我在写一个国家级挑战杯竞赛项目的论文，本次竞赛我们选的题目的主办方是中国南方航空股份有限公司，题目名称是基于多源数据的航班态势推演与智能恢复，我负责的内容简单概括是，基于主办方提供的ZGGG广州白云机场的2025年5月的起降数据，基于前文其他合作者（文中不要提到“合作者”，认为是同一团队的共同成果）的对于机场航班的态势推演和预测结果，进行机场起飞排队仿真，仿真结果和真实情况对比，推演效果良好。后续还加入了因恶劣天气或其他特殊因素导致的机场航班停飞、塔台效率降低的情况，根据仿真结果分析这些特殊情况对机场积压情况的影响。 以上是情况简介。

 我负责的部分，统一设定为一个模型：未来24小时航班运行状态推演模型。

涉及两个评审指标： 指标2:出港积压发生时段偏移误差。计算 出港积压 发生时段偏移误差，即模型推演的积压发生时段（广州机场出港延误航班量首次超过10班的时段）和积压化解时段（延误航班量最后一次降到10班以下的时段）与实际的偏差；积压发生到化解时段称为“积压时段”，验证以下4个子项： 1）推演的积压时段：与实际前后偏差不超过1个时段

如预测10点-15点，实际为11点-16点，则合格 2）推演的积压时段持续时长：与实际一致 3）积压最高峰：发生时段和积压量（所在时段的原计划航班量+前面时段因超容延误到这个时段的航班量）要与实际一致，最高峰航班积压量与实际偏差不高于15%。 4）推演的航班最晚运行时段：与实际要一致

说明：以上计算仅需细化到小时，按时段进行推演

指标3:停止起降情景下出港积压化解偏移误差。1、情景设定：demo程序提供手动设定停止起降时段功能，项目组需任意设定广州机场停止起降时段（2个时段以内）。

2、停止起降时段：可通过提供的1个月航班表，查找7点到23点之间发生过停止起降时段，作为设定时段；航班实际时刻作为计算实际积压时段计算依据。

3、根据情景设定推演航班演进延误情况，计算出港积压化解偏移误差，积压时段定义及计算参考指标2，需验证的4个子项与指标2保持一致。

在指标3这里，需要注意的是，没有原始数据并没有提供明确的停飞时间段，所以无法去比对 设置了停飞时段的仿真结果和真实停飞情况下的积压情况差异，所以这个部分主要是，基于指标2对于正常情况对积压预测推演仿真结果良好，然后成果实现设置停飞时间段的仿真，利用仿真结果给实际情况提供一定的技术指导和建议。  我需要你帮忙根据这些提供的内容，生成相关部分的段落和论文内容，行文专业，阐述明晰，用语精准，有需要公式或者图表的地方也直接生成（如果无法生成那么也给出一个对该图表的描述，然后我来生成）。其中指标2属于3.2章节，指标3属于3.3章节。

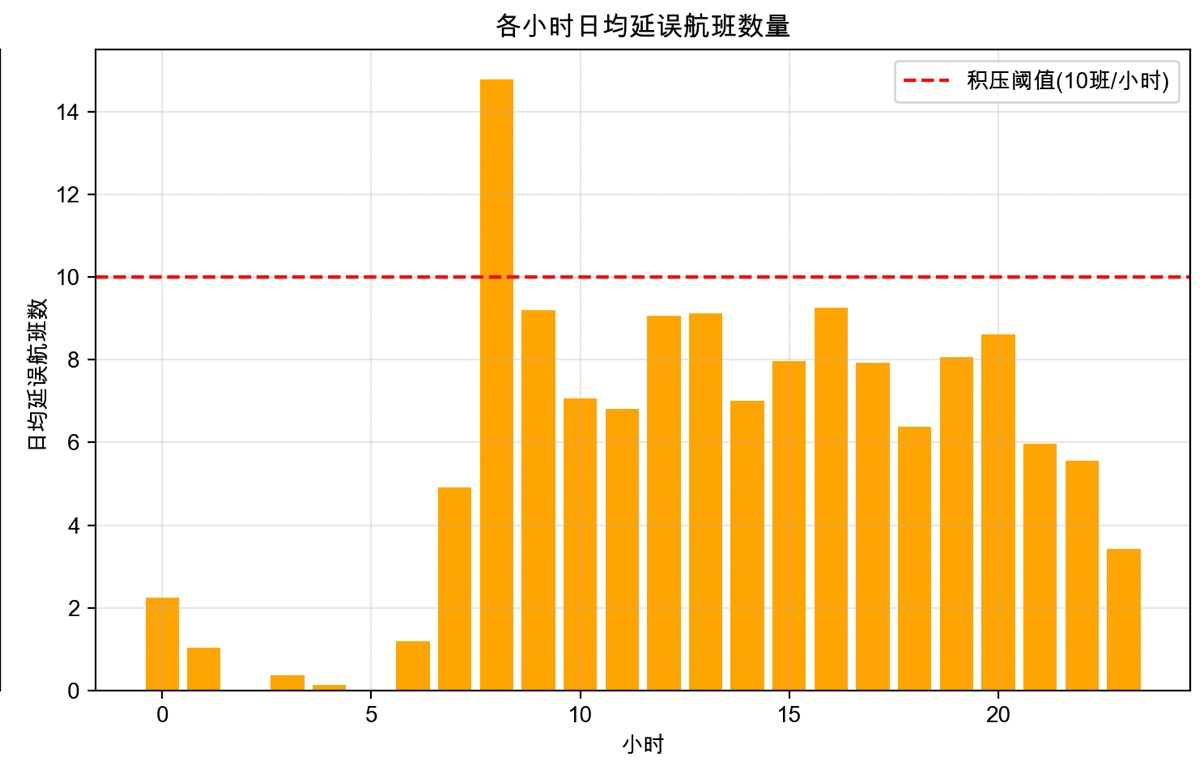
 接下来是我的工作内容的一些详细记录，以供参考（其中内容因为进度不断更新，所以前后文可能有更新迭代的地方，你要在参考同时确定一个最终的论文上展示的仿真思路，线路明确）。

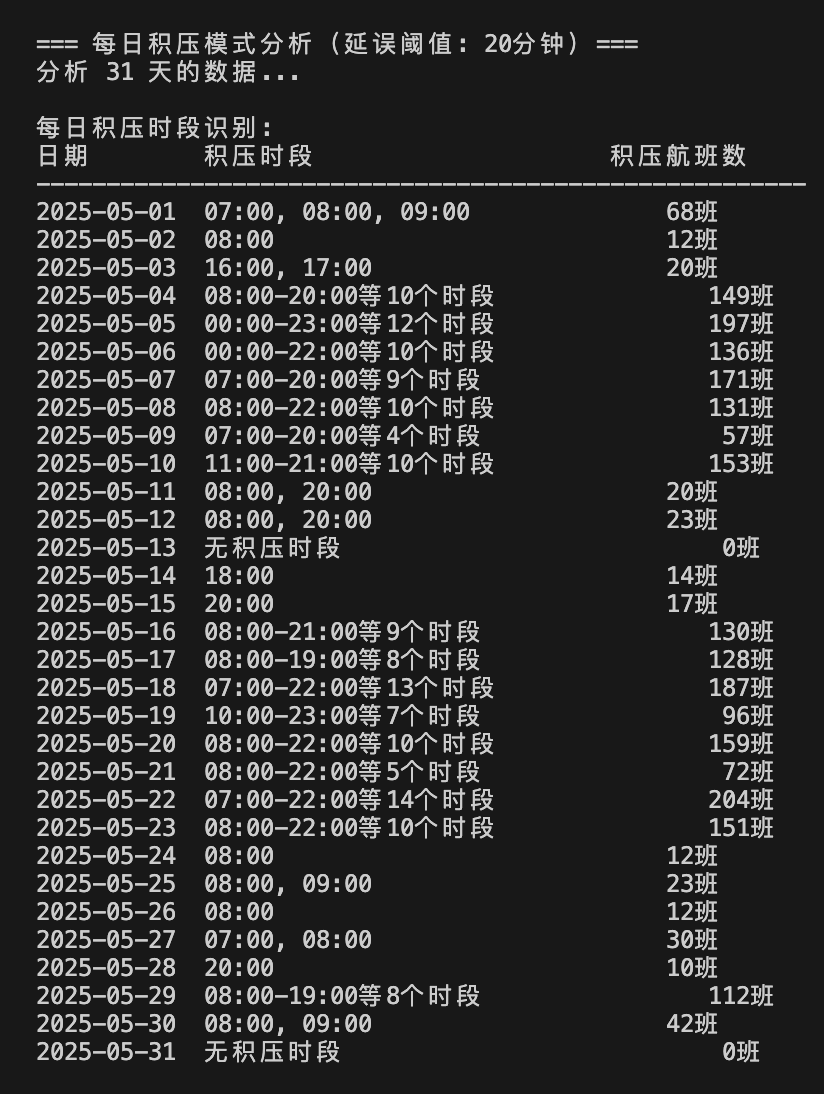
航站航班数据分析：

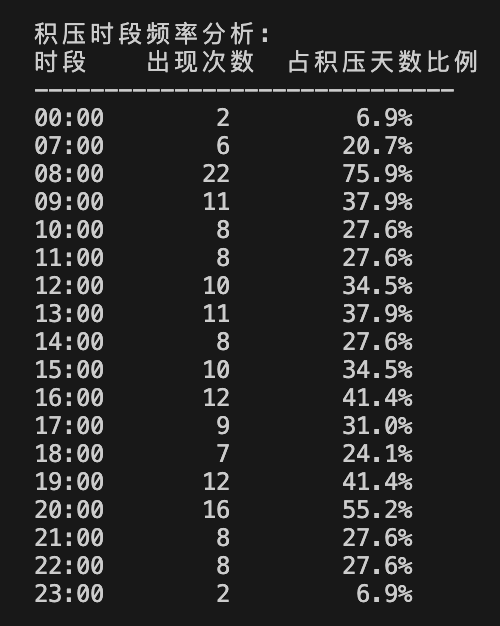
优化积压情况的识别

先不看仿真输出，只看实际航班的数据。如果实际情况中，每一天的航班能识别出：恶劣天气导致的停飞、异常延误的情况，然后识别出每个时段延误的航班数量，发现由此得出的积压时段符合理想情况，而不是现在的，从早上七点到晚上十一点一直是积压。那么，这个判定航班延误、判定积压时段的标准，也就能直接套用到仿真输出结果里了。

但是数据分析识别结果：







有积压的天数: 29/31 天 (93.5%)

平均每天积压时段数: 5.9 个

平均每天积压航班数: 87.4 班

机场数据分析

官方提供的航班数据里，给了计划离港、实际离港、实际起飞，三个时间点，

以及暂时还没用过的三个指标，原计划离港时间、调时标识、取消时间，这三个不知道有没有用

实际离港肯定是早于实际起飞的，计划离港则有早有晚，不好捉摸。

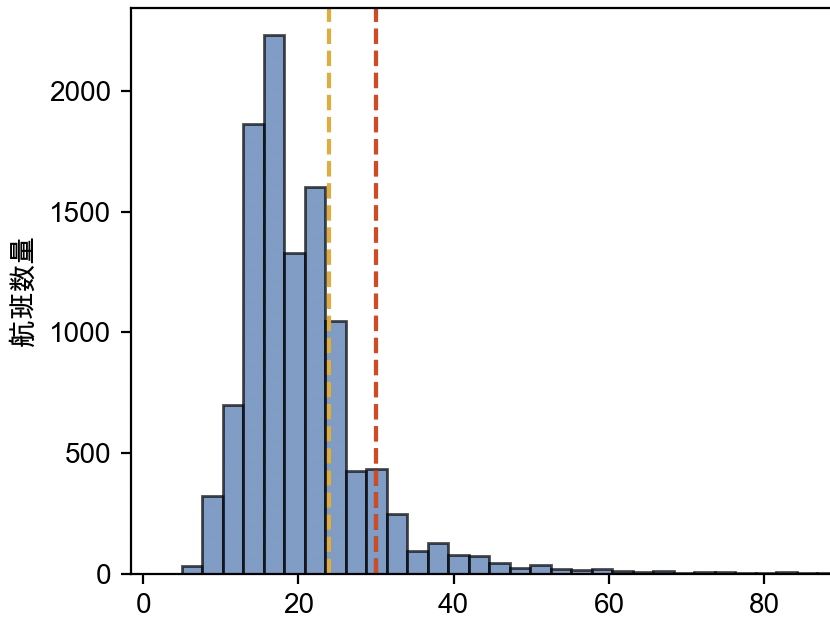
现在考虑的是，直接观察实际离港和实际起飞之间等待时间长的，说明耗时久？但是挑取部分数据，发现也并不是这样，理论上的高峰时段，这两个实际时间之间的差距也没有很大，就是平时范围内（注意我说的这些都不一定准确，如果有需要可以去验证或者忽略）

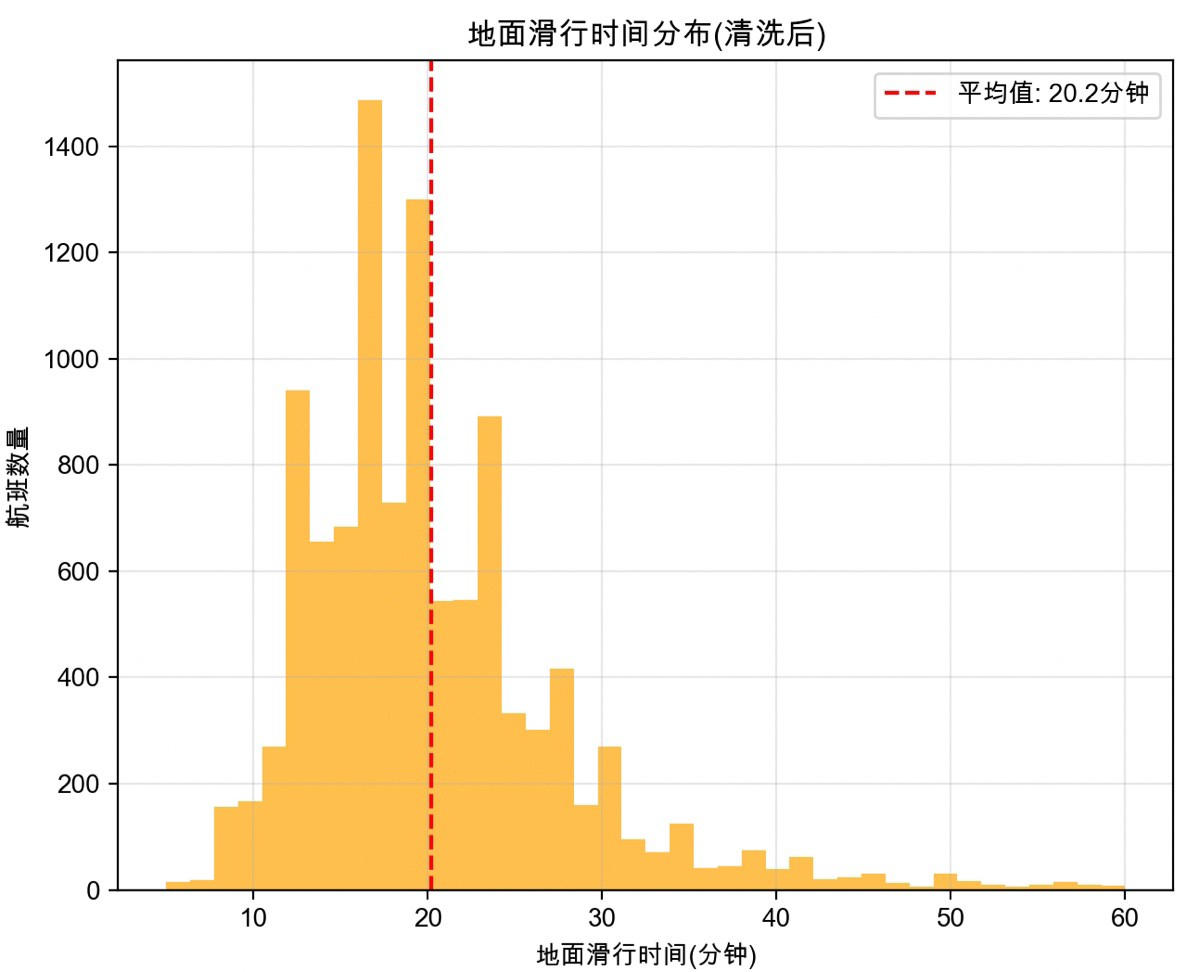
或者是，按照计划离港的时间，在考虑延误的问题？比如计划十点整离港，结果实际离港是十点十五分，实际起飞十点半，于是认为是延误？

在想怎么根据这些数据来判定，某一航班是否是延误的呢？

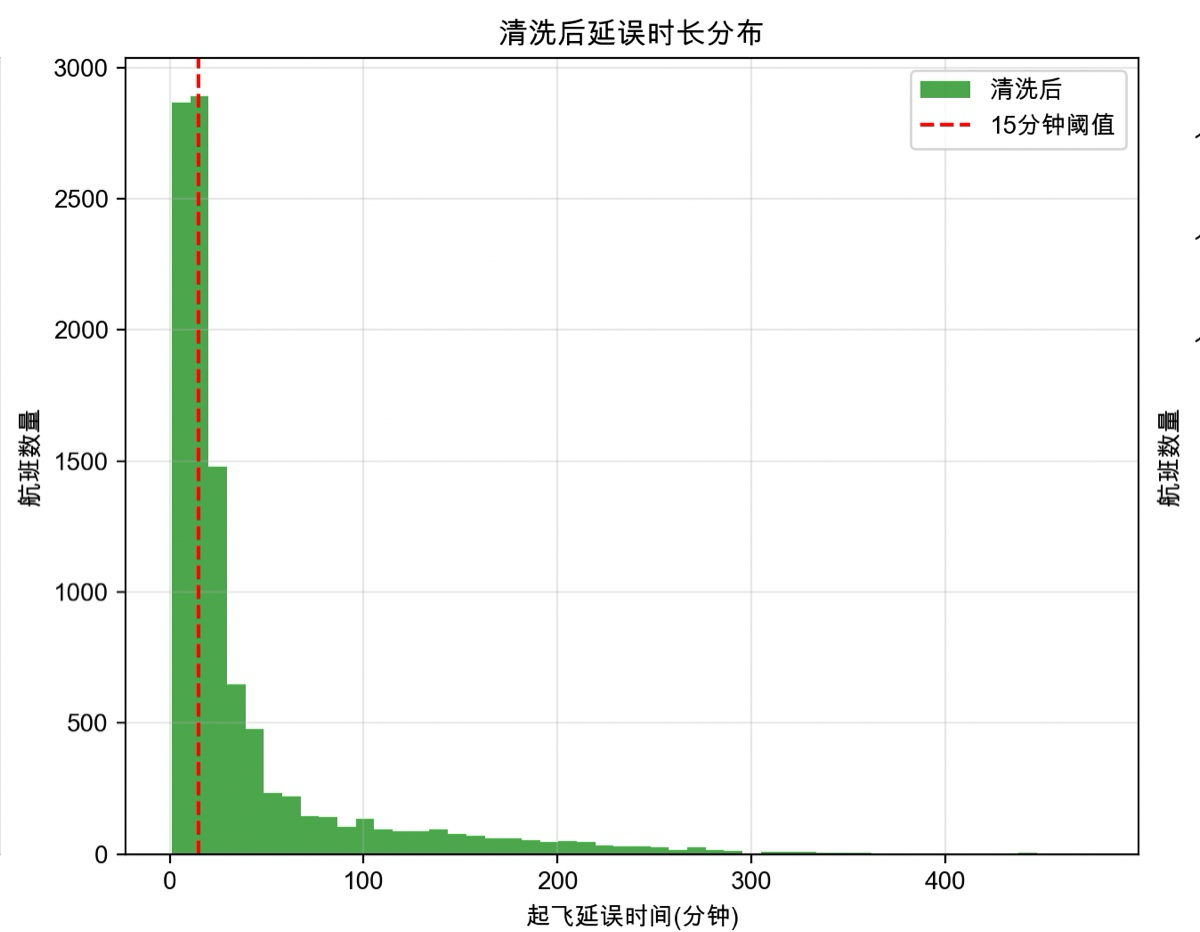
（然后就能仿真中捕捉到超过10班延误的，积压开始时段了）

这是实际离港-实际起飞之间的时长-航班数量统计





离港到起飞的时长：



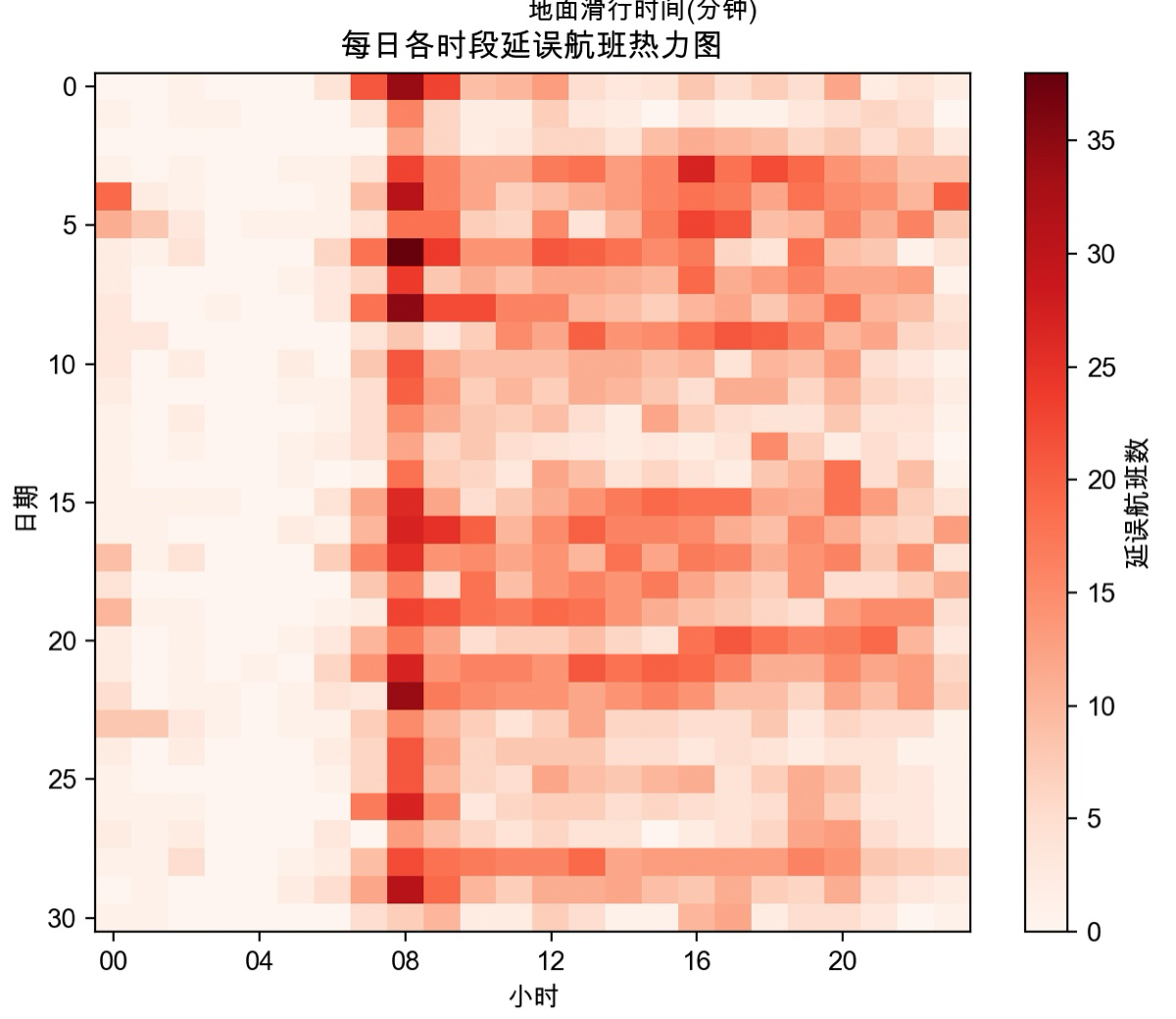
想到：时长延误过多的，就是因为恶劣天气延迟起飞的。这种情况下，就按照这个航班实际离港时间，的附近时刻，认为是恶劣天气结束了，期间的积压航班和之后的计划航班都开始起飞，于是恶劣天气结束之后的这个时段，如果积压数量超过10，可以算作积压时段，等到恢复正常了，积压时段消失。

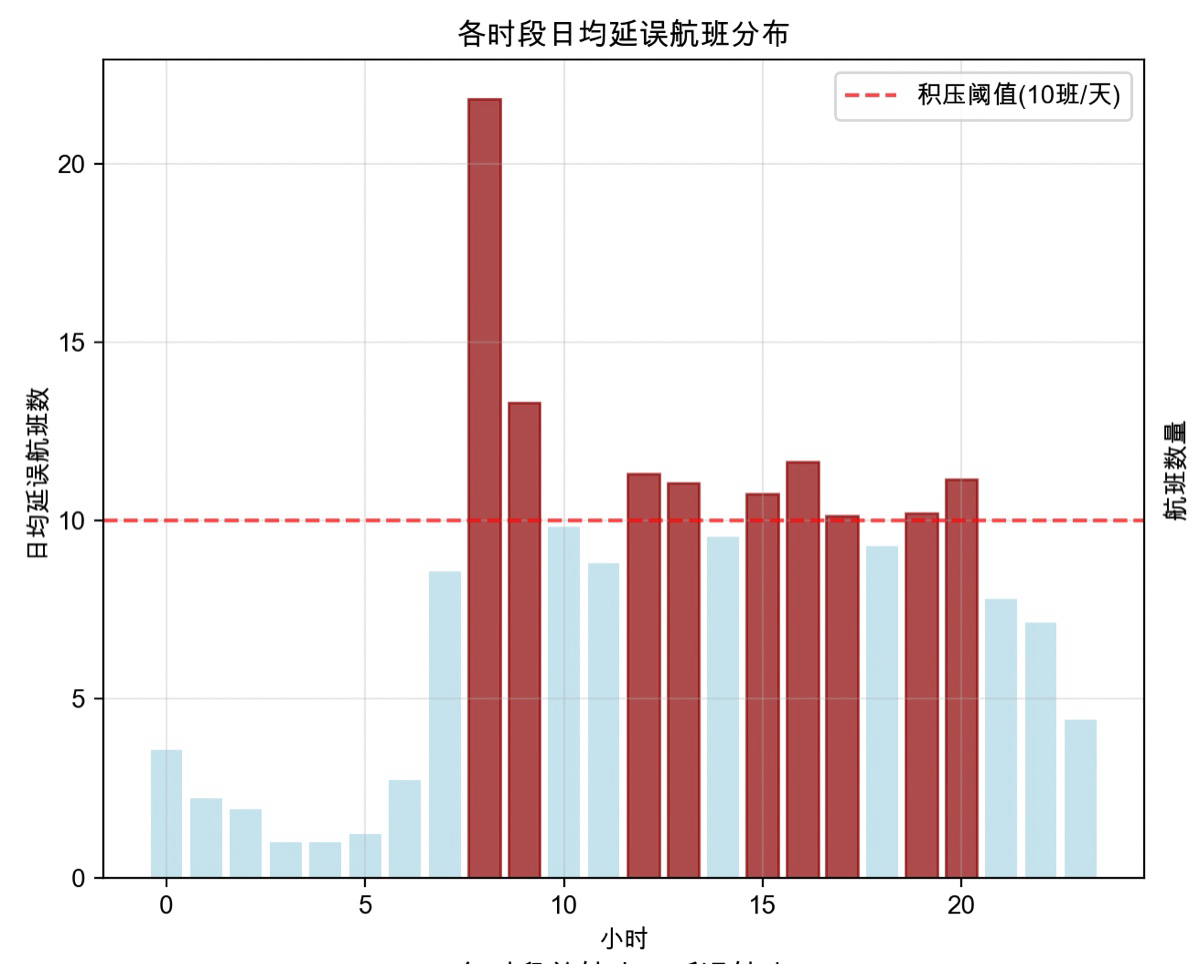
于是，这样的停飞航班的延误时长，应该从停飞时段停止后开始算起（恶劣环境停飞时段要自己识别）（停飞不一定是所有飞机都不能飞）。

then，识别出恶劣天气对应时段（该时段多个航班未按计划起飞且延误时间过长，之后在大致同一时段起飞），确定停飞时段和重新开放起飞时段。

在开放起飞的时间点，进行是否存在积压航班的判断（超过10班）

设置每日和日均的延误情况，





机场出港起飞仿真相关内容：

ZGGG机场跑道背景调查

目前有4条跑道，第五条在建

网络调查显示日均飞机起降1546架次，数据分析结果基本无误

目前不具备三跑并行能力，最多只能两架飞机同时起降

不应只看起飞、降落的出入港耗时（taxi-in/out），还应该参考另一个指标ROT跑道占用时间

比方说，飞机在10:00:00出港，耗时20分钟，是和塔台对接、脱离廊桥、测试发动机、滑行到跑道入口等待起飞等等的流程，此时并没有开始上跑道。10:20:00出港完毕，开始起飞，进入ROT时段（几十秒），maybe 10:21:00起飞完毕，如果不考虑尾流干扰等话，10:00:00-10:20:00，10:21:00之后，都是可以有飞机直接起飞/降落的

仿真因素设定

首先是ROT，飞机起降，跑道占用耗时

除了ROT之外，还要考虑的是尾流干扰（起飞1min的rot之后，等待气流稳定的时间，重中轻分别963-0秒）【非独立跑道需加入冲突约束时间，算在尾流干扰的这段不能起飞的时间里】~~、滑行跑道的占用排队情况（理解：飞机脱离廊桥后滑行到跑道入口，平均需要一分钟时间，如果排队等待更久，好吧这个算在taxi-in/out里就行~~

然后附带一个“固定”的起飞之前的出港时长taxi-out/降落之后的入港时长taxi-in，18min

然后按照计划离港时刻，给航班飞机们排队，仿真起飞就行

一般情况仿真思路

进行ZGGG机场飞机起飞仿真。

读取脱敏xlsx文件中和ZGGG机场有关的全部起飞数据，即ZGGG作为实际出发航站的航班，注意不需要考虑ZGGG作为抵达航站的降落航班。读取起飞航班的计划离港时间&实际起飞时间

这里注意，理论上，离港时间应早于起飞时间

（这里可以进行一波数据分析，实际起飞是否都早于计划离港，实际落地时间是不是基本早于计划到港时间，二者时间差距是不是大概二十分钟，可以考虑延误航班的标识

见下方，附1-机场数据分析

从ZGGG出发的航班可以分成两类。一类是正常按照计划执行的；另一类是由于恶劣天气等原因，航站针对受影响航班，设定停飞时段。等恶劣天气度过后，再去起飞（这样子的被延误了的航班，从停飞结束时刻开始，和原本就计划在这个时刻起飞的飞机一起，进入排队队列，按照原计划离港时刻先后进入队列。

所以需要先进行数据的分析，分出正常按计划飞行的航班，以及受恶劣天气影响停飞、而后再逐一起飞的航班，并识别确定停飞时间段（停飞时段结束后受影响航班开始起飞）

接下来说明积压相关问题，

积压发生时段，按照规则要求，就是ZGGG出港延误航班量首次超过10班的时段（按小时为单位），延误的识别标准设定成一个参数，不要固定成15分钟，超过了就是延误。后续可以调整判定延误的标准，来使仿真结果和实际结果贴近。

仿真的方案就按照正常的队列queue仿真即可。

先根据5月航班脱敏xlsx文件得到要用到的航班信息，包括航班号、机型、唯一序列号、实际离港时间。其他数据如果有需要也可以使用，但主要是前面分析恶劣天气、停飞时段、积压时候会用到，不确定接下来的仿真是否还要用到其他时间信息。

然后设定每次航班的rot（跑道占用时间）、尾流干扰等参数的值，可以根据机型进行大致的耗时分类

然后进行排队仿真，输出仿真结果。

接下来，根据仿真结果，和实际情况比对看是否相符【比对的内容是推演的积压时段以及持续时长是否和实际情况一致、航班积压量的最高峰和实际的积压最高峰的偏差是否在15%以内】，如果不一致，继续调整参数的值，包括rot、尾流干扰、延误的判定时长（可能要调整到比十五分钟更小，以出现更多的判定延误航班）

但现在存在的问题是，当前代码把真实和仿真情况中的，很多情况**都识别成了积压时段**。实际上不应该每个时段都在积压，也就是最初的积压判定，或许存在问题。

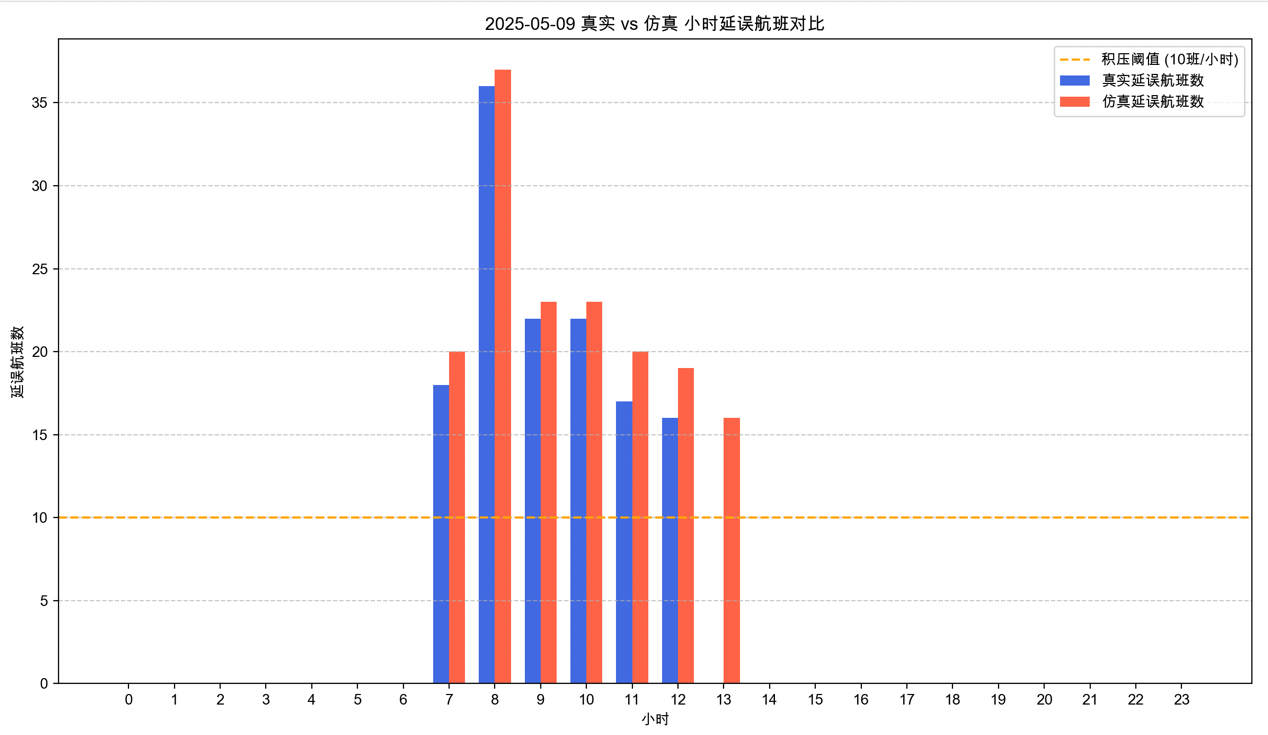
保留以往所有文件，后续的分析生成新的代码文件。

积压时段应该持续1-5小时为宜，从早上七点到后续晚上十一点一直积压，真实数据不应该识别成这样，然后仿真数据按照新的标准的话，或许也会有变化。

尝试调整积压时段的分析办法

规则指出，一个时段如果有10个以上的航班被积压，就会算成积压时段。

所以不应该去调高 判定进入积压时段的积压量的阈值，而是，或许应该调高判定单个航班是否被延误的阈值，来降低积压情况，减少全天积压时段数量。此外，无论是真实数据还是仿真数据，都不要用日均的做法，进行每日每个时段的判定，因为日均还是会有平均导致的模糊影响，可能有些天的16点是高峰，而有些天16点是流量少，结果一平均，两天都被认为刚刚高于积压判定的10个航班的情况。



继续优化仿真代码：

1、跑道排队仿真使用五月份的全部数据进行仿真，而不是只选某一天

2、官方建议判定延误的区间是15分钟以上，可以据此数据围绕15分钟，适当缩小或者放大时间，来优化仿真输出。

3、计算起飞延误分钟时候，有没有剔除过长延误的特殊情况（也就是因为恶劣天气导致的停飞的延误），对于这些恶劣天气的停飞航班，延误时长按照停飞结束状况恢复的那一刻，认为是计划离港，用实际起飞相减得到这类航班的起飞延误分钟。这样子期望起飞延误应该会小于当前的输出，50分钟。也就是把天气停飞的识别往前移。

4、在仿真结果统计的代码部分之后，继续加入新的模块。仿真数据和真实数据的对比分析模块。比对仿真输出和真实情况的“积压时段”的区间，左右端点的误差是否在1小时内；比对积压时段里的积压量峰值，差距是否在15%内；

在尝试将这个判定标准也化用到仿真中，但是输出结果不理想：因为仿真时候航班的排队延误太长了，平均七十到一百多分钟，而实际数据分析得到的平均延误是三四十分钟（这里面还包括了不少没有被识别成恶劣天气停飞，但实际上就是某些原因导致的停飞延误，迟了十多个小时——我们认为的延误应该是排队时候，计划起飞到实际起飞的延误差异，不考虑特殊情况导致的过长非排队延误）

但是这个仿真的问题还没能优化，暂时猜测是，排队耗时过长——跑道taxi-out时长设定十几分钟，而其他参数只有几分钟，影响不均衡。

后续尝试直接在排队耗时里去掉这个出港耗时，或许能有效改善排队时长。

以及在仿真中，也对那些恶劣天气和停飞时段加以考虑，来和真实数据实现更高的匹配程度，目前仿真直接按照航班计划进行排队，没有设置停飞时段。

以下是停飞时段相关互动内容的输入内容设定，后续先放在代码里，终端input，可以接到web页面中进行互动。

互动模块输入内容详细设计

基本设置

- 仿真日期范围

- 开始日期 (YYYY-MM-DD)

- 结束日期 (YYYY-MM-DD)

- 延误判定阈值 (分钟) [默认 15]

- 积压判定阈值 (班/小时) [默认 10]

航班计划数据

- 输入方式选择：

- 选项 A：使用现有 5 月数据（默认）

- 选项 B：上传 .xlsx 航班数据

停飞时段配置

- 逐一设置每个停飞事件（默认列表空，无停飞设置）：

- 停飞日期 (YYYY-MM-DD)

- 停飞开始时间 (HH:MM)

- 停飞结束时间 (HH:MM)

塔台效率配置

- 默认效率：1.0 (100%)

- 逐一设置每个低效时段（默认列表空，无低效时段设置）：

- 影响日期 (YYYY-MM-DD)

- 开始时间 (HH:MM)

- 结束时间 (HH:MM)

- 效率系数 (0.1–1.0)

- 影响类型（单选）：

- 随机延误（随机选择航班延误）

- 按顺序延误（按起飞顺序延误）

- 优先级延误（大型机优先，小型机延误）

把互动的input内容，设定排队各类参数、仿真起止日期、是否设置停飞时段和塔台低效等功能加载到最初指标显示良好的函数中

最终设定比较合理的方案是：延误阈值15分钟（这样真实延误的积压时段合理）、积压阈值固定为10班、滑行时间15分钟（此时仿真延误如图所示比较合理，如果变成16分钟，那么积压量就会过度上涨；如果设置14分，那么积压量则几乎没有）。重新调整代码，把以上数据设置为默认值。

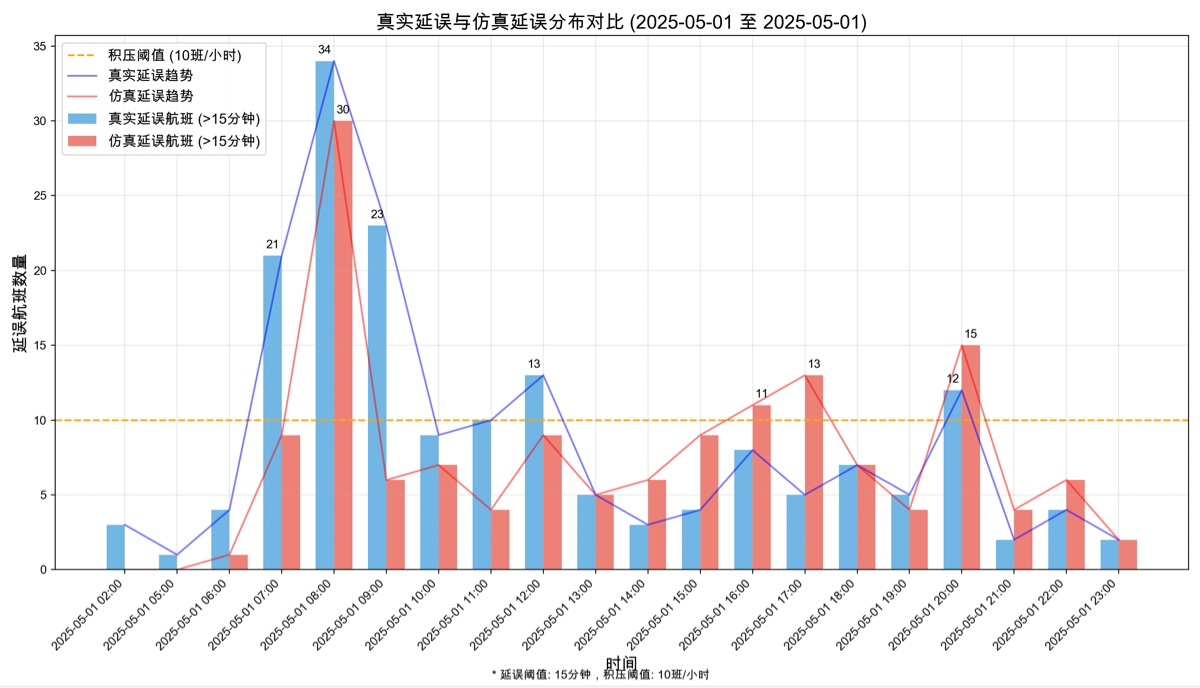
接下来处理方案：

根据实际数据可以得出航站在某日某时段的积压情况，然后分离出高峰、平峰还是低峰。得到一个日期&时段 对应 三类峰值的映射。

然后对仿真的不同日期时段分别设置不同的排队参数，来调整排队情况，使得仿真结果整体和实际结果拟合。

继续优化三个峰段的参数设置。

得到最终5月1日的仿真：



需要注意，最终的指标2，四个要求，应该是都成功达到了，你可以根据上文的步骤，详细给出仿真具体的步骤。以及对仿真结果的分析并说明指标达标。