Risk-Informed Resource-Constrained Scheduling for Refueling Outage Management

D. Mandelli[a*], C. Wang[a]

[a] Idaho National Laboratory (INL), 2525 N. Fremont Ave, 83415 Idaho Falls (ID), USA

Abstract: Refueling outages of nuclear power plants are high-consequence operations in which unplanned schedule extensions incur significant economic losses. This paper presents a risk-informed outage scheduling and decision-support framework developed to address the principal limitations of current outage management tools. First, outage scheduling is formulated as a Resource-Constrained Project Scheduling Problem (RCPSP) that enforces workforce availability, shared equipment access, radiation dose budgets, consumable inventories, plant system isolation mutual-exclusion rules, and regulatory surveillance windows as hard constraints, capturing the resource contention that critical path methods leave to manual resolution. Second, schedule uncertainty is quantified by constructing empirical activity duration distributions from historical work order records retrieved by semantic similarity, then propagating these distributions through the resource-constrained scheduling model via Monte Carlo simulation to produce completion risk curves, activity criticality indices, and schedule path sensitivities. Third, a decision-support module evaluates emergent work findings encountered during outage execution, combining knowledge-graph-based schedule impact assessment with evidence-traceable ESCALATE / PROCEED / DEFER recommendations for the outage coordinator. Applied to a 66-task PWR refueling outage test case, the choice of scheduling priority rule alone changes the projected outage duration by 71 hours (a \$4.4 million difference) while automated replanning after a simultaneous crew disruption and scope addition completes in under one second. These results demonstrate that risk-informed scheduling makes visible decision consequences that deterministic critical path methods cannot expose.

Keywords: outage management, resource-constrained project scheduling, schedule risk, emergent work

---

1. INTRODUCTION

Each refueling outage is among the most resource-intensive phases of a nuclear plant's operational life: thousands of activities are executed over a 2–3 week period by plant staff and external contractors. Beyond operational complexity, outages have a direct economic impact on plant finances, e.g., a light water reactor loses approximately \$1.0–1.5 million in generation revenue for each day it remains offline. The International Atomic Energy Agency documents [1, 2] that most such extensions are preventable: structured planning disciplines can reduce the standard refueling outage duration by approximately 25%.

Current industry practice relies primarily on critical path method (CPM) tools such as Oracle Primavera P6.[1] Industry assessments confirm that CPM-based scheduling remains the dominant approach despite its well-documented limitations in resource-intensive outage environments [3]. These tools are effective at capturing task dependencies and identifying the logical critical path, but their fundamental limitation is that resource constraints (workforce availability by skill type, shared equipment, radiation dose budgets, consumable inventories, plant system isolation rules, and regulatory surveillance windows) are displayed as advisory histograms rather than enforced during schedule generation. Resolution of resource conflicts is left to the outage planner's manual judgment, a process that scales poorly to the

---

[*] Corresponding author, diego.mandelli@inl.gov

[1] Oracle Primavera P6; `https://www.oracle.com/construction-engineering/primavera-p6/`.

thousands of interdependent tasks typical of a large refueling outage.

Two addtional limitations compound the resource conflict problem: first, deterministic scheduling produces a single schedule with no quantification of completion risk, i.e., the outage manager knows the planned end date but not the probability of achieving it, which activities most threaten the schedule across the range of possible outcomes, or how sensitive the end date is to individual activity duration variations. The second limitation is the lack of structured decision support for emergent work (unplanned scope emerging from equipment findings during the outage). The manager must simultaneously assess technical urgency, regulatory standing, schedule impact, and resource availability, typically under time pressure.

This paper presents LOGOS [2], an integrated scheduling and decision-support (combining scheduling, uncertainty quantification, and decision support into a unified workflow for outage coordinators) framework that addresses these three limitations.

2. OUTAGE SCHEDULING AS A PROBABILISTIC DECISION PROBLEM

2.1. Decision Phases and Their Information Needs

Outage management involves two qualitatively different decision phases. In the planning phase, decisions concern scope, sequencing, resource allocation, and contingency sizing: how should work be ordered given skill-type availability profiles, shared equipment conflicts, and regulatory constraints? How much schedule margin is sufficient, and where should it be positioned? These decisions are made weeks to months before the outage begins, and their quality determines the feasibility of the published outage schedule.

In the execution phase, the plan meets physical reality. Activity durations deviate from estimates. Crew assignments change due to absenteeism. Equipment failures require replacement. Engineering walkdowns uncover additional scope. The critical question shifts from “what is the plan?” to “given what has happened, what is the best recovery path and how long will it take?”

Current deterministic tools answer neither question with quantified risk information. LOGOS provides a unified framework for both phases, as illustrated in Figure 1.

2.2. RCPSP Formulation

Outage scheduling is here formulated as a Resource-Constrained Project Scheduling Problem [4, 5]. Let $J = \{1, \dots, n\}$ be the set of activities and $s_j \geq 0$ the start time decision variable for each activity. The objective is to minimise the project makespan:

$$\min\ C_{\max} = \max_{j \in J} (s_j + p_j) \tag{1}$$

subject to seven constraint families that capture the nuclear outage domain.

Precedence with lags (ℓ_{ij} is a mandatory inter-task lag, e.g. cooldown hold):

$$s_j \geq s_i + p_i + \ell_{ij}, \quad (i, j) \in E \tag{2}$$

Renewable resources (workforce by skill type and shared equipment; $R_k(t)$ is time-varying to capture shift calendars and crew surge schedules):

$$\sum_{j \in J(t)} r_{jk} \leq R_k(t) \quad \forall k, t \tag{3}$$

---

[2] LOGOS GitHub repository: https://github.com/idaholab/logos

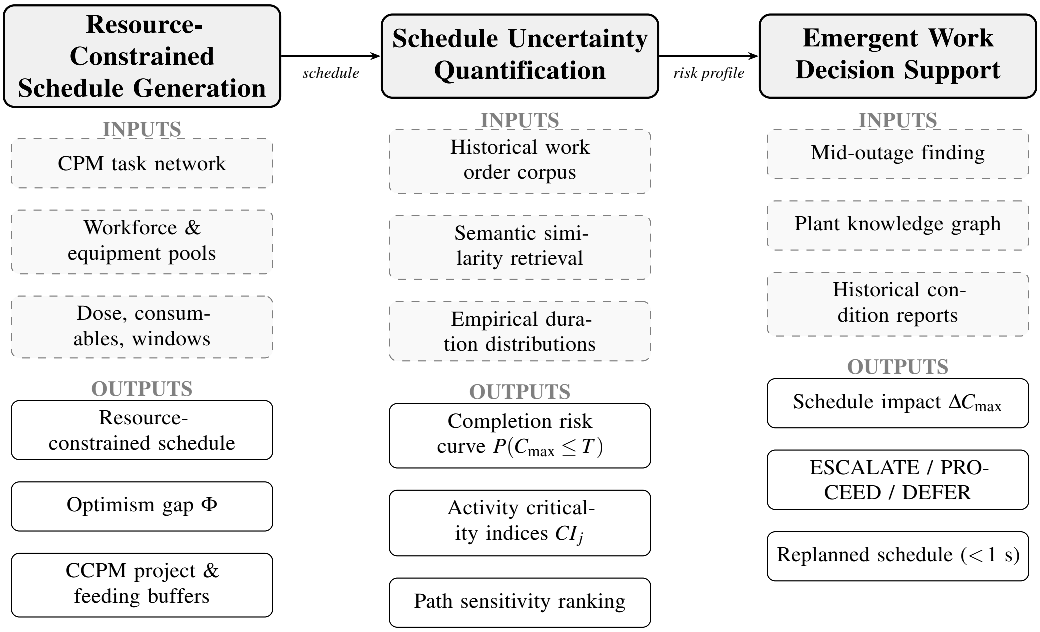

Figure 1: LOGOS framework: three integrated capabilities covering the planning phase (left, centre) and execution phase (right). The resource-constrained schedule feeds the uncertainty quantification module; the resulting risk profile and the current schedule state feed the emergent work decision-support module.

Consumable resources (anti-contamination suits, specialty gases, and other non-replenishing inventories C_c):

$$\sum_{j \in J} c_{jc} \leq C_c \tag{4}$$

Regulatory time windows (Technical Specification surveillance deadlines and NRC hold point release windows):

$$w_j^{\min} \leq s_j \leq w_j^{\max} - p_j \tag{5}$$

Radiation dose budgets (ALARA limits B_s per skill class s; δ_j is activity dose rate, n_{js} is crew count from class s):

$$\sum_{j \in J_s} \delta_j \cdot n_{js} \cdot p_j \leq B_s \quad \forall s \tag{6}$$

Plant system mutual exclusion (redundant safety trains cannot be concurrently out of service):

$$L_{j_1}(t) \cap L_{j_2}(t) = \emptyset \quad \forall\, j_1 \neq j_2 \text{ with conflicting system states} \tag{7}$$

Operationally, mutual exclusion is enforced by a reference-counted lock on each plant system state: an activity requiring an exclusive system state is blocked until no other in-progress activity holds a conflicting state, and the lock is released when the reference count returns to zero on activity completion.

This formulation is NP-hard in general [4]. The nuclear-specific extensions (time-varying resource availability, dose budgets, and system mutual exclusion) increase the practical search space further. Exact integer programming solvers can be employed for small problems since they can handle constraint (3) but do not model dose budgets or system state exclusion, and their solution times on realistic outage instances exceed operational requirements. Primavera enforces constraint (1) but treats constraints (2)-(6) as advisory only.

3. RESOURCE-CONSTRAINED SCHEDULE GENERATION

3.1. The Optimism Gap

The CPM duration is a lower bound computed under the assumption of unlimited resources. LOGOS enforces finite resources, producing a realistic makespan. The “optimism gap ratio” defined as

$$\Phi = C_{\max}^{\text{LOGOS}} \,/\, C_{\max}^{\text{CPM}} \tag{8}$$

quantifies how much longer the outage realistically takes once resource contention is enforced. $\Phi > 1$ is not a scheduling failure: it is the true cost of resource competition that CPM hides. In the Unit 2 case (Section 6), Φ ranges from 1.04 to 1.34 depending on the priority rule applied, representing 10 to 81 additional hours beyond the CPM lower bound.

3.2. Event-Driven Scheduling Engine

LOGOS implements an event-driven heuristic scheduler. The simulation clock advances only to the next meaningful event (task completion, resource availability change, surveillance window opening, shift boundary, or consumable restock delivery). At each event, the scheduler identifies all activities whose predecessors are complete and whose earliest start time has been reached, ranks them by a priority rule, and assigns them while a unified feasibility check confirms that all constraint pools simultaneously permit the assignment.

LOGOS implements a parallel Schedule Generation Scheme (SGS): at each event, all feasible candidates are ranked by the active priority rule and assigned greedily in rank order while the unified feasibility check holds; when no further assignment is possible, the clock advances to the next event. Replanning after a mid-outage disruption uses the same mechanism: LOGOS freezes the start and end times of completed and in-progress activities, releases their resource allocations at the appropriate times, and re-runs the parallel SGS on the remaining unstarted activities from the current resource state, producing a revised schedule in seconds without disturbing committed work.

Twenty-two priority rules are available across five families: time-float rules (Latest Finish, Latest Start, Earliest Finish, Shortest Processing Time), graph topology rules (Most Total Successors, Greatest Rank Positional Weight), resource demand rules, slack-displacement rules, and machine-learned hyper-heuristic rules derived by genetic programming from historical scheduling instances [6, 7]. Because each rule produces a complete schedule in seconds, the best rule for a given outage configuration is identified by running all rules and selecting the minimum-makespan result.

3.3. Resource-Constrained Critical Chain and Buffers

After scheduling, LOGOS constructs an augmented project graph by adding contention edges to the original precedence network: if two activities compete for the same constrained resource and one must wait for the other, a directed edge is inserted between them. The longest path through this augmented graph is the *resource-constrained critical chain*: the true schedule-limiting sequence, which may differ substantially from the CPM critical path when resource contention is significant.

Activities are classified by their actual total float TF_{actual}, computed over augmented-graph successors. Activities with $TF_{\text{actual}} = 0$ are on the chain and shown in red on the Gantt display; activities with $TF_{\text{actual}} \approx 0$ are one disruption away from criticality (orange); activities with positive actual float are buffered (blue). This classification is more operationally meaningful than CPM float, which can misclassify an activity as non-critical when its float is illusory because a resource constraint will block it in practice.

LOGOS inserts Critical Chain Project Management buffers [8] at two locations:

- A *project buffer* at the end of the critical chain provides the primary schedule reserve, sized by the sum-of-squares method over nominal and aggressive duration estimates for chain activities.

- A *Feeding buffers* protect each merge point where a non-critical chain feeds into the critical chain. Buffer consumption during execution is monitored by a traffic-light protocol (green: $<50\%$, yellow: 50–100%, red: $>100\%$), giving the outage coordinator a continuous readout of schedule health without requiring Gantt-level inspection.

4. SCHEDULE UNCERTAINTY QUANTIFICATION

4.1. Empirical Duration Distributions from Plant History

Deterministic scheduling treats activity durations as fixed values. In practice, the time required to complete a maintenance task depends on many factors (e.g., as-found equipment condition, crew composition, tooling availability, and plant-specific configuration); all these factors introduce variability that a single point estimate cannot capture.

We employ here DACKAR [3] to construct empirical duration distributions for each activity type by querying the plant's historical work order database [9]. For a target activity j with description text d_j, a semantic similarity search retrieves the k most analogous completed work orders:

$$\mathscr{W}_j = \text{top-}k\big\{\, w \in \mathscr{W}_{\text{hist}} : \text{sim}(d_j, d_w)\big\} \tag{9}$$

where $\text{sim}(\cdot,\cdot)$ is cosine similarity over a transformer-based sentence encoder fine-tuned on nuclear maintenance text, and $\mathscr{W}_{\text{hist}}$ is the historical work order corpus. Retrieved work orders with anomalous durations (statistical outliers identified by interquartile range filtering) are excluded before the empirical sample is assembled. The actual durations of the retained retrieved work orders form the empirical distribution for activity j, capturing the full range of outcomes, including the right-tail overruns that deterministic estimates systematically exclude.

This data-driven approach is reproducible and auditable: the retrieved work order sample can be inspected and the similarity query reviewed for relevance. It also updates automatically as new work orders accumulate in the plant's corrective action program records, without requiring expert re-elicitation.

4.2. Monte Carlo Risk Propagation

Schedule risk is quantified by propagating the empirical duration distributions through the full resource-constrained scheduling model via Monte Carlo simulation using RAVEN [10]. Each simulation trial draws independent samples $\tilde{p}_j \sim F_j$ for each activity j and executes a complete LOGOS scheduling run with those durations, producing a sampled makespan $\tilde{C}_{\max}$. Repeating for N trials yields the empirical distribution of project makespan from which three risk metrics are extracted.

The completion time distribution $P(C_{\max} \leq T)$ gives the probability that the outage completes by any target date T. For the planned return-to-service date T^*, the value $P(C_{\max} \leq T^*)$ is the *schedule reliability* (the probability of on-time completion given duration uncertainty and resource contention).

The *criticality index* CI_j for activity j is the fraction of Monte Carlo trials in which j lies on the resource-constrained critical chain:

$$CI_j = \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}\big[j \in \text{chain}_n\big] \tag{10}$$

A high criticality index ($CI_j \geq 0.80$) identifies activities that drive the end date across most uncertainty scenarios, representing the schedule's structural vulnerabilities. CPM identifies only activities with

---

[3] Digital Analytics, Causal Knowledge Acquisition and Reasoning (DACKAR) GitHub repository: https://github.com/idaholab/DACKAR

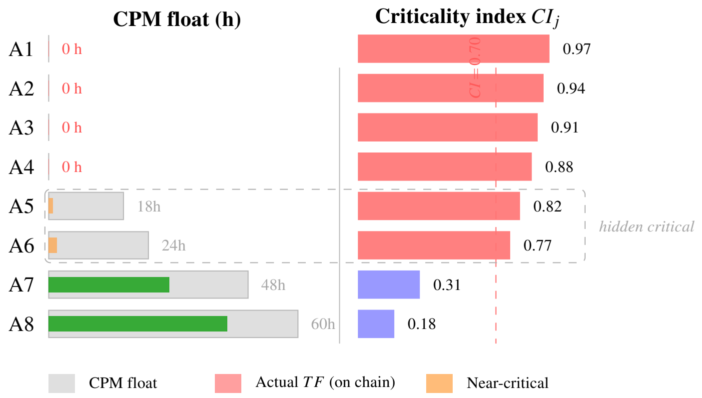

Figure 2: CPM total float (gray bars), actual total float under resource constraints (coloured overlay), and Monte Carlo criticality index CI_j (right panel) for eight representative activities. Values are illustrative, chosen to be consistent with the Unit 2 case results (10 activities with $CI_j \geq 0.70$, of which 4 are on the CPM critical path and 6 carry positive CPM float). Activities A5 and A6 carry 18–24 hours of CPM float but near-zero actual float and $CI_j \geq 0.77$: resource contention eliminates their apparent CPM buffer in the majority of Monte Carlo scenarios, making them *hidden critical* activities that CPM-only analysis would not flag for priority treatment.

zero float in the nominal schedule; the criticality index reveals the broader set of activities that become critical as durations vary, providing a richer basis for contingency allocation decisions.

Path sensitivity analysis identifies which activity duration variations most strongly influence the project makespan. For each activity j, the Spearman rank correlation between $\tilde{p}_j$ and $\tilde{C}_{\max}$ across Monte Carlo trials is computed. The ranked sensitivity list prioritises which activities warrant tighter duration estimation, additional standby resources, or pre-positioned contingency.

Figure 2 illustrates the relationship between CPM total float, actual total float under resource constraints, and Monte Carlo criticality index for eight representative activities from the Unit 2 case. Activities A1–A4 are identified as critical by both CPM and the resource-constrained analysis. Activities A5 and A6 carry positive CPM float (18 and 24 hours respectively) but have near-zero actual float and high criticality indices ($CI = 0.82$ and 0.77): resource contention eliminates their apparent CPM buffer in most scenarios. These are the *hidden critical* activities, invisible to CPM-only analysis and the primary source of unexpected schedule extensions. Activities A7 and A8 are genuinely buffered under both analyses.

4.3. Risk-Informed Contingency Positioning

Traditional outage contingency is allocated as a flat percentage of total planned duration (typically in the 10-20% range) appended to the project end date. This approach provides no guidance on what probability of on-time completion it achieves or where within the schedule the margin is most needed.

The Monte Carlo completion risk curve directly quantifies the contingency required to achieve a target schedule reliability. If the P70 completion time (the time by which 70% of Monte Carlo trials finish) is 12 hours beyond the nominal schedule, an outage manager targeting 70% reliability needs exactly that 12-hour buffer. CCPM project and feeding buffers, sized from empirical duration uncertainty, serve as the structural vehicle for this allocation, positioning margin at the end of the resource-constrained critical chain and at merge points where it is most effective, rather than distributing it uniformly across all activities.

5. EMERGENT WORK DECISION SUPPORT

During outage execution, equipment findings (inspection results, condition report items, and unexpected scope additions) arrive continuously and require triage decisions that balance technical urgency, regulatory obligations, schedule impact, and resource availability. The DACKAR decision-support module provides structured, evidence-traceable assistance for these decisions.

5.1. Schedule Impact Assessment

When an emergent work item is submitted, the module first quantifies its schedule impact. The item is characterised by its estimated duration, required skill type and headcount, required equipment, plant system, and any regulatory classification. The module queries the plant knowledge graph [11, 9] to identify other work items currently on the schedule that share the same equipment, skill pool, plant system, or physical zone as the emergent item. Schedule impact is evaluated by inserting the emergent item into the current schedule state and executing a replanning pass to compute the change in projected makespan $\Delta C_{\max}$.

5.2. Evidence Retrieval and Causal Context

For emergent items originating from equipment condition findings, the module retrieves historical condition reports and work orders analogous to the current finding using the same semantic similarity search described in Section 4.1. Retrieved records provide causal context: has this failure mode been observed before, were there precursor signals, and did prior corrective actions close the root cause or merely the symptom? This context informs both urgency classification (i.e., is this a known recurring pattern requiring escalation, or a first-occurrence degradation?) and scope estimation, since analogous historical findings may reveal additional corrective work not apparent from the initial inspection.

5.3. ESCALATE / PROCEED / DEFER Recommendations

The module synthesises impact assessment and retrieved evidence into a recommendation from three options, each with its evidence basis documented:

- **ESCALATE**: The finding exceeds a criticality threshold based on regulatory classification, safety function impact, or schedule impact magnitude ($\Delta C_{\max}$ above a configurable limit). Immediate coordinator review is required before adjacent work continues.
- **PROCEED**: The finding is within acceptance criteria, historical analogs confirm it is routine, schedule impact is below threshold, and resources are available. Incorporate the work into the running schedule and continue.
- **DEFER**: The finding does not affect current operability, resources are unavailable, or schedule impact would be unacceptable. Document for the next maintenance window; record the technical basis for deferral in the corrective action program.

Each recommendation is accompanied by the computed $\Delta C_{\max}$, the specific regulatory basis where applicable, the retrieved historical analogs supporting the classification, and any open uncertainties the coordinator should verify before accepting the recommendation. The full evidence chain is stored in the run manifest, creating an auditable record of the triage decision that satisfies corrective action program documentation requirements.

5.4. Proactive Risk Screening

LOGOS also includes a pre-outage risk screening capability that identifies high-risk components before the outage begins. Condition reports, work orders, and maintenance records for equipment scheduled for work in the upcoming outage are retrieved and analysed for indicators of elevated failure risk: recurring corrective maintenance, overdue preventive maintenance tasks, adverse trends in condition monitoring data, and prior incomplete corrective actions. Components meeting configurable risk thresholds are flagged in the pre-outage schedule for expedited inspection planning, additional material

Table 1: Unit 2 PWR refueling outage case parameters.

Parameter	**Value**
Total activities	66
Workforce skill types	3 (mechanics, electricians, welders)
Peak workforce	16 workers
Shared equipment items	3 (polar crane, EC rig, refueling machine)
Consumable item type	1 (anti-contamination suits)
Surveillance time windows	5
Hold points (QA and NRC)	3
System isolation mutual-exclusion rules	2
CPM critical path duration	238 h (9.9 days)
Target outage window	408 h (17 days)

pre-staging, or enhanced contingency allocation. This shifts risk management from reactive to proactive, reducing the incidence of mid-outage emergent scope.

6. APPLICATION: PWR UNIT 2 REFUELING OUTAGE

6.1. Case Description

The developed methods are here applied to a 66-task PWR refueling outage test case. The problem instance exercises all seven constraint families and is representative of the scale and complexity of a typical short-cycle refueling outage. Table 1 summarises the instance parameters.

The mechanical crew expands from 8 to 12 workers at day 5, modelled as a time-varying resource pool. Containment access is blocked for the first 24 hours (cooldown). The reactor cavity is restricted to one concurrent activity at all times. System isolation mutual-exclusion rules prevent both ECCS trains and both spent fuel pool cooling trains from being simultaneously out of service. Anti-contamination suit inventory starts at 50 units at 4 units per task, with a restock delivery at hour 96. Radiation dose budgets are enforced for radiological control technician personnel. One activity (steam generator tube eddy-current inspection) offers a crash execution mode that reduces duration from 24 to 16 hours at the cost of additional personnel and equipment.

6.2. Schedule Generation Results

Table 2 presents the resource-constrained makespan produced by six representative priority rules. The best-performing rule (Latest Finish, LF) yields a scheduled duration of 248 hours, an optimism gap ratio of $\Phi = 1.042$, representing 10 hours of resource-contention overhead above the 238-hour CPM lower bound. All five Technical Specification surveillance windows are satisfied in this schedule despite resource pressure.

The worst-tested rule (Greatest Rank Positional Weight, GRPW) yields 319 hours ($\Phi = 1.34$), an 81-hour excess above the CPM lower bound. The spread between best and worst rules is 71 hours, representing a \$4.4 million difference at \$1.5 million per day replacement power cost, attributable entirely to priority rule selection. This consequence is invisible in CPM-based method output and cannot be identified without explicit resource-constraint modelling.

The CCPM project buffer for the LF schedule, sized from empirical duration uncertainty, is 42.2 hours. Five feeding buffers protecting merge points on the resource-constrained critical chain range from 5.0 to 45.6 hours. These structured margins replace a flat contingency estimate with analytically grounded values positioned where they are most effective.

Table 2: Priority rule comparison on the Unit 2 case. All schedules satisfy the five surveillance windows. Financial impact at \$1.5M/day.

Priority Rule	Duration (h)	Ratio Φ	Gap vs. LF
Latest Finish (LF)	248	1.042	—
Earliest Finish (EF)	261	1.097	+13 h / \$0.8M
Greatest Res. Demand (GRD)	274	1.151	+26 h / \$1.6M
Shortest Processing Time (SPT)	290	1.218	+42 h / \$2.6M
Most Total Successors (MTS)	305	1.281	+57 h / \$3.6M
Gr. Rank Pos. Weight (GRPW)	319	1.340	+71 h / \$4.4M
CPM lower bound	238	1.000	—

6.3. Multi-Mode Analysis

We consider here the situation where the steam generator eddy-current inspection offers a crash mode that reduces duration from 24 to 16 hours by deploying an additional radiological control technician and a second inspection rig. A planner examining CPM total float would conclude this saves 8 hours off the outage. DACKAR analysis shows the opposite: this activity carries 50 hours of CPM float but is not on the resource-constrained critical chain in any scenario. Deploying crash mode diverts constrained personnel and equipment from activities that are on the chain, extending the projected makespan by 2 hours. This result demonstrates that multi-mode decisions must be evaluated against the resource-constrained schedule, not the logical precedence network alone.

6.4. Schedule Uncertainty and Criticality Indices

Monte Carlo analysis propagates empirical duration distributions through the resource- constrained schedule. The findings for the Unit 2 case were the following:

- Completion risk: The P50 (median) completion time for the LF schedule is 252 hours and the P90 is 271 hours, indicating that contingency planning should anticipate up to 33 additional hours beyond the CPM lower bound in 10% of outage scenarios. The 408-hour target window provides high schedule reliability, but the margin above the P90 completion time provides the useful contingency budget for the outage coordinator.
- Criticality indices: Ten activities have criticality indices $CI_j \geq 0.70$. Four of these are on the CPM critical path; six have positive CPM total float but appear on the resource-constrained chain in the majority of Monte Carlo trials because resource contention eliminates their apparent float under realistic duration variability. These six activities would not be prioritised by a planner relying on CPM alone and represent the primary hidden risk in the nominally “buffered” schedule.
- *Path sensitivity:* The five activities with the highest Spearman rank correlation between their duration and the project makespan are candidates for proactive resource pre-positioning or expedited work authorisation, since duration reductions on these activities translate most directly to schedule improvement.

6.5. Mid-Outage Replanning

We evaluate the scenario where at hour 72 of outage execution, two disruptions arrive simultaneously: an emergent turbine inspection scope addition (3 hours, 2 mechanics) and a crew reduction from 12 to 5 mechanics due to illness. DACKAR promptly replans the remaining schedule from a snapshot of in-progress and completed work. The revised projected makespan increases from 248 to 290 hours (a 42-hour extension representing a \$2.6 million economic loss). All five surveillance windows remain satisfied in the replanned schedule.

Without automated replanning, recovering a manually constructed schedule after a disruption of this magnitude typically requires 2–4 hours of outage coordinator effort. The same replanning computation that identifies the 42-hour impact also evaluates which remaining activities can be re-sequenced or mode-switched to partially offset the extension, providing the coordinator with a prioritised set of recovery options alongside the impact assessment.

7. CONCLUSIONS

This paper presents LOGOS, a risk-informed framework for nuclear refueling outage scheduling and decision support, implemented using DACKAR for semantic retrieval and RAVEN for Monte Carlo simulation. Three capabilities advance outage management from deterministic planning toward probabilistic, evidence-based decision-making.

Resource-constrained schedule generation enforces the six constraint families that critical path methods leave to manual resolution. The optimism gap ratio Φ quantifies the resource-contention overhead hidden in every CPM schedule. On the Unit 2 case, priority rule selection alone changes the projected duration by 71 hours, a $4.4 million consequence invisible in CPM output.

Monte Carlo uncertainty quantification constructs empirical activity duration distributions from historical work order data and propagates them through the full resource-constrained scheduling model. The resulting completion risk curve, activity criticality indices, and path sensitivity ranking give outage managers quantified answers to planning questions that deterministic scheduling cannot address: what is the probability of meeting the return-to- service date, which activities are structural vulnerabilities across the range of possible duration outcomes, and how much contingency is required to achieve a target schedule reliability?

The emergent work decision-support module evaluates mid-outage findings against the current schedule state, retrieves historical analogs, and computes schedule impact to produce evidence-traceable ESCALATE/PROCEED/DEFER recommendations. Mid-outage replanning after a simultaneous crew disruption and scope addition completes in under one second, replacing hours of manual effort with a quantified impact assessment and a prioritised set of recovery options.

Current technology readiness is TRL 4: core capabilities are implemented and validated on representative synthetic outage data. Near-term development efforts will focus on integration with plant work management systems to eliminate manual data entry, calibration of duration distributions against plant-specific historical records, and validation of the emergent work pipeline against completed corrective action program records from operating plants.

The broader principle shown in this paper is that risk-informed outage management is achievable with data that already exists at every generating station (work order histories, CMMS records, and plant equipment models) combined with computational methods that are standard today.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy, Office of Nuclear Energy, under the Light Water Reactor Sustainability (LWRS) program. AI tools assisted with the development of the presented framework and the revision of this paper. All technical analyses, designs, results, and conclusions were developed and verified by the authors.

REFERENCES

[1] International Atomic Energy Agency, “Good practices for outage management in nuclear power plants,” Tech. Rep. IAEA-TECDOC-498 or IAEA-TECDOC-621, IAEA, Vienna, Austria, 2002.

[2] International Atomic Energy Agency, “Indicators for management of planned outages in nuclear power plants,” Tech. Rep. IAEA-TECDOC-1490, IAEA, Vienna, Austria, 2006.

[3] Electric Power Research Institute, “Methodologies for optimal outage scheduling: A framework and assessment,” Tech. Rep. EPRI 3002029703, EPRI, Palo Alto, CA, 2024.

[4] J. Blazewicz, J. Lenstra, and A. Rinnooy Kan, “Scheduling subject to resource constraints: Classification and complexity,” *Discrete Applied Mathematics*, vol. 5, no. 1, pp. 11–24, 1983.

[5] R. Kolisch and S. Hartmann, “Experimental investigation of heuristics for resource-constrained project scheduling: An update,” in *European Journal of Operational Research*, vol. 174, pp. 23–37, 2006.

[6] S. Hartmann and D. Briskorn, “A survey of variants and extensions of the resource-constrained project scheduling problem,” *European Journal of Operational Research*, vol. 207, no. 1, pp. 1–14, 2010.

[7] V. Valls, F. Ballestin, and S. Quintanilla, “Justification and RCPSP: A technique that pays,” *European Journal of Operational Research*, vol. 165, no. 2, pp. 375–386, 2005.

[8] E. Goldratt, *Critical Chain*. Great Barrington, MA: North River Press, 1997.

[9] C. Wang, D. Mandelli, and J. Cogliati, “Technical language processing of nuclear power plants equipment reliability data,” *Energies*, vol. 17, no. 7, 2024.

[10] A. Alfonsi, C. Rabiti, D. Mandelli, J. Cogliati, C. Wang, P. Talbot, D. Maljovec, and C. Smith, “RAVEN user manual,” Tech. Rep. INL/EXT-15-34123, Idaho National Laboratory, 2021.

[11] R. Liu, R. Fu, K. Xu, X. Shi, and X. Ren, “A review of knowledge graph-based reasoning technology in the operation of power systems,” *Applied Sciences*, vol. 13, no. 7, 2023.