

Intelligent Job Scheduling for HPC Systems:

A Statistical Evaluation of Deep Reinforcement Learning Approaches

Justin M. Cheney

2026-04-02

University of the Western Cape

Honours 2026 - Submission 1

Problem Context & Literature Gap

HPC Scheduling Challenge

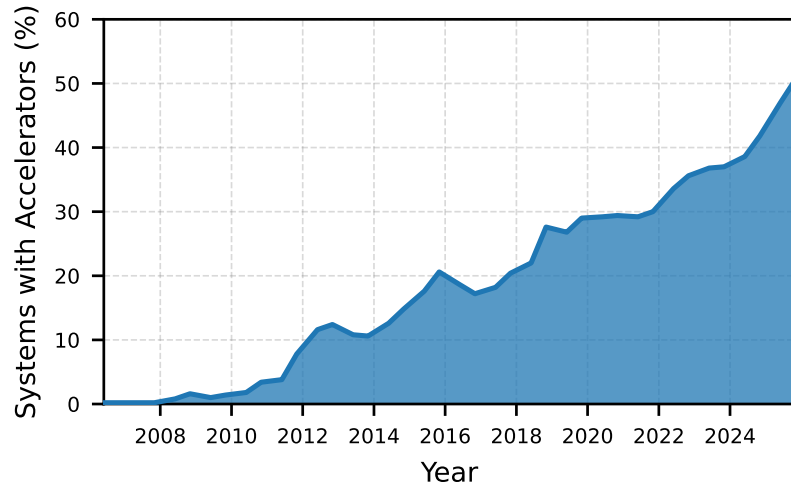


Figure 1: TOP500 heterogeneous systems growth

- Traditional heuristics inflexible
- Meta-heuristics don't generalise
- DRL offers adaptive policy learning

Critical Literature Gap

Systematic Survey (59 papers, 2024-2025)

- 27 papers (46%) use PPO-based approaches
- Only 12 papers (20%) empirically justify this choice
- Most cite “stability” without comparative evidence

Research Implication

Methodological convergence without empirical validation suggests **algorithmic path dependence**

Research Questions & Methodology

Primary Research Question

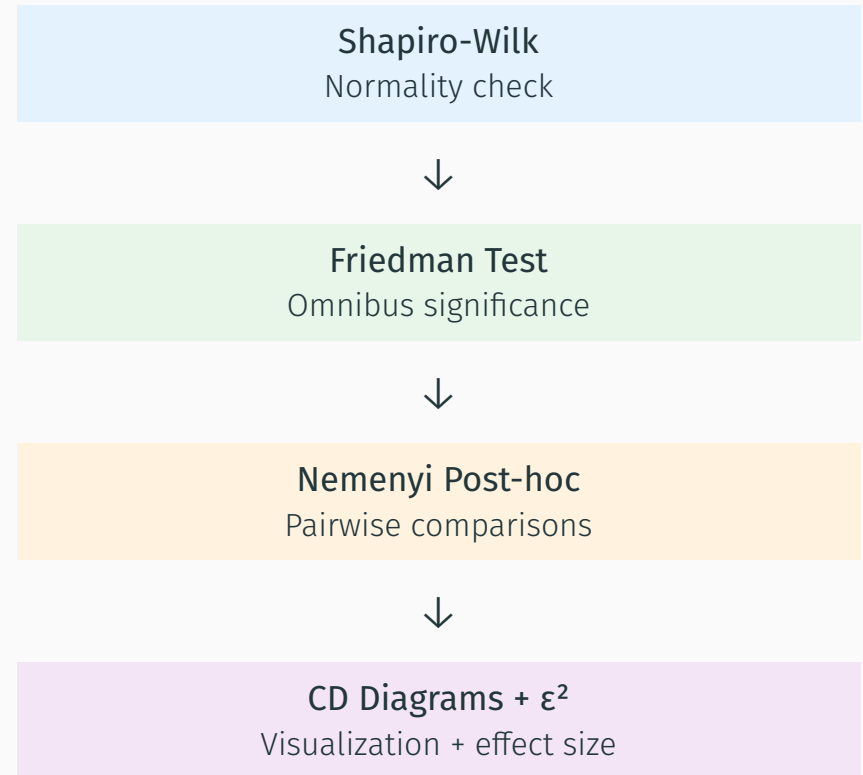
Is the field's preference for masked PPO variants empirically justified compared to other DRL algorithm families?

Algorithm Selection

Algorithm	Family	Masking
MaskablePPO	Policy Gradient	✓
MaskableDQN	Value-Based	✓
Vanilla PPO	Policy Gradient	✗
A2C	Actor-Critic	✗

4 algorithms × 2 real Slurm traces (84k + 28k jobs)

Statistical Framework



Reference: Demšar 2006 methodology

Current Progress

Completed Infrastructure

- ✓ Gymnasium simulation environment
HPCsim with action masking support
- ✓ Classical baseline evaluation
42 combinations (7 selectors × 3 allocators × 2 partitions)
- ✓ Statistical analysis framework
Friedman + Nemenyi pipeline designed
- ✓ Paper writing
Introduction + Literature Review sections complete
- ✓ Visualization pipeline
visualise.py for multi-run analysis

Baseline Results Summary

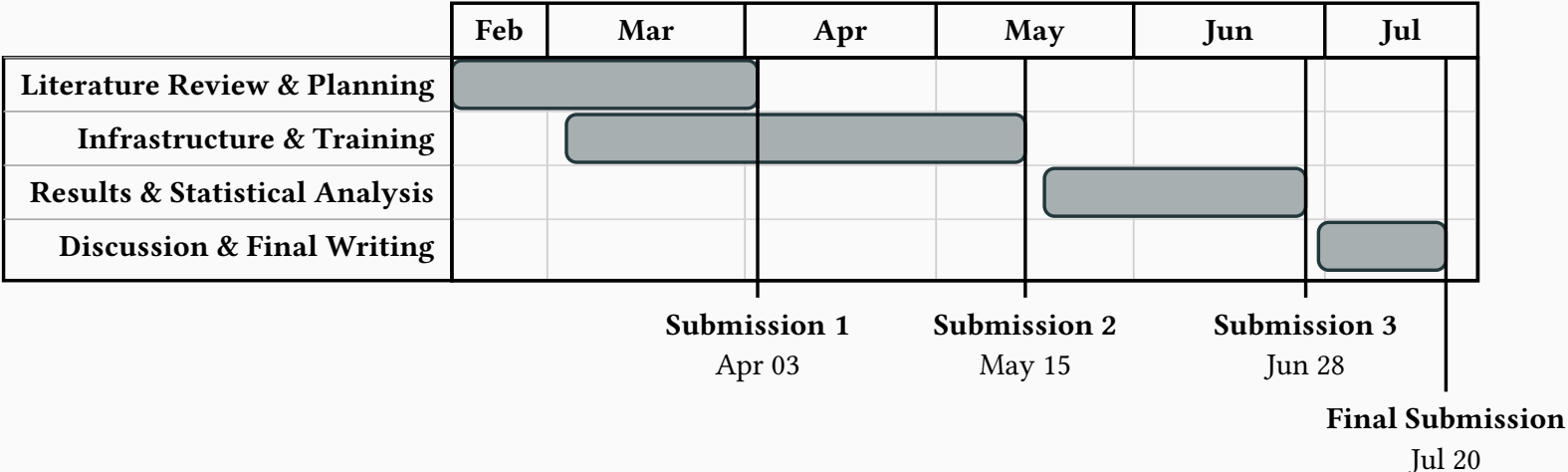
Physical Cluster (84k CPU jobs):

Algorithm	Avg Wait (s)	Slowdown
LCFS + best_fit	1,851	12.71
UNICEP + best_fit	1,955	12.78
SJF + best_fit	1,988	13.32
FCFS + best_fit	2,098	15.21

Reference Point:

DRL must beat 1,851s avg wait to justify complexity

Research Plan & Timeline



Next Steps (6 Weeks - Submission 2)

- 1. Weeks 1-2: Setting up DRL training infrastructure
- 2. Week 3-5: DRL training & setting up statistical analysis (Friedman tests, CD diagrams)
- 3. Weeks 6: Writing Submission 2 & Update Presentation/Website

Computational Resources

- Training Budget:
~400 GPU-hours estimated
- Wall-clock time:
~9 days with parallelization
- GPU allocation:
2-week period (to be confirmed)

Expected Contribution

Primary Contribution

First rigorous statistical comparison of DRL algorithm families for HPC scheduling on real heterogeneous workloads

If PPO wins:

Empirical validation for field's convergence

If alternative wins:

Challenge conventional assumptions

Broader Impact: Foundation for distillation and deployment in resource-constrained environments

Thank You