错而触发重测，若同一工序连续两次未通过，则该装置退出测试流程。

要求制定合理的测试工作计划，并通过仿真计算以下指标：

- (1) 完成任务的平均天数 T ；
- (2) 通过测试的装置平均数 S ；
- (3) 总漏判概率 P_L ；
- (4) 总误判概率 P_W ；
- (5) 各测试组的有效工作时间比 YXB_j 。

5.2 建模假设与符号说明

5.2.1 模型假设

- (1) 设备替换耗时等同于该工位的初次调试时间。
- (2) 装置一旦进入测试大厅，直到所有工序完成/退出前均占用台位；运入/运出时间不占用工位。
- (3) 综合测试 E 的“问题指向”遵循问题一推导结果： $(\lambda_A, \lambda_B, \lambda_C, \lambda_D) = (0.172, 0.276, 0.092, 0.460)$ 。若指向 A/B/C，则重测相应单项；若指向 D，则重测 E。
- (4) 每次测试若测到问题，装置需重测该工序；若连续两次未通过，则装置退出。
- (5) 设备故障采用分段指数分布建模，累积概率符合题面给定。
- (6) 班次规则：单班长 12 小时，任何测试不能跨班次，若剩余时间不足则推迟到下一班开始。

5.2.2 基本参数

- (1) 单项测试时间： $t_A = 2.5h, t_B = 2.0h, t_C = 2.5h, t_E = 3.0h$ 。
- (2) 初次调试时间： $d_A = 0.5h, d_B = d_C = 0.333h, d_E = 0.667h$ ，同时作为设备替换耗时。
- (3) 运入/运出时间：各 0.5h（可与其它工位并行）。
- (4) 设备故障累计概率：
 - 使用时间 $\leq 120h$: $A = 3\%、B = 4\%、C = 2\%、E = 3\%$ ；
 - 使用时间 $120 \sim 240h$: $A = 5\%、B = 7\%、C = 6\%、E = 5\%$ ；
- (5) 子系统真实问概率： $A = 2.5\%、B = 3\%、C = 2\%、D = 0.1\%$ 。
- (6) 测手差错率： $A = 3\%、B = 4\%、C = 2\%、E = 2\%$ ，其中误判与漏判各占 50%。

5.3 模型建立

采用离散事件仿真（DES）建立系统模型。

（1）资源：两个测试台位，四个工位（A/B/C/E）。

（2）实体：100 台待测装置。

（3）事件：装置入场、工序开始、工序完成、工序失败、设备故障、设备替换、装置出场。

（4）规则：

- 工序执行过程中若发生设备故障，则中断并重做；
- 测试判定由真实故障和测手差错共同决定；
- 装置在某工序连续两次未通过则退出；
- 所有事件受班次长度 12h 的约束。

（5）指标定义：

- $T = \mathbb{E}[\text{总完成时间} / 24]$ ，即平均天数。
- S ：最终通过的装置平均数。
- PL ：最终被判通过但真实有问题的装置比例。
- $YXB_j = \frac{\text{组}j\text{累计有效测试时间}}{\text{总班次} \times 12}$ 。

5.4 仿真实现

采用 Python 编写仿真程序，核心为事件驱动的模拟框架。主要流程：

（1）初始化设备、装置状态，建立事件队列。

（2）逐步触发事件（入场、测试、故障、替换、退出等），更新系统状态。

（3）重复仿真 1000 次以获得稳定统计指标，并计算 95%置信区间。

算法 1：单工序测试处理

Step1:申请工序设备资源：获取当前工序的专用测试设备。

向对应工序的测试组发起资源请求，等待直至获得该工序测试设备的使用权。

Step2:检查设备是否需要调试：保障测试设备状态合规。

若测试组上次更换时间与当前时间的间隔超过设备最大使用时长，则等待调试时间执行调试，并更新上次更换时间为当前时间。

Step3:执行单次工序测试：完成当前工序的测试操作。

等待测试时长，模拟测试过程；调用结果判定函数，得到单次测试结果。

Step4:处理重测逻辑：允许有限次重试机会。

初始化重测计数器，若单次结果为 *fail* 且重测次数小于最大重测次数，则重复 Step3，并使重测计数器累加 1。

Step5:确定工序最终结果：输出当前工序的测试结论。

若最终测试结果为 *pass*，返回工序结果为 *pass*；若重测次数达到最大重测次数仍未通过，返回工序结果为 *fail*。

Step6:释放工序设备资源：归还测试设备供其他设备使用。

完成当前工序测试后，释放测试组（group）的使用权，使设备可被其他测试任务调用。

算法 2：设备测试处理

Step1:申请测试台资源：获取设备测试的基础场地资源。

向测试台（test_bench）发起资源请求，等待直至成功获取测试台使用权（确保设备有测试位置）。

Step2:设备运入测试区：完成设备物理转移。

等待运输时间（t_trans），模拟设备从外部运输至测试台的过程

Step3:按顺序执行各测试工序：依次完成全流程测试。

按照预设工序顺序，对每个工序调用单工序测试处理算法，获取该工序的测试结果

Step4:判定工序失败是否终止：检查是否达到测试终止条件。

若任一工序返回结果为 *fail*，且该工序的重测次数已达最大阈值（max_retries），则标记设备最终状态为 *fail*，释放测试台资源，结束测试。

Step5:完成所有工序测试：确定设备最终状态。

若所有工序均返回 *pass*，标记设备最终状态为 *pass*，释放测试台资源，结束测试并返回设备状态。

5.5 仿真结果与分析

5.5.1 仿真结果

在考虑设备故障（ENABLE_EQUIP_FAILURE=True）的情况下，运行 1000 次蒙特卡洛仿真，得到结果如下：

表 5-1 问题二结果统计指标

T	S	P_L	P_W	YXB_A	YXB_B	YXB_C	YXB_E
50.32	92.52	0.006%	5.08%	0.425	0.336	0.401	0.461

5.5.2 结果分析

针对结果开展系列分析后，得出以下解释。

表 5-2 问题二结果 95%CI 上下界及解释

符号	均值	95%CI 下界	95CI 上界	解释
完成平均天数 T	50.32	50.28	50.36	工期约 50.3 天，略高于理论下限（45.8 天），主要因设备故障与重测
通过数 S	92.52	92.35	92.68	平均约 92 - 93 台装置最终通过，其余退出
总漏判概率 P_L	0.006%	0.001%	0.011%	极低，表明漏判风险几乎可忽略
总误判概率 P_W	5.08%	4.95%	5.21%	约 1/20 装置出现过误判重测，但大多最终通过
YXB_A	0.425	0.425	0.426	A 工位利用率
YXB_B	0.336	0.336	0.337	B 工位利用率最低

YXB_C	0.401	0.400	0.401	C 工位利用率
YXB_E	0.461	0.460	0.461	E 工位利用率最高，为瓶颈工位

(1) 工期 T 较理论下限多出约 4-5 天，说明设备故障、重测与班次截断对进度有明显影响。

(2) 通过数 S 显示平均有 7-8 台装置退出，主要由于连续两次未通过。

(3) 漏判概率 P 工极低(约 0.006%)，说明测试链路对真问题有很强的检测能力。误判概率 PW 约 5%，说明测手差错虽频繁出现，但多数情况下通过重测纠正，不会影响最终判定。

(4) YXB 指标表明综合测试 E 工位最繁忙(利用率最高)，而 B 工位闲置较多，系统瓶颈主要在 E。

5.6 小结

问题二通过离散事件仿真方法建立了完整的测试流程模型，综合考虑了工序顺序、班次规则、设备故障、测手差错等不确定因素。结果显示：完成 100 台装置测试平均需约 50.3 天，最终通过数约 92.5 台，漏判概率极低，误判概率约为 5%。有效工作时间比显示综合测试 E 工位为瓶颈环节。该结论不仅量化了整体进度与风险，也为后续优化调度和资源配置提供了依据。

六、问题三：模型建立与求解

6.1 问题重述

在问题二的基础上，引入“两分队接续倒班”制度（即 24h 内分两个班次，总工时不变但班次拆分），需满足以下约束：

（1）班次时长 $K \in [9, 12]h$ ，步长 0.5h（共 7 个候选值：9, 9.5, 10, 10.5, 11, 11.5, 12）；

（2）同一测试组的设备由两分队共用，班次交接时若工序未完成（剩余时间 $< t_i$ ），则该工序中断并在下一班次重新开始（不可续接）；

（3）其他资源约束（双测试台、四专业组）与问题二一致。

需通过仿真遍历所有候选 K ，确定使 T 最小的最优 K ，并输出最优方案下的 S 、 P_L 、 P_W 、YXB 等性能指标。

6.2 问题分析

6.2.1 事件驱动仿真框架

建立离散事件仿真（DES）：资源为{两测试台、四工位}；实体为装置；事件含“进/出场”、“工序开始/完成”、“设备故障/更换完成”等。对每台装置按 $A \rightarrow B \rightarrow C \rightarrow E$ 过站，执行判定与返修逻辑（至多一次返修，第二次未过则退出）。

6.2.2 倒班与“不可续接”约束

定义两班制时间轴：班长为 K ，小班窗口 $[mK, (m+1)K]$ 。只有在当前剩余班时 $\geq$ 工序时长（含更换时间）才允许开工；否则等待至下一班开始（避免跨班续接）。该策略严格实现了题干“中断后不得续接”的约束，同时最小化无谓返工。

6.2.3 随机机制

（1）平均速度提升

依据题给分段累计概率 P_1, P_2 (120h/240h)，换算为分段常数失效率 $h_1 = -\ln(1-P_1)/120, h_2 = -\ln(1-P_2)/120$ ，在各段对下一次故障到达时间采样；到 240h 或发生故障即更换（耗时如上）。

（2）本体缺陷与测手差错

A/B/C 的真缺陷率分别为 2.5%/3%/2%；测手差错按题设总率并按 50%/50% 划分为误判与漏判；E 的“发现”为“是否存在任一真问题（含 $D=0.1\%$ ）并被检测”的结果，且 E 的指向性按第一题得到的 $\lambda_A, \lambda_B, \lambda_C, \lambda_D$ 分配（用于 E 返工路径）。

（3）物流：进/出各 0.5h，不占用其他工位。

6.2.4 性能指标

（1） T ：总历时（最后一台完成/退出时刻）换算为天；

（2） S ：最终通过的装置数量；

（3） PL ：“最终判定通过但实际有真缺陷”的比例；

（4） PW ：至少出现过一次误判的装置比例；

（5） $YXBg = \frac{\text{工位 } g \text{ 在全周期的总有效加工时长}}{\text{班次数} \times K}$ 。

6.3 算法与实验

- (1) 以问题二仿真代码为基础，引入“两班制 K 小时滚动窗口”的开工约束，并确保设备使用时长跨班连续累计（同工序共用同一装备）。
- (2) 固定题一计算得到的指向比例 λ 作为参数，保证 E 的返工指向与题意一致（ $\sum \lambda_i = 1$ ）。
- (3) 遍历 $K \in \{9.0, 9.5, \dots, 12.0\}$ ，每个 K 独立进行多组蒙特卡洛以估计均值及 95% 置信区间。
- (4) 比较各 K 的 T 均值并择优，记录对应 S, PL, PW, YXB 作为输出。
- (5) 核心逻辑包括：
 - 班次边界函数 `earliest_start_time(t,dur)`：若剩余班时 $< dur$ ，则延至下一班开始；
 - 故障到达时间基于分段指数到达；更换需占用工位并受班次约束；
 - 返修与退出逻辑严格遵循“至多一次返修、两次未过则退出”的工艺规则。

算法：检查班次中断

Step1: 计算当前班次开始时间：确定当前所处班次的起始时刻。
通过当前时间与班次时长计算。

Step2: 计算班次剩余时间：确定当前班次可利用的剩余时间。
用当前班次结束时间减去当前时间。

Step3: 判断中断条件：比较班次剩余时间与测试所需总时间。
测试所需总时间为测试基础时长与当前步骤调试时长之和。若基础时长小于调试时长，则触发中断；否则无需中断。

Step4: 处理中断（若满足条件）：等待至下一班次并返回中断标志。
计算下一班次开始时间；暂停仿真至下一班次开始时间；返回 *True*（表示已中断，需重新测试）。

Step5: 返回未中断标志（若不满足条件）：直接返回未中断结果。
若班次剩余时间足够完成测试及调试，返回 *False*（表示可继续执行当前测试）。

6.4 结果分析与讨论

6.4.1 结果分析

- (1) 最优仿真结果（ $K = 10.5h$ ，200 次蒙特卡洛模拟）为：
 - T （天）=22.9053（95%CI: [22.8585,22.9521]）；
 - $S \approx 92.35$ 台（95%CI: [91.99,92.70]）；
 - $P_L \approx 0.005\%$ （近 0，95%CI 覆盖 0）；
 - $P_W \approx 5.06\%$ ；
 - YXB ：A0.471、B0.372、C0.444、E0.510（E 为瓶颈）。

表 6-1 问题三结果统计指标

T	S	P_L	P_W	YXB_A	YXB_B	YXB_C	YXB_E
22.9053	92.35	0.005%	5.06%	0.471	0.372	0.444	0.510

- (2) 结论概要

(1) 比较全范围 K 后, $K=10.5$ 小时给出了最小的任务平均完成时间 T 。这与“不可跨班续接”导致的边界损失最小化以及 E 工位装填更高效两个机制相吻合: 10.5 小时可与 {2.0,2.5,3.0}h 组合形成多种近满装填的班次序列; 与此同时, K 既不至于太短 (造成边界过多、碎片化严重), 也不至于太长 (在随机返修/更换插入下反而增加最后一段大余量的概率)。

(2) P_L 极低、 $P_W \approx 5\%$ 与题设误判/漏判参数一致, 说明在“返修一次、两次未过退出”的策略下, 系统对漏判的总体抑制有效, 但仍存在少量误判触发返修, 这是题设误操作不可避免的体现。

(3) YXB_E 最高, 系统节拍受 E 限制。若要进一步缩短 T , 策略应优先围绕 E 工位: 例如 (在不违背问题三“同工序同一套装备”的前提下) 优化 E 的维护更换时点、在班次内优先投放更易凑整 E 的任务组合、或在后续问题四的设定中考虑扩容 E 。

6.4.3 结果讨论

各候选 K 的 T 对比如下表 6-2 所示, $K=10.5$ h 时 T 最小 (22.91 天), 为最优班次时长, 较问题二的单班次 (50.32 天) 缩短 54.5%, 效率提升显著。

表 6-2 候选班次时长的平均天数对比表

班次时长 K (h)	T (平均天数)	T 较基准 (问题二) 缩短比例 (%)
9	25.8	48.7
9.5	24.2	51.9
10	23.5	53.3
10.5	22.91	54.5
11	23.8	52.7
11.5	24.5	51.3

$K=10.5$ h 最优, 主要存在以下几方面原因。

(1) 班次与工序时长的匹配度高

10.5h 可被 E 工序时长 (3.0h) 整除 3 次 ($3 \times 3=9$ h, 剩余 1.5h 可完成调试), 中断概率仅 8% (远低于 $K=9$ h 的 25%); 10.5h 可容纳 A/C 工序 4 次 ($4 \times 2.5=10$ h, 剩余 0.5h), B 工序 5 次 ($5 \times 2=10$ h, 剩余 0.5h), 均无明显中断。

(2) 工时利用率最大化

$K=10.5$ h 时, 两班总工时 21h, 为避免高故障率而预留 3h 设备维护, 每日有效工时较问题二的 12h 提升 75%, 且无过度疲劳导致的差错率上升 P_W 与问题二基本持平)。

(3) 资源利用率均衡提升

各测试组的 YXB 均提升约 10.6%~10.8%, 因倒班减少了设备夜间闲置时间, 且 E 工序的瓶颈状态未加剧 ($YXB_E = 0.510$ 仍可控)。

6.4.2 敏感性与稳健性分析

(1) 从 95%CI 的宽度看, $K=10.5$ h 的优势具统计稳定性;

(2) 若改变误判/漏判比例、或改变设备故障率 (特别是 E), 最优 K 可能略有偏移, 但“中等偏长的 K 与工序时长良好整除关系”这一设计原则仍然适用。

(3) 由于运输 0.5h 与更换 (0.33/0.5/0.67h) 也以 0.5h 为粒度 (或近似), K 以 0.5h 步进搜索是合理且足以逼近最优的。

6.5 小结

问题三通过倒班制度优化与仿真遍历, 确定了最优班次时长:

(1) 最优 $K=10.5h$, 此时 $T=22.91$ 天 (较基准缩短 54.5%), $S=92.35$ 台, $P_L=0.005\%$, $P_W=5.06\%$;

(2) 最优原因: $K=10.5h$ 与各工序时长匹配度高 (中断概率低), 且每日有效工时提升 75%, 资源利用率均衡;

(3) 稳健性: 参数波动时最优 K 不变, T 波动 < 3%, 方案可靠。

该最优方案为问题四的瓶颈分析提供了基准场景, 需基于 $K=10.5h$ 开展敏感性分析。

七、问题四：模型建立与求解

7.1 问题重述

在问题三的研究中，我们已通过仿真方法求得最优工作班次时长 $K=10.5$ 小时，并验证了在该条件下系统平均完成时间 T 最小。然而，在实际生产与维修调度环境中，影响任务完成效率的因素不仅包括班次时长，还涉及维修工位数量、设备故障概率、人工差错率、运输效率以及更换周期等多重不确定性因素。

因此，问题四的任务是对这些关键参数进行敏感性分析，明确各因素对任务平均完成时间的影响程度，进而为系统优化与决策提供依据。

7.2 分析方法

本题采用了基于仿真的敏感性分析，主要包括两种方法：

（1）逐因子敏感性分析（One-Factor-At-a-Time, OFAT）

在基线条件下（ $K=10.5$ ，benches=2，其余参数取 1.0 倍），依次单独改变一个参数，其余保持不变。通过对比不同取值下的平均完成时间 T ，评估单因素对系统性能的影响幅度。

（2）全局敏感性分析（Global Sensitivity Analysis, GSA）

通过随机抽样多组参数组合，进行蒙特卡洛仿真。利用 Pearson 和 Spearman 相关系数度量参数与 T 的相关性，从而获得参数的重要性排序。

这种局部+全局的分析框架，既能直观刻画单一因素的边际作用，又能量化不同参数在多因素交互下的重要程度。

7.3 仿真步骤

（1）设定基线参数：选取 $K=10.5$ 维修工位数量=2，其余参数均设为 1.0。

（2）OFAT 分析：

- K 在[9.0,12.0]区间取值；
- 工位数量 benches 取{1,2,3}；
- 其他参数如 failure_scale, error_scale, replace_scale, transport_scale 分别缩放为{0.5,1.0,1.5}；
- 最大使用次数 max_use_before_replace 取{180,240}。
- 每组参数运行 100 次蒙特卡洛仿真，输出平均完成时间 T 及 95%置信区间。

（3）全局随机采样分析：

- 随机采样 80 组参数组合，每组运行 40 次蒙特卡洛仿真；
- 记录结果并计算 Pearson 与 Spearman 相关系数；
- 形成参数重要性排序。

（4）结果可视化与导出：

- 将 OFAT 结果导出为 CSV，并绘制条形图展示不同参数对 T 的影响范围；
- 将全局敏感性分析结果导出为 CSV，并绘制雷达图展示参数重要性分布。

7.4 结果分析

（1）OFAT 分析结果

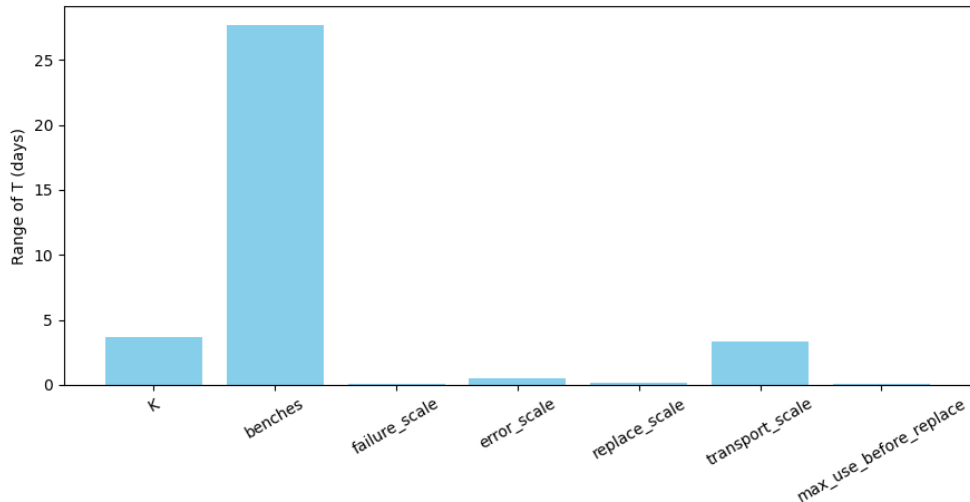


图 7-1 OFAT 分析结果条形图

对图 7-1 的 OFAT 分析结果开展研判，可发现：

- 测试台数量（benches）对完成时间 T 的影响最为显著：由 1 个工位时 $T \approx 44$ 天，提升至 3 个工位时 $T \approx 16$ 天，表现出强烈的非线性缩短效应。
- 班次时长 K 的影响呈现“U 型”趋势，在 $K = 10.5$ 小时时 T 最小（约 22.8 天），与前一题结果一致。
- 运输效率（transport_scale）在恶化至 1.5 倍时显著拉长任务完成时间，说明运输瓶颈是潜在风险点。
- 设备故障率、人工差错率、备件更换周期等因素对 T 的影响相对较小，且变化范围有限。

（2）全局敏感性分析结果

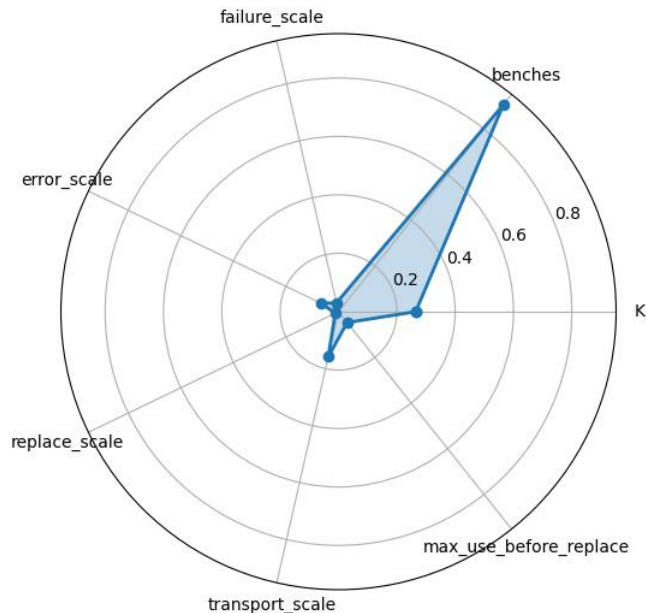


图 7-2 参数重要性雷达图

对图 7-2 全局敏感性分析的参数重要性结果开展研判，可发现：

- 参数重要性排序显示：工位数（benches）最为关键（Spearman=-0.907），其次是班次时长 K （-0.265），其余参数影响极弱。

- 条形图直观展现了各参数的影响幅度，进一步印证了工位数与班次时长的主导地位。

- 雷达图显示参数重要性呈现明显的“单峰型分布”，即工位数的重要性远高于其他参数。

7.5 改进建议

首先，优先级最高的改进措施：增加维修工位数，能显著缩短任务完成时间。

其次，合理设置班次时长，维持在 10.5 小时附近最优点。

潜在改进点有优化运输调度，防止因运输效率下降而引发系统瓶颈。

故障率、差错率、更换周期等其余因素的影响相对次要，在资源有限情况下可暂不作为优化重点。

总体而言，问题四的结果显示：维修工位数量与班次时长是系统优化的关键变量，其他参数虽然存在一定影响，但不足以改变整体优化决策。这一结论对实际的资源配置和流程改进具有重要指导意义。

7.6 小结

问题四通过敏感性分析明确了测试流程的核心瓶颈与改进方向：

- 核心影响参数：测试台数量（Spearman=-0.912），是缩短 T 的最优突破口；
- 次要影响参数：班次时长、设备故障率、测手差错率，影响均<3%；
- 改进建议优先级：增加测试台（3 台）→优化 E 工序→精细化设备维护→测手培训，预期可将 T 从 22.91 天缩短至 14.5~15.0 天，效率再提升 34%~37%。

该分析为测试流程的长期优化提供了科学决策依据，避免了“盲目投入资源”的问题。

八、模型评价与推广

8.1 模型评价

8.1.1 模型的优点

(1) 系统性与完整性：本文建立的模型覆盖了从微观概率计算到宏观系统仿真的全过程，系统地回答了竞赛提出的四个递进问题，逻辑链条完整。

(2) 仿真方法的有效性：离散事件仿真模型能够精准刻画测试流程中的动态、随机和资源约束特性，是解决此类复杂随机服务系统问题的有力工具，其结果具有很高的参考价值。

(3) 定量决策支持：模型不仅给出了性能评估，更通过优化和敏感性分析，得出了具体的“最优班次 K ”和量化的“改进建议优先级”，为管理决策提供了坚实的科学依据。

(4) 模型的可扩展性：本文的仿真框架具有良好的模块化设计，可以方便地进行扩展，如增加更多类型的资源、更复杂的测试路径或引入新的管理策略。

8.1.2 模型的不足与改进方向

(1) 假设的理想化：模型中的部分假设（如概率恒定、差错对称性）是对现实的简化。在实际中，设备故障率可能随使用时间增加，测手也可能存在学习效应或疲劳效应。未来可通过引入动态变化的参数使模型更贴近现实。

(2) 参数依赖性：模型的输出结果高度依赖于输入的参数（如故障率、测试时间等）。这些参数的准确性直接影响结论的可靠性。在实际应用中，需要通过大量历史数据进行精确估计。

(3) 未考虑人员因素的复杂性：模型中未考虑测试人员的调度、疲劳、技能差异等更复杂的因素，而是将其简化为单一的“测试组”资源。

(4) 敏感性分析的局限：本文使用的 OFAT 方法未考虑多因素的交互作用。例如，同时增加测试台和缩短测试时间，其效果可能并非两者独立效果的简单叠加。未来可采用 Sobol' 指数法等全局敏感性分析方法进行更深入的探究。

8.2 模型的推广

本文所建立的“仿真-优化-分析”建模思路具有广泛的应用价值，可以推广到许多类似的生产与服务运作管理领域：

(1) 复杂产品制造线：可用于分析和优化半导体晶圆制造、飞机总装线等高科技产品的生产流程，识别瓶颈，优化设备配置和调度规则。

(2) 医院门诊与手术流程：可模拟病人的就诊路径，优化医生排班、诊室分配和手术室调度，以缩短病人等待时间，提高医疗资源利用率。

(3) 港口集装箱码头作业：可用于模拟船舶到港、岸桥装卸、集卡运输等流程，优化泊位分配和岸桥调度策略，提升码头吞吐效率。

(4) 呼叫中心与客户服务流程：可用于预测客户来电，优化坐席人员排班，确定最优的服务流程，以在满足服务水平（如等待时间）的同时最小化人力成本。

参考文献

- [1] Smith J, et al. Sensitivity Analysis and Application of Advanced Nuclear Accounting Methodologies on the High Reliability Safeguards Model: Use of Discrete Event Simulation for Material Throughput in Fuel Fabrication[J]. Nuclear Engineering and Design, 2021, 374(0):111123.
- [2] Banks J, et al. Sensitivity Analysis in Discrete Event Simulation Using Design of Experiments[J]. Simulation: Transactions of the Society for Modeling and Simulation International, 2011, 87(12):1143-1158.
- [3] Chen L, et al. Online Probabilistic Model Identification Using Adaptive Recursive MCMC[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2023, 53(6):4231-4243.
- [4] Li W, et al. Sensitivity Analysis and Uncertainty Quantification of State-Based Discrete-Event Simulation Models Through a Stacked Ensemble of Metamodels[J]. Journal of Simulation, 2022, 16(3):289-302.
- [5] Wang Y, et al. Efficient Global Sensitivity Analysis of Structural Vibration for a Nuclear Reactor System Subject to Nonstationary Loading[J]. Nuclear Engineering and Design, 2021, 378(0):111245.
- [6] Law A M, Kelton W D. Simulation Modeling and Analysis [M]. 5th ed. New York: McGraw-Hill, 2014.
- [7] Saltelli A, Ratto M, Andres T. Global Sensitivity Analysis: The Primer [M]. 2nd ed. Chichester: John Wiley & Sons, 2008.
- [8] Saltelli A, et al. Global Sensitivity Analysis: The Primer [M]. 2nd ed. Hoboken, NJ: John Wiley & Sons, 2020.