

การวิจัยกลไกการจัดการความล่าช้าของระยะเวลาการก่อสร้างโครงการในภาคใต้ของจีน

— โดยอาศัยการวิเคราะห์แบบ data-driven

A Study on the Management Mechanisms of Construction Project Schedule Delays in

South China: A Data-Driven Analysis

华南地区建筑项目工期延误的管理机制研究——基于数据驱动分析

เต็ง จิงเหวิน Deng Jingwen 邓景文¹

เฉิน ชิวหลี่ Chen Hsiu-Li 陈绣里²

หลัว จวินกัง Luo Jungang 罗俊罡³

บทคัดย่อ

ความล่าช้าของกำหนดเวลาในโครงการก่อสร้างเป็นปัญหาการจัดการที่สำคัญซึ่งส่งผลกระทบต่อประสิทธิภาพของโครงการและความสามารถในการแข่งขันของบริษัทก่อสร้าง งานวิจัยเกี่ยวกับความล่าช้าในโครงการในอดีตมักอาศัยการตัดสินใจจากผู้เชี่ยวชาญหรือวิธีการทางสถิติแบบเชิงเส้น ซึ่งมีข้อจำกัดในการอธิบายความสัมพันธ์ที่ซับซ้อนของปัจจัยความเสี่ยงหลายประการ เพื่อเพิ่มประสิทธิภาพในการระบุและคาดการณ์ความเสี่ยงของความล่าช้า งานวิจัยนี้ได้พัฒนารอบการวิเคราะห์ตามแนวคิดของทฤษฎีระบบที่ซับซ้อน โดยพิจารณาปัจจัยสำคัญ 5 มิติ ได้แก่ ความซับซ้อนของโครงการและเทคโนโลยี โครงสร้างสัญญาและการกำกับดูแล ทรัพยากรและการจัดองค์กรก่อสร้าง ความไม่แน่นอนของสภาพแวดล้อมภายนอก และความสามารถด้านการบริหารจัดการ จากข้อมูลโครงการก่อสร้างอาคารพลเรือนจำนวน 126 โครงการ ได้สร้างแบบจำลองการทำนายระดับความล่าช้าโดยใช้เทคนิคการเรียนรู้ของเครื่องแบบ XGBoost และใช้การวิเคราะห์ความสำคัญของตัวแปรร่วมกับวิธี SHAP เพื่ออธิบายผลลัพธ์ของแบบจำลอง ผลการศึกษาพบว่าแบบจำลองการ

¹ นักศึกษาหลักสูตรบริหารธุรกิจมหาบัณฑิต สาขาบริหารธุรกิจ (หลักสูตรภาษาจีน) วิทยาลัยนานาชาติไทย-จีน เพื่อการจัดการมหาวิทยาลัยหอการค้าไทย , Master's Student in Master of Business Administration (Business Administration), Thai-Chinese International School Of Management, University of the Thai Chamber of Commerce , 泰国商会大学, 泰-中国际管理学院, 工商管理硕士研究生

² วิทยาลัยนานาชาติไทย-จีนเพื่อการจัดการ, มหาวิทยาลัยหอการค้าไทย, มหาวิทยาลัยหมิงฉวน , Thai-Chinese International School Of Management, University of the Thai Chamber of Commerce, Ming Chuan University, 泰-中国际管理学院, 泰国商会大学, 铭传大学

³ วิทยาลัยนานาชาติไทย-จีนเพื่อการจัดการ, มหาวิทยาลัยหอการค้าไทย, มหาวิทยาลัยหมิงฉวน , Thai-Chinese International School Of Management, University of the Thai Chamber of Commerce, Ming Chuan University, 泰-中国际管理学院, 泰国商会大学, 铭传大学

เรียนรู้ของเครื่องมีความแม่นยำสูงในการทำนายระดับความล่าช้า และสามารถระบุปัจจัยความเสี่ยงสำคัญได้อย่างมีประสิทธิภาพ โดยความถี่ของการเปลี่ยนแปลงแบบ ความเสถียรของการจัดหาวัสดุ ประสิทธิภาพการประสานงานของผู้รับเหมาช่วง และคุณภาพของการวางแผนกำหนดเวลา เป็นปัจจัยสำคัญที่มีผลต่อความเสี่ยงของความล่าช้า จากมุมมองด้านการจัดการ ความล่าช้าในโครงการไม่ได้เกิดจากปัญหาทางเทคนิคเพียงอย่างเดียว แต่ยังสะท้อนถึงความสามารถขององค์กรในด้านการจัดสรรทรัพยากร การประสานงานของห่วงโซ่อุปทาน และการกำกับดูแลโครงการ งานวิจัยนี้แสดงให้เห็นว่าการนำแนวทางที่ขับเคลื่อนด้วยข้อมูลมาใช้ในการบริหารโครงการก่อสร้างสามารถช่วยส่งเสริมการเปลี่ยนผ่านจากการตัดสินใจตามประสบการณ์ไปสู่การตัดสินใจเชิงอัจฉริยะ และเป็นแนวทางสำหรับการพัฒนาระบบเตือนภัยความเสี่ยงด้านความล่าช้าในโครงการ

คำสำคัญ: โครงการก่อสร้าง; ความเสี่ยงของความล่าช้า; การเรียนรู้ของเครื่อง; XGBoost; การจัดการกำหนดเวลาโครงการ

Abstract

Construction project schedule delay is a critical management issue that affects project performance and the competitiveness of construction enterprises. Traditional studies on project delays mainly rely on expert judgment or linear statistical methods, which are limited in capturing the complex interactions among multiple risk factors. To improve delay risk identification and prediction, this study develops a multidimensional analytical framework based on complex system theory. Five key dimensions are considered, including project and technical complexity, contract and governance structure, resources and construction organization, external environmental uncertainty, and managerial capability. Using data from 126 civil construction projects, an XGBoost machine learning model is established to predict delay levels. Feature importance analysis and SHAP methods are further employed to interpret the model results. The findings indicate that the machine learning model achieves high prediction accuracy and effectively identifies critical delay risk factors. Variables such as design change frequency, material supply stability, subcontractor coordination efficiency, and schedule planning quality have significant impacts on delay risks. From a managerial perspective, project delays are not only technical problems but also reflect the enterprise's capability in resource allocation, supply chain coordination, and project governance. The study demonstrates that integrating data-driven approaches into construction project management can facilitate the transition from experience-based decision-making to intelligent decision-making, providing practical insights for establishing delay risk early-warning systems and improving project management performance.

Keywords: Construction projects; Schedule delay risk; Machine learning; XGBoost; Project schedule management

摘要

建筑工程项目工期延误是影响项目绩效与企业竞争力的重要管理问题。传统延误研究多依赖经验判断或线性统计方法，难以有效刻画多因素耦合条件下的复杂风险结构。为提升工期延误风险识别与预测能力，本文基于复杂系统理论构建建筑工程项目工期延误影响因素框架，从项目与技术复杂性、合同与治理结构、资源与施工组织、外部环境不确定性以及管理能力五个维度建立指标体系。在此基础上，选取 126 个民用建筑工程项目样本数据，采用 XGBoost 机器学习算法构建工期延误多分类预测模型，并通过特征重要性分析与 SHAP 方法解释模型结果。研究表明：机器学习模型在延误等级预测中具有较高的准确性与稳定性，能够有效识别关键风险因素；其中设计变更频率、材料供应稳定性、分包协调效率及进度计划合理性对延误风险具有显著影响。从企业管理视角看，工期延误不仅源于施工技术问题，更与企业资源配置能力、供应链协同水平及项目治理结构密切相关。研究表明，将数据驱动方法引入工程项目管理，有助于实现由经验决策向智能决策的转变，为施工企业建立进度风险预警机制和优化管理决策提供了重要参考。

关键词：建筑工程项目；工期延误风险；机器学习；XGBoost；进度管理

1. 引言

1.1 研究背景

随着经济全球化与城市化进程的不断推进，建筑行业在推动社会经济发展和城市基础设施建设方面发挥着越来越重要的作用。近年来，我国建筑业保持了较快的发展速度，大量基础设施项目、住宅建设项目以及公共建筑项目陆续实施，建筑工程项目的规模和复杂程度不断提高（赵丽娜 & 王卓甫，2016）。与此同时，工程项目管理面临的挑战也日益增加，其中，工期延误问题成为影响工程项目顺利实施的重要因素之一。

工期延误是指工程项目未能在合同约定或计划安排的时间内完成施工任务的现象。大量研究表明，建筑工程项目延期不仅会导致项目成本增加，还会对施工企业的资源配置效率、合同履行能力

以及市场信誉产生负面影响 (Assaf & Al-Hejji, 2006)。在一些情况下,严重的工期延误甚至可能引发合同纠纷、经济索赔以及法律风险,从而进一步增加项目管理的复杂性 (Sambasivan & Soon, 2007)。因此,如何有效识别和控制建筑工程项目中的工期延误风险,已经成为工程管理领域关注的重要问题。

针对工期延误问题,国内外学者开展了大量研究。早期研究主要集中在识别影响工期的关键因素,通过问卷调查和案例研究等方法对不同地区建筑项目的延期原因进行分析 (Kaming et al., 1997)。研究表明,建筑项目工期延误通常由多种因素共同作用产生,包括设计变更、资金支付延迟、施工组织不当、材料供应不足以及不可预见的外部环境变化等 (Doloi et al., 2012)。此外,项目参与方之间沟通协调不足、合同管理不完善以及项目管理能力不足等管理因素也可能导致项目进度受到影响 (Chan & Kumaraswamy, 1997)。

随着工程项目复杂程度的不断提高,单一因素分析方法逐渐难以全面解释工期延误的形成机制。一些学者开始从系统视角对工程项目延误问题进行研究,强调不同风险因素之间的相互作用关系 (Flyvbjerg et al., 2003)。例如,有研究指出,项目延期往往并非由单一原因导致,而是由多个技术因素、管理因素以及外部环境因素共同作用形成的复杂系统问题 (Williams, 2005)。因此,在研究工期延误问题时,需要从多维度视角进行综合分析。

近年来,随着信息技术和数据分析方法的发展,数据驱动研究方法逐渐被引入工程管理领域。工程项目在实施过程中会产生大量与施工进度、资源配置以及项目管理相关的数据,这些数据为研究项目延期问题提供了新的研究基础 (Cheng et al., 2012)。机器学习等数据分析技术能够在复杂数据环境中识别变量之间的潜在关系,从而提高工程项目风险预测的准确性 (Son et al., 2019)。与传统统计方法相比,机器学习方法在处理多变量非线性关系方面具有明显优势,因此逐渐被应用于工程进度预测与风险管理研究 (Marzouk & El-Rasas, 2014)。

然而,现有研究在数据驱动分析方面仍存在一定不足。一方面,部分研究主要关注算法模型的预测精度,对预测结果在工程管理实践中的应用关注较少 (Kim et al., 2019)。另一方面,不同地区建筑项目在施工环境、管理模式以及市场条件等方面存在差异,因此延误因素及其影响程度也可能存在显著差别 (Doloi et al., 2012)。目前针对特定区域建筑项目的实证研究仍然相对有限,特别是在建筑活动高度集中的华南地区,相关研究仍有进一步深化的空间。

华南地区是我国经济发展最为活跃的区域之一,建筑工程项目数量多、类型丰富,工程建设活动十分频繁。在快速发展的背景下,施工企业在项目管理过程中面临着更加复杂的管理环境,包括项目规模扩大、技术复杂性提高以及资源协调难度增加等问题 (王卓甫, 2015)。在这种情况下,如何通过科学方法识别建筑项目工期延误的关键因素,并建立有效的风险预警机制,成为施工企业提

升项目管理水平的重要课题。

因此，从数据驱动视角对华南地区建筑项目工期延误问题进行研究，不仅能够丰富工程项目管理领域的相关理论，也能够为施工企业提升项目进度管理能力提供实践参考。

1.2 研究目的

基于上述研究背景，本文以华南地区建筑工程项目为研究对象，旨在通过数据驱动分析方法识别影响建筑项目工期延误的关键因素，并探索其在企业项目管理中的作用机制。具体而言，本文的研究目的主要体现在以下几个方面。

(1) 通过梳理现有文献，总结建筑工程项目工期延误的主要影响因素，并在此基础上构建适用于建筑项目延期风险分析的指标体系。通过对项目复杂性、合同与治理结构、资源配置、施工组织以及外部环境等多个维度进行综合分析，形成系统化的研究框架。

(2) 利用工程项目实际数据，采用数据驱动分析方法对建筑项目工期延误风险进行预测分析。通过建立预测模型，识别对项目进度影响较大的关键变量，从而提高工期延误风险识别的准确性和科学性。

(3) 在模型分析结果的基础上，从企业管理视角探讨工期延误形成的管理机制。通过分析不同因素在项目实施过程中的作用路径，揭示项目延期背后的管理问题，并为施工企业优化项目管理流程提供依据。

1.3 研究的意义

在理论层面上，本文从多维度视角构建建筑项目工期延误风险分析框架，将项目复杂性、管理能力以及外部环境等因素纳入统一分析体系，从而拓展了工程项目延期研究的分析视角(Williams, 2005)。同时，本文将数据驱动方法引入建筑项目工期延误研究，通过机器学习等数据分析技术识别关键影响因素，有助于丰富工程管理领域的的数据驱动研究方法。

此外，本研究结合企业管理视角，对工期延误问题进行分析，不仅关注技术执行层面的影响因素，还探讨了项目治理结构和管理能力在延误形成中的作用机制。这种综合分析方法有助于进一步深化对工程项目延期问题的理解。

在实践层面上，本研究通过对实际工程项目数据的分析，识别影响建筑项目工期延误的关键因素，并提出相应的管理启示。研究结果能够为施工企业在项目实施过程中进行进度风险识别提供参考，从而帮助企业提前采取管理措施，减少项目延期风险。

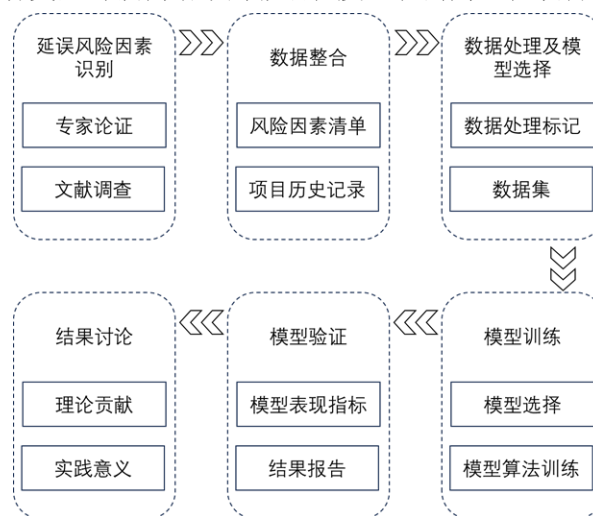
同时，本研究提出的工期延误风险分析框架和预测方法，也可以为企业建立项目进度预警系统提供基础。通过将数据分析结果应用于企业项目管理实践，施工企业能够更有效地进行资源配置和施工组织，从而提升整体项目管理水平。

2. 研究方法

2.1 研究框架

针对建筑工程项目工期延误问题的复杂性与多因素特征，本文构建了基于数据驱动分析的研究框架。研究整体思路包括指标体系构建、数据采集与预处理、模型构建以及结果解释四个主要步骤。

首先，通过梳理建筑项目工期延误相关研究文献，总结影响项目进度的主要因素，并结合建筑工程项目管理实践，构建工期延误风险指标体系。其次，收集华南地区建筑工程项目数据，对原始数据进行清洗与标准化处理，并完成指标变量编码。随后，基于数据驱动分析方法构建预测模型，以识别影响建筑项目工期延误的关键因素。最后，通过模型评估指标对模型预测效果进行验证，并利用模型解释方法分析各变量对预测结果的影响程度，从而为工程项目管理提供决策参考。



如图 1 所示，整体研究框架可概括为：因素识别 → 数据整合 → 数据处理与模型选择 → 模型训练 → 模型验证 → 结果讨论。

图 1 本文研究框架

2.2 延误因素识别与指标构建

建筑工程项目工期延误通常受到多方面因素影响，包括项目自身特征、施工组织管理以及外部环境等（Sambasivan & Soon, 2007；王卓甫，2015）。在综合已有研究成果和工程管理实践经验的基础上，本文从五个维度构建工期延误风险指标体系。

（1）项目与技术复杂性因素

项目规模、施工技术复杂程度以及工程结构类型等因素都会对施工进度产生影响。一般而言，项目规模越大、施工技术越复杂，其施工组织难度和管理难度也相应增加（Doloi et al., 2012；王雪青，2018）。

（2）合同与治理结构因素

合同管理方式、业主决策效率以及项目参与方之间的协调机制都会影响项目实施进度。合同管

理不完善或决策流程过于复杂可能导致施工过程中的协调成本增加 (Chan & Kumaraswamy, 1997; 陈勇强 & 王雪青, 2009)。

(3) 资源配置与施工组织因素

施工人员数量、设备配置情况以及材料供应稳定性是影响工程进度的重要因素。资源配置不合理或供应不及时, 容易导致施工计划被迫调整 (Assaf & Al-Hejji, 2006; 丁荣贵 & 张水波, 2011)。

(4) 管理能力因素

施工组织能力、项目管理团队经验以及进度控制机制对工程进度具有重要影响。良好的项目管理体系能够有效降低工期延误风险 (Kerzner, 2017; 刘士军 & 戚安邦, 2010)。

(5) 外部环境因素

政策环境、市场环境以及气候条件等外部因素也可能对施工进度产生一定影响。例如极端天气或政策调整可能导致施工活动受到限制 (Kaming et al., 1997; 马卫东 & 何伯森, 2008)。

通过上述五个维度的综合分析, 形成建筑项目工期延误风险指标体系, 为后续数据分析和模型构建提供基础。

2.3 研究样本

本文以华南地区建筑工程项目为研究对象。华南地区建筑市场规模较大, 工程建设活动频繁, 具有较高的研究代表性。研究样本主要来源于建筑企业项目管理资料以及工程项目数据库。

在数据筛选过程中, 首先剔除信息缺失较为严重或数据异常的项目记录, 以保证样本数据的完整性与可靠性。最终获得 126 个建筑工程项目样本作为研究数据集。样本项目主要包括住宅建筑、商业建筑等类型的民用建筑项目, 如表 1 所示。

项目特征 分类 / 数值范围	频数 (n=78)	占比 (%)
项目地点 一线城市 (华南区域)	52	41.3
二线城市 (华南区域)	44	34.9
三线及以下城市 (华南区域)	30	23.8
项目体量 建筑面积 < 30,000 m ²	57	45.2
30,000 m ² - 80,000 m ²	37	29.4
> 80,000 m ²	32	25.4

表 1 项目基本信息

项目特征	分类 / 数值范围	频数 (n=78)	占比 (%)
延误情况	无延误	29	23.0
	轻度延误 ($\leq 10\%$)	46	36.5
	中度延误 (10% - 20%)	42	33.3
	严重延误 ($> 20\%$)	9	7.1
施工单位性质	国有大型施工企业	58	46.0
	民营施工企业	68	54.0

表 1 项目基本信息 (续)

样本数据主要涵盖项目基本信息、施工管理信息以及项目进度信息等内容，这些数据为分析建筑项目工期延误风险提供了重要依据。

2.4 数据处理及编码

在模型构建之前，需要对原始数据进行预处理，以提高数据分析的有效性。首先，对原始数据进行缺失值处理和异常值识别。对于少量缺失数据，通过均值填补或删除方式进行处理；对于明显异常的数据记录，则根据实际情况进行修正或删除。

其次，对不同类型变量进行统一编码处理。对于数值型变量，采用标准化方法进行处理，以减少不同量纲对模型训练的影响。对于分类变量，则通过数值编码方式进行转换，以便机器学习模型进行识别和计算。

此外，为提高模型训练效果，对部分变量进行特征构造与变量筛选，以减少冗余变量对模型预测结果的干扰 (Goodfellow et al., 2016)。

2.5 模型构建及方法选择

在模型构建方面，本文采用 XGBoost (Extreme Gradient Boosting) 算法 作为主要分析方法。XGBoost 是一种基于梯度提升树的集成学习算法，具有较强的预测能力和良好的泛化性能，被广泛应用于工程管理和风险预测研究 (Chen & Guestrin, 2016)。

与传统统计方法相比，XGBoost 在处理非线性关系和多变量交互方面具有明显优势，同时能够通过特征重要性分析识别影响结果的关键变量。此外，该算法在训练效率和模型稳定性方面也具有较好的表现 (Chen & Guestrin, 2016)。

在模型训练过程中，将样本数据划分为训练集和测试集。训练集用于模型学习和参数优化，测试集用于验证模型预测效果。通过多次交叉验证的方法优化模型参数，以提高模型预测的可靠性。

(Hastie et al., 2009)。

2.6 模型调参方法

本研究以 Python 作为主要分析平台，利用 Scikit-learn 与 XGBoost 库完成数据预处理及模型构建。为确保模型在准确率与泛化能力之间取得良好平衡，本文采用网格搜索（Grid Search）结合交叉验证（5-fold cross validation）进行参数调优。主要调参变量如表 2 所示。

参数名称	含义说明	设置说明或调优范围
max_depth	单棵树的最大深度	3 - 8（步长为 1）
eta	学习率（收缩步长）	0.01 - 0.3（步长为 0.05）
n_estimators	树的数量（最大迭代轮数）	100 - 300
subsample	每轮训练子样本占比	0.6 - 1.0（步长为 0.1）
colsample_bytree	每棵树随机采样的特征占比	0.6 - 1.0（步长为 0.1）
lambda	L2 正则化项系数	默认 1
alpha	L1 正则化项系数	默认 0
objective	多分类目标函数	multi:softprob
eval_metric	模型评估指标（多类分类）	mlogloss、merror 等

表 2 XGBoost 模型主要调参变量

2.7 模型评估指标与解释方法

为评估模型预测性能，本文采用混淆矩阵、准确率（Accuracy）、精确率（Precision）、召回率（Recall）以及 F1 值等指标对模型分类效果进行评价（Sokolova & Lapalme, 2009）。其中，准确率用于衡量模型整体预测正确率，精确率和召回率用于评估模型在识别延误项目方面的能力，而 F1 值则综合反映模型预测性能。

此外，为进一步解释模型预测结果，本文采用特征重要性分析和 SHAP（Shapley Additive Explanations）方法对模型进行解释。SHAP 方法能够量化各变量对模型预测结果的贡献程度，从而帮助识别影响建筑项目工期延误的关键因素（Lundberg & Lee, 2017）。

通过模型评估和解释分析，可以更清晰地理解不同因素在工期延误风险形成中的作用机制，并为施工企业优化项目管理提供数据支持。

3. 研究结果

3.1 样本与描述性分析

本文选取华南地区建筑工程项目作为研究对象，在数据清洗与筛选后共获得 126 个有效样本。样本项目主要包括住宅建筑、商业建筑等民用建筑项目类型，这些项目在规模、施工复杂程度以及资源配置等方面具有一定差异，因此能够较好地反映建筑项目工期延误风险的多样性。

在样本结构方面，大部分项目施工周期集中在 12 个月至 36 个月之间，项目规模以中大型项目为主。从项目复杂性角度看，部分项目涉及较多专业交叉施工或结构复杂施工工艺，这在一定程度上增加了施工组织难度。

在延误情况方面，样本项目中存在一定比例的工期延期现象。描述性统计结果表明，项目延误通常与设计变更频率、材料供应稳定性以及施工组织协调效率等因素存在一定关联。这一结果与既有研究结论基本一致，即建筑工程项目延期往往由多种技术与管理因素共同作用形成[24]。

从指标分布情况来看，不同变量之间存在一定差异。例如，在资源配置维度中，部分项目材料供应稳定性较高，而另一些项目则存在供应不及时的问题；在管理能力维度中，不同项目团队在进度计划制定与风险响应能力方面也存在明显差别。这些差异为后续模型识别关键影响因素提供了数据基础。

本文在综合工期延误理论研究成果与前文文献综述的基础上，从项目与技术复杂性，合同与治理结构，资源与施工组织，外部环境，管理能力五个维度构建工期延误影响因素指标体系如表 3 所示。该体系既涵盖传统进度管理研究中反复验证的重要变量，又兼顾华南地区气候条件与高强度城市建设环境的区域特征。

一级维度	二级指标（变量）	变量类型
技术复杂性	结构形式复杂度、关键工序数量、设计变更次数、专业交叉程度	有序型
合同与治理结构	风险分担、变更、索赔机制	有序型
资源与施工组织	班组经验、材料供应、设备供应、监理协调响应、劳动力缺口	有序型
外部环境	气候、突发事件、监管制度、审批流程	有序型
管理能力	进度计划、信息沟通、风险响应、决策及时性	有序型

表 3 延误风险因素指标体系

3.2 模型预测效果

在完成数据预处理和变量编码后，本文利用 XGBoost 算法构建建筑项目工期延误预测模型。为保证模型评估的可靠性，将样本数据划分为训练集和测试集，并通过交叉验证方法对模型参数进行优化。

模型预测结果表明，该模型在测试数据上的整体预测表现较为稳定。从混淆矩阵结果来看，大部分样本能够被正确分类，表明模型在识别工期延误风险方面具有较好的效果。

在模型评估指标方面，模型准确率（0.75）达到较高水平，同时精确率（0.65）与召回率（0.62）也表现良好。F1 值（0.64）作为综合评价指标，进一步反映了模型在识别延误项目方面的整体性能 [25]。

从图 2 可以看出，模型预测误差主要集中于轻度延误与中度延误之间，两类样本约有 20% 出现交叉错判。这一现象表明，在实际工程项目中，轻度与中度延误在风险暴露强度上具有较强连续性，其驱动因素（如监管、协调效率和资源配置）往往呈现渐进式变化，而非清晰断裂。这使得基于静态指标构建的分类模型在边界区间内更容易产生混淆。

		预测值			
		正常	轻度延误	中度延误	严重延误
实际值	正常	25	4	0	0
	轻度延误	5	38	3	0
	中度延误	0	4	34	4
	严重延误	0	1	2	6

图 2 工期延误预测混淆矩阵

3.3 模型解释与关键风险因素

为了进一步理解模型预测结果，本文采用特征重要性分析与 SHAP 方法对模型进行解释。通过分析不同变量对模型预测结果的贡献程度，可以识别影响建筑项目工期延误的关键因素。

如图 3 所示，分析结果表明，在所有指标中，设计变更频率、材料供应稳定性以及施工组织协调能力等因素对项目延误风险的影响较为显著。这表明在工程项目实施过程中，技术方案变化和资

源供应问题是影响施工进度的重要因素。

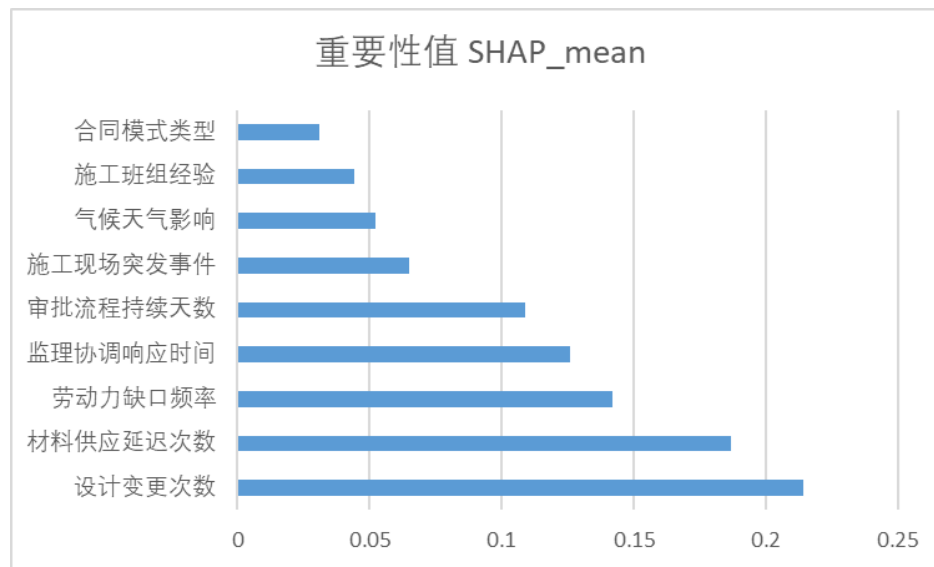


图 3 XGBoost 特征重要性排序条形图

3.4 交互检验

为了进一步分析不同因素之间的交互关系，本文基于模型解释结果对关键变量之间的交互效应进行了检验。如图 4 所示，部分变量之间存在明显的交互影响。

基于 SHAP 交互的计算结果，图 4-4 展示了关键风险因素之间的平均交互贡献强度。结果表明，设计变更次数与材料供应延迟次数之间的交互效应最为显著，其 SHAP 交互均值达到 0.082，明显高于其他变量组合，说明两类风险在项目实施过程中存在显著的联动放大效应。

	设计变更次数	材料供应延迟次数	劳动力缺口频率	监理协调响应时间	审批流程持续天数
设计变更次数	-	0.082	0.067	0.053	0.041
材料供应延迟次数	0.082	-	0.061	0.049	0.038
劳动力缺口频率	0.067	0.061	-	0.044	0.032
监理协调响应时间	0.053	0.049	0.044	-	0.028
审批流程持续天数	0.041	0.038	0.032	0.028	-

图 4 关键风险因素两两交互的 SHAP 热力图

4. 对策及建议

结合前文模型分析结果可以发现，华南地区建筑项目工期延误风险主要来源于设计变更、材料供应稳定性、施工组织协调以及企业管理能力等方面。同时，由于华南地区建筑市场活跃、城市建设密度高、气候条件复杂，项目管理面临更高的不确定性。因此，施工企业需要从企业治理、供应链管理以及数字化管理等多个方面提升项目进度管理能力。基于研究结果，本文提出以下管理对策与建议。

4.1 提升前期决策与设计管理水平

研究结果表明，设计变更是影响建筑项目工期延误的重要因素之一。在华南地区，由于城市建设密度较高，建筑项目往往涉及复杂结构设计、地下空间开发以及多专业协同施工，因此设计阶段的不确定性较大（王雪青,& 陈勇强，2012）。如果设计方案在施工过程中频繁调整，往往会导致施工计划被迫修改，从而影响项目整体进度。

因此，施工企业和项目业主应加强项目实施前期的设计管理。首先，在项目规划阶段应开展充分的技术论证，通过多轮方案评审和专家咨询机制，对工程设计方案进行优化，以减少后期施工阶段的设计调整。对于结构复杂或技术要求较高的项目，可以通过 BIM 技术进行模拟分析，以提前识别潜在问题。

其次，应建立严格的设计变更控制机制。在施工阶段，如果出现设计调整，应通过技术评估和进度影响分析后再进行决策，从而避免频繁变更对施工进度造成冲击。通过强化设计阶段管理，可以从源头上降低工期延误风险。

4.2 提升建筑供应链的稳定性

华南地区建筑市场规模庞大，工程项目数量众多，建筑材料需求量大且来源广泛。许多项目材料需要跨区域采购，因此材料供应稳定性对施工进度具有重要影响。研究结果表明，材料供应不稳定是导致项目施工中断的重要原因之一。

为降低这一风险，施工企业应加强供应链管理。首先，可以通过建立长期合作供应商体系，与核心材料供应商建立稳定合作关系，以提高材料供应可靠性。稳定的供应关系不仅能够降低采购成本，也能够减少供应波动对项目进度的影响（张水波 & 丁荣贵，2013）。

其次，企业应加强材料需求预测与库存管理。通过信息化管理系统对施工进度与材料需求进行动态分析，可以提前制定采购计划，从而避免材料供应滞后问题。

此外，在华南地区大型项目建设中，施工企业还可以通过建立区域性材料供应平台，实现供应链资源共享，从而提高材料供应效率。

4.3 提高复杂施工环境下的组织协同能力

华南地区许多建筑项目位于城市核心区域，施工空间有限，同时周边交通和居民环境对施工活动存在一定限制。这种复杂的施工环境对项目组织协调能力提出了更高要求。

因此，施工企业需要通过优化项目组织结构来提高协同效率。首先，可以建立跨部门协同机制，通过定期项目协调会议及时解决施工过程中出现的问题。通过加强信息沟通，可以减少不同施工单位之间的协调成本（赵丽娜 & 王卓甫，2016）。

其次，应明确项目管理团队的职责分工，并建立清晰的责任体系。通过制定项目管理责任清单，可以减少管理过程中的职责冲突和信息传递延迟问题，从而提高项目执行效率。

此外，在大型工程项目中，可以通过设立项目管理办公室（PMO）对项目进度、资源配置以及风险管理进行统一协调。PMO 能够在项目实施过程中发挥信息整合与决策支持作用，从而提高整体管理效率。

4.4 加强气候风险管理与施工计划优化

华南地区气候条件具有明显特点，全年降雨量较大，并且夏季常受到台风等极端天气影响。这些气候因素可能对施工活动产生直接影响，例如强降雨可能导致施工暂停，台风天气则可能影响高空作业安全。

因此，在项目管理过程中，施工企业需要将气候风险纳入施工计划管理。首先，在施工进度计划编制阶段，应充分考虑季节性气候变化对施工进度的影响。例如，在雨季期间应减少对天气条件依赖较高的施工工序。

其次，可以通过动态进度管理方式，根据气象预测信息及时调整施工计划，从而减少天气因素对施工进度影响。

此外，在大型工程项目中，可以通过建立气候风险预警机制，对极端天气进行提前预判，从而提高项目风险应对能力。

4.5 推进数字化进度管理与数据驱动决策

随着建筑行业数字化转型的推进，数据驱动管理逐渐成为提升项目管理效率的重要工具。本文研究通过数据分析模型识别了影响建筑项目工期延误的关键因素，这表明数据分析方法在工程项目管理中具有重要应用价值。

因此，施工企业可以通过建设数字化项目管理平台，实现对项目进度的实时监控。通过整合施工进度数据、资源配置数据以及项目管理信息，可以为管理人员提供更加全面的决策支持（周勇 & 丁士昭，2014）。

同时，可以将预测模型嵌入项目管理系统，对项目进度风险进行动态评估。当系统识别到潜在

延期风险时，可以提前发出预警，从而帮助管理人员及时采取应对措施。

此外，通过对历史项目数据进行分析，企业可以不断优化项目管理策略，逐步形成以数据为基础的的项目管理模式。

5. 研究限制及未来研究建议

尽管本文通过数据驱动方法对华南地区建筑项目工期延误风险进行了分析，并识别了若干关键影响因素，但由于研究条件和数据来源的限制，仍然存在一定不足，未来研究可从以下几个方面进一步深化。

（1）样本规模与区域范围存在一定局限。

本研究主要选取华南地区建筑工程项目作为研究样本，样本数量为 126 个，虽然能够在一定程度上反映该区域建筑项目进度管理的基本特征，但总体样本规模仍然有限。同时，不同地区在市场环境、管理模式以及施工条件方面存在差异，因此研究结论的普适性仍有待进一步验证。未来研究可以通过扩大样本规模，并纳入更多地区建筑项目数据，从而提高研究结果的代表性和稳健性。

（2）研究方法以量化分析为主。

本文主要采用数据驱动模型对建筑项目工期延误风险进行预测与分析，研究结果主要体现为不同风险因素的重要性排序。虽然这种方法能够较为客观地识别关键影响因素，但对于企业管理实践中的具体决策过程和组织行为机制仍缺乏深入解释。未来研究可以结合深度访谈、案例研究等定性研究方法，对项目管理人员和施工企业管理者进行访谈，从而进一步探讨工期延误背后的管理机制。

（3）指标体系仍有进一步完善空间。

本文构建的工期延误风险指标体系主要从项目复杂性、合同治理结构、资源配置、施工组织以及外部环境等方面进行分析。虽然这些指标能够在一定程度上反映建筑项目延误风险的主要来源，但在实际工程项目中，还可能受到企业战略决策、组织文化以及数字化管理水平等因素的影响。未来研究可以在现有指标体系基础上进一步拓展相关变量，从而构建更加全面的分析框架。

（4）数据驱动模型的应用仍有深化空间。

本文采用机器学习方法对建筑项目延误风险进行预测，并通过模型解释方法识别关键因素。然而，随着建筑行业数字化水平的提升，未来研究可以进一步结合大数据分析和智能决策系统，将预测模型嵌入企业项目管理平台，实现实时风险监测与动态决策支持，从而提升数据驱动管理的实践价值。

参考文献

- 陈勇强, 王雪青. (2009). 建设项目工期延误原因及其对策分析. 建筑经济, 30(4), 78 - 81.
- 丁荣贵, 张水波. (2011). 工程项目管理研究的发展趋势与展望. 工程管理学报, 25(1), 1 - 6.
- 刘士军, 戚安邦. (2010). 建设项目进度风险识别与控制研究. 建筑经济, 31(6), 93 - 96.
- 马卫东, 何伯森. (2008). 建筑工程项目进度延误因素分析. 建筑管理现代化, 22(3), 42 - 45.
- 王雪青, 陈勇强. (2012). 工程项目管理研究热点及发展趋势分析. 工程管理学报, 26(2), 1 - 7.
- 王雪青. (2018). 工程项目管理学. 中国建筑工业出版社.
- 王卓甫. (2015). 建设项目工期延误原因分析及对策研究. 建筑经济, 36(5), 45 - 48.
- 张水波, 丁荣贵. (2013). 工程项目风险管理研究综述. 工程管理学报, 27(4), 1 - 8.
- 赵丽娜, 王卓甫. (2016). 建设项目进度风险评价指标体系研究. 建筑经济, 37(7), 85 - 88.
- 周勇, 丁士昭. (2014). 建筑工程项目进度管理关键因素分析. 建筑经济, 35(9), 96 - 100.
- Assaf, S. A., & Al-Hejji, S. (2006). Causes of delay in large construction projects. *International Journal of Project Management*, 24(4), 349 - 357.
- Chan, D. W. M., & Kumaraswamy, M. M. (1997). A comparative study of causes of time overruns in Hong Kong construction projects. *International Journal of Project Management*, 15(1), 55 - 63.

- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785 – 794).
- Cheng, M. Y., Tsai, H. C., & Sudjono, E. (2012). Conceptual cost estimates using evolutionary fuzzy hybrid neural network. *Expert Systems with Applications*, 39(10), 8979 – 8987.
- Doloi, H., Sawhney, A., Iyer, K., & Rentala, S. (2012). Analysing factors affecting delays in Indian construction projects. *International Journal of Project Management*, 30(4), 479 – 489.
- Flyvbjerg, B., Holm, M. S., & Buhl, S. (2003). How common and how large are cost overruns in transport infrastructure projects? *Transport Reviews*, 23(1), 71 – 88.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. Cambridge, MA: MIT Press.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning*. New York: Springer.
- Kaming, P. F., Olomolaiye, P. O., Holt, G. D., & Harris, F. C. (1997). Factors influencing construction time and cost overruns on high-rise projects in Indonesia. *Construction Management and Economics*, 15(1), 83 – 94.
- Kerzner, H. (2017). *Project management: A systems approach to planning, scheduling, and controlling*. Hoboken, NJ: Wiley.
- Kim, B., Lee, H., & Kim, H. (2019). Predicting construction project duration using machine learning techniques. *Automation in Construction*, 104, 104 – 113.
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems* (Vol. 30).

- Marzouk, M., & El-Rasas, T. (2014). Analyzing delay causes in Egyptian construction projects. *Journal of Advanced Research*, 5(1), 49 – 55.
- Sambasivan, M., & Soon, Y. W. (2007). Causes and effects of delays in Malaysian construction industry. *International Journal of Project Management*, 25(5), 517 – 526.
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427 – 437.
- Son, H., Kim, C., & Kim, C. (2019). Fully convolutional network-based building extraction from high-resolution satellite images. *Automation in Construction*, 96, 45 – 58.
- Williams, T. (2005). Assessing and moving on from the dominant project management discourse in the light of project overruns. *IEEE Transactions on Engineering Management*, 52(4), 497 – 508.