2026 European Conference on Computing in Construction
Corfu, Greece
July 12-15, 2026

PREDICTIVE FRAMEWORK FOR DOMINANT DELAY RISKS IN METALLURGICAL PROJECTS

Shivam Pandey, Sahil Garg
Indian Institute of Technology Delhi, Delhi, India

Abstract

Indian steel and aluminium megaprojects frequently experience multi-year schedule overruns despite detailed planning and risk registers. This paper presents a predictive delay-risk framework that learns from 20 completed projects using contextual similarity and interaction-aware modelling of 39 delay events. Projects are encoded through execution context, while events are represented using normalised delay impacts and a structured dependency network derived from empirical co-occurrence and expert judgement. A prototype tool generates probabilistic completion envelopes and ranked, interaction-aware risk lists for new projects. Initial results align with observed completion behaviour and highlight a small set of high-leverage risks for targeted mitigation.

Introduction

Industrial projects in the steel and aluminium sector continue to experience persistent schedule overruns, often extending several years beyond baseline plans. Although prior research identifies numerous delay causes such as approvals, procurement disruptions, contractor performance, and labour issues, these are typically analysed as independent factors or through isolated case studies.

In practice, project teams rely on schedules, Earned Value indicators, and qualitative risk registers, which do not indicate which risks will dominate or how combinations of risks propagate delays. This limits proactive decision-making and leads to reactive control.

This study addresses this gap by proposing a predictive framework that identifies dominant delay risks and models their interaction using limited but structured historical data.

Empirical dataset and delay structure

The framework is developed using an empirical dataset of 20 completed steel and aluminium projects, including greenfield plants and brownfield expansions. Observed schedule overruns range from moderate (6–12 months) to multi-year delays.

A structured library of 39 delay events is defined, covering approvals, procurement, logistics, execution, and governance-related disruptions. Each event is characterised by its phase and delay impact using triangular estimates. Projects are linked to events through a project event matrix, enabling consistent modelling of delay behaviour.

Methodology

Contextual Similarity Learning

Each project is represented by a context vector including plant type, project nature, capacity, region, climate, and contracting model. Similarity between a new project and historical projects is computed using weighted categorical matching and Gaussian similarity for numerical attributes.

Similarity scores are transformed into normalized weights, ensuring that contextually similar projects exert greater influence on delay estimation.

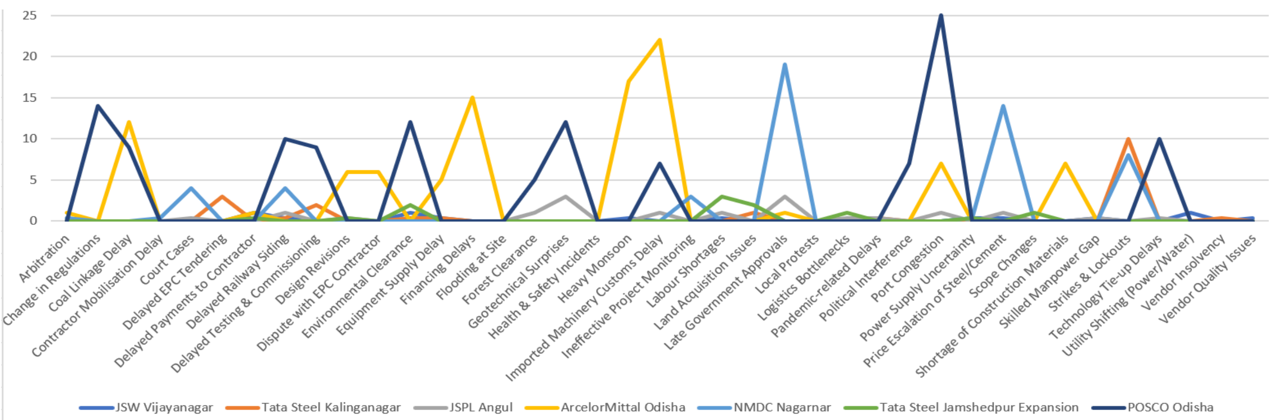

Figure 1: Multi-project delay profile for dominant events showing mean schedule impact (months)

Dominance Estimation of delay events

Delay events are treated as contributors to total delay rather than independent occurrences. For each project, event impacts are normalised to represent their relative contribution to overall delay. For a new project, dominance scores are computed by aggregating these contributions across similar projects.

Table 1: Top delay events ranked by probability, contribution, and mean impact.

Delay events	Probability	Contribution (%)	Mean impact (month)
Price Escalation of Steel/Cement	0.21	8.8	1.04
Shortage of Construction Materials	0.39	8.0	0.94
Delayed Payments to Contractor	0.61	6.6	0.61
Local Protests	0.39	6.2	0.39

Interaction and Amplification Modelling

Delay events are often interdependent. To capture this, a dependency network is defined representing trigger response relationships between events. Amplification factors, derived from empirical co-occurrence and expert judgement, adjust probabilities and impacts of dependent events during simulation.

This enables modelling of cascading effects, where upstream risks propagate across project phases and create clustered delay scenarios.

Predictive Simulation

The model is evaluated using Monte Carlo simulation. Event impacts are sampled from triangular distributions, modified by dependency relationships, and aggregated to estimate total project delay.

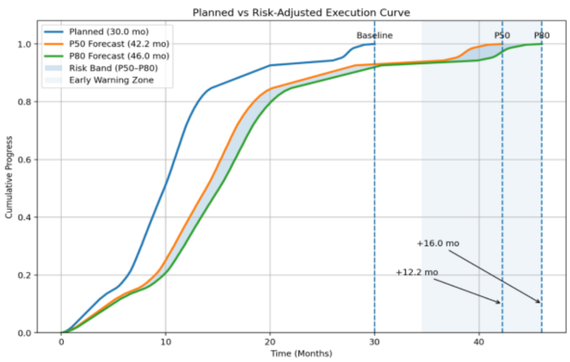

Figure 2: Baseline vs probabilistic completion curves (P50, P80) with risk band

Results and Implementation

The predicted P50 and P80 completion times fall between baseline and actual values, indicating realistic estimation of delay magnitude and uncertainty. Results show that a small subset of risks dominates total delay, while interaction modelling captures cascading effects that produce heavy-tailed outcomes.

The framework is implemented as a Python/Streamlit prototype, generating completion curves, dominant risk rankings, and phase-wise delay insights to support targeted mitigation.

Conclusions

This study presents a predictive framework for delay behaviour in metallurgical projects using contextual similarity, dominance estimation, and interaction-aware simulation. The results show that delay is driven by a small set of dominant and interacting risks, and that modelling these interactions improves realism. The framework enables early identification of high-impact risks, supporting proactive project control.

References

MOSPI, “Flash Report on Programme Implementation,” Ministry of Statistics and Programme Implementation, Government of India, New Delhi, India, 2024.

Reza Azimi, SangHyun Lee, Simaan M. AbouRizk, and Amin Alvanchi, “Framework for an automated and integrated project monitoring and control system,” Automation in Construction, 2013.

A. Shibani & K. Arumugam, “Avoiding Cost Overruns in Construction Projects in India,” Procedia Engineering, 2015.

PMI (India), “Study on Project Schedule and Cost Overruns,” Project Management Institute – India Research Whitepaper, 2017.