

UNIVERSITY of the WESTERN CAPE
IYUNIVESITHI yaseNTSHONA KOLONI
UNIVERSITEIT van WES-KAAPLAND

HONOURS RESEARCH · UNIVERSITY OF THE WESTERN CAPE

deep reinforcement learning for hpc job scheduling

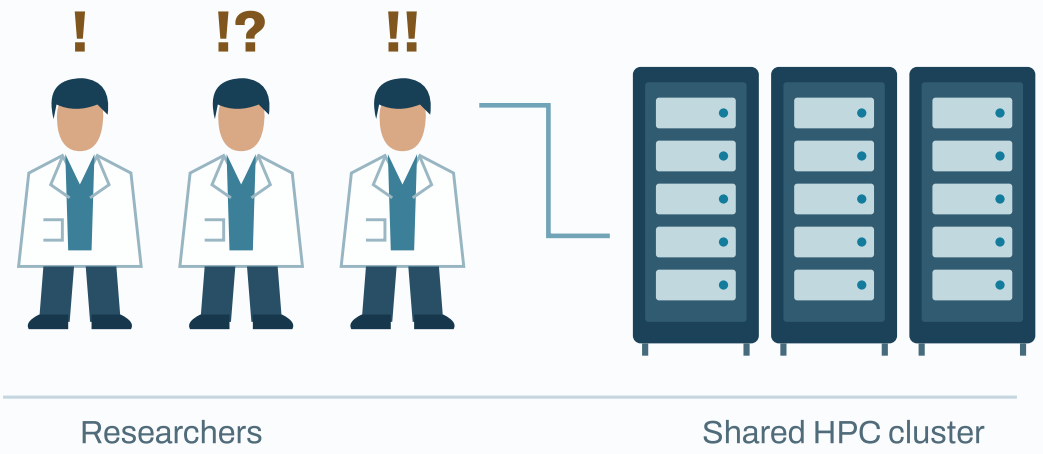
A statistical evaluation

Justin M. Cheney · Supervisor: Prof. Clement N. Nyirenda

This research forms part of a paper published in the SACAIR 2026 online proceedings. [1]

[1] Cheney & Nyirenda (2026). [8] UWC emblem and wordmark.

everyone has a job they want to run.



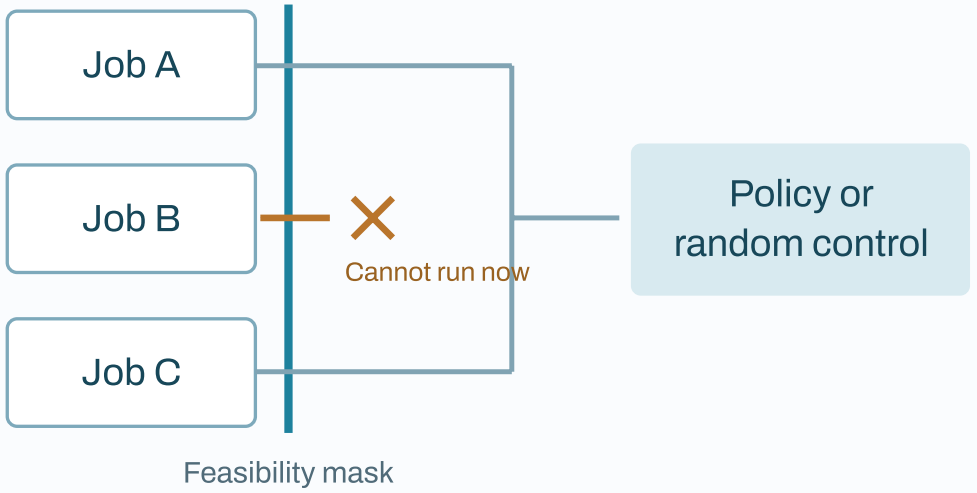
The scheduler decides which job runs next and which resources it receives.

Rules such as First-Come, First-Served and Shortest Job First give jobs different priorities.

Reducing the average wait does not necessarily reduce the longest wait.

[2, 3] Wang et al. (2025); Sutton & Barto (2018). Background and scheduling formulation.

which valid job should the model choose?



DRL learns job selection from reward feedback. I compared PPO, A2C and DQN.

The mask removes infeasible choices. Masked random chooses from the remaining jobs without learning.

Does learning improve those choices enough to justify its cost?

[1, 3] Cheney & Nyirenda (2026); Sutton & Barto (2018). Study design and RL background.

METHODOLOGY

what i compared

PPO, A2C and DQN,
each with and without masking.

6 DRL treatments

Ten seeds per treatment

4 references

SJF, LCFS, UNICEP and masked random

Same allocator. No backfill in the primary comparisons.

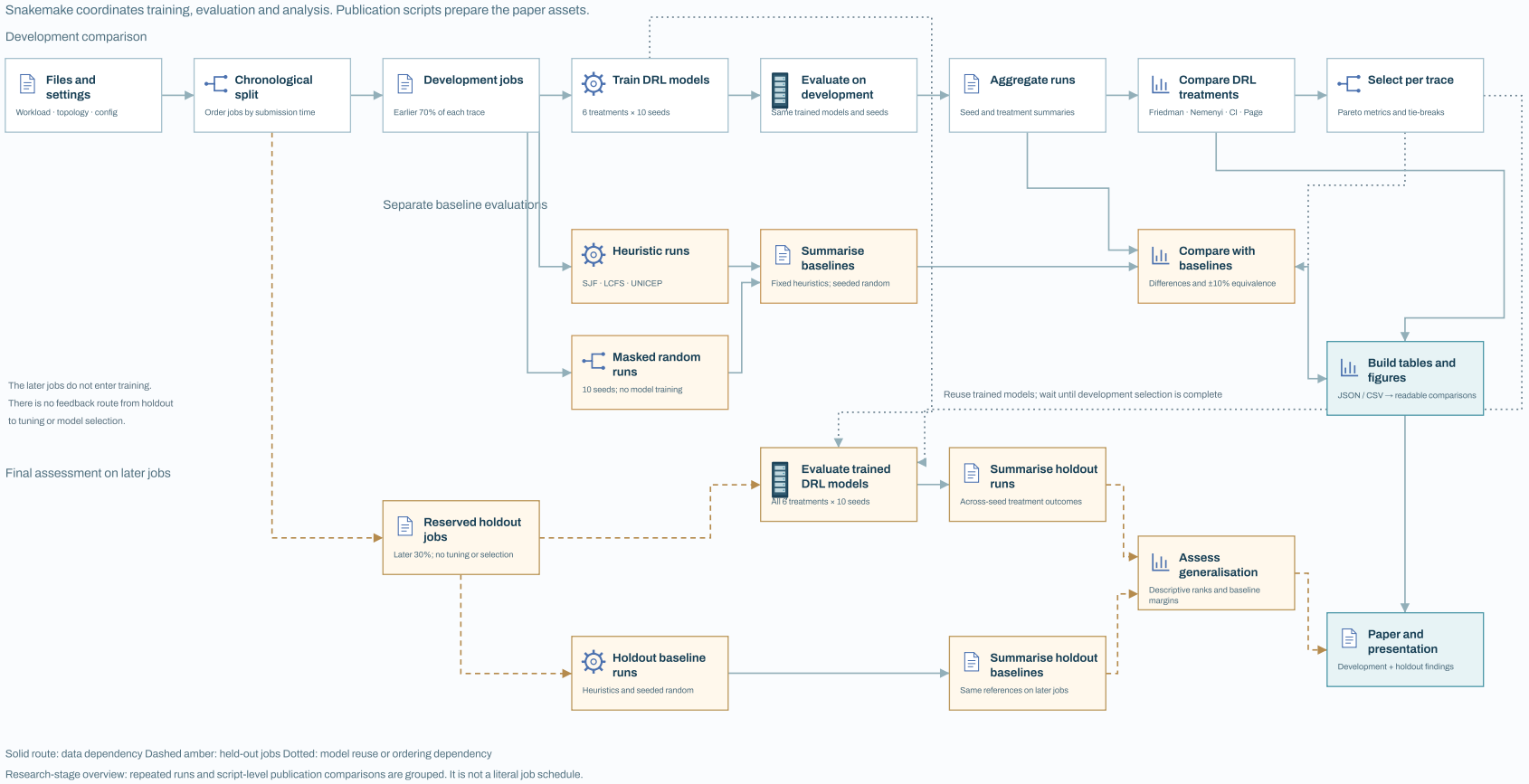
Physical: a homogeneous CPU-only trace.
Deeplearn: a heterogeneous trace including GPUs.



I measured waiting time, bounded slowdown, training time and decision latency.

[1, 2] Study Methodology, Tables 1 and 2; HPCSim: Wang et al. (2025).

how the experiment becomes a set of results



The full research workflow Development comparisons, baseline evaluations and final assessment on later jobs feed the reported findings.

[1, 2, 4–7] Study workflow; HPCSim; rank tests, correction and equivalence methods.

all masked treatments completed evaluation.

Completed evaluation on both traces

Algorithm	With masking	Without masking
PPO	✓	✗
A2C	✓	✗
DQN	✓	✗

The three unmasked treatments stalled and were stopped by the simulator's guard.

Their incomplete results remain in the six-treatment rank tests, which limits what those tests establish.

Finishing still doesn't tell us whether learning helped.

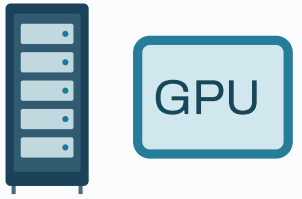
[1] Study Results: completion and truncation reporting conventions.

what did learning add?



Physical

None of the masked DRL treatments improved on random across the three primary metrics. Several comparisons favoured random.



Deeplearn

MaskablePPO improved all three primary metrics against masked random.

That contrast does not establish that GPUs or heterogeneity caused the difference.

[1] Study Results §§4.1 and 4.2: average waiting, average and maximum bounded slowdown.

how did ppo compare?

MaskablePPO led average-waiting ranks among complete treatments on both traces.

Nemenyi comparisons did not find a significant difference from MaskableA2C.

Training time and decision time differ.

Among masked models, MaskablePPO had the lowest recorded mean training time. MaskableDQN had the lowest mean decision latency.

On deeplearn, MaskablePPO was equivalent to UNICEP within ±10% on the three primary metrics.

[1] Study Results, Discussion and training/inference cost tables.

BEFORE I WRAP UP

questions?

I'd like to open the floor.

CLOSING

Before we ask whether a scheduler can learn, we need to ask what it has learned to do, and whether that is better than simply making a valid choice.

Thank you · Justin M. Cheney

references

1. **Cheney, J. M., & Nyirenda, C. N. (2026).** *Deep Reinforcement Learning for HPC Job Scheduling: A Statistical Evaluation.* SACAIR 2026 online proceedings.

2. **Wang, L., Rodriguez, M. A., & Lipovetzky, N. (2025).** Optimizing HPC scheduling: a hierarchical reinforcement learning approach for intelligent job selection and allocation. *The Journal of Supercomputing.* doi:10.1007/s11227-025-07396-3

3. **Sutton, R. S., & Barto, A. G. (2018).** *Reinforcement Learning: An Introduction.* 2nd ed. MIT Press.

4. **Carrasco, J., García, S., Rueda, M. M., Das, S., & Herrera, F. (2020).** Recent trends in the use of statistical tests for comparing swarm and evolutionary computing algorithms: Practical guidelines and a critical review. *Swarm and Evolutionary Computation, 54*, 100665. doi:10.1016/j.swevo.2019.100665

5. **Demšar, J. (2006).** Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research, 7*, 1–30. jmlr.org/papers/v7/demsar06a.html

6. **Holm, S. (1979).** A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65–70.

7. **Lakens, D. (2017).** Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science, 8*(4), 355–362. doi:10.1177/1948550617697177

8. **University of the Western Cape.** Official emblem and trilingual wordmark, from the UWC website. Original artwork retained; locally stored for offline display.

Slide footnotes use these reference numbers. The schematic illustrations are original.