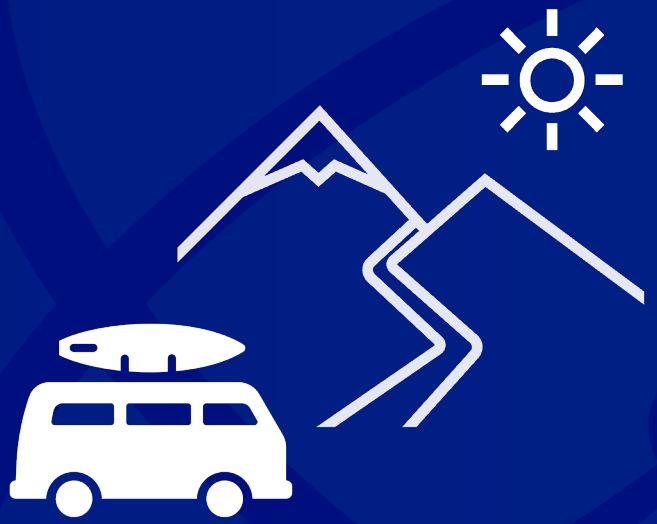
Los Alamos NATIONAL LABORATORY

Are We There Yet?

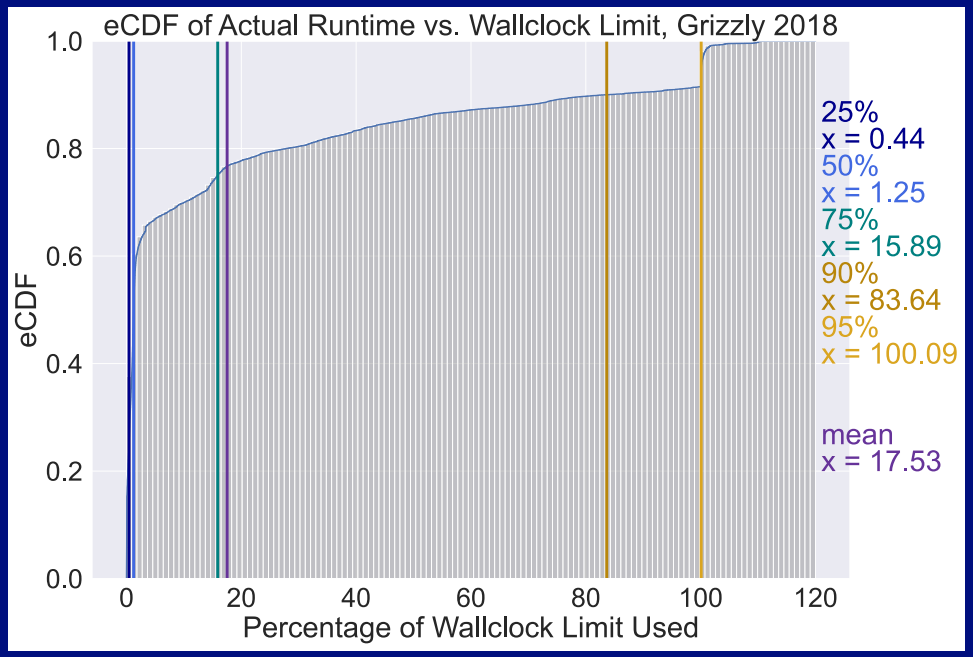
Predicting the Queue Wait times and Job Runtimes for HPC Jobs

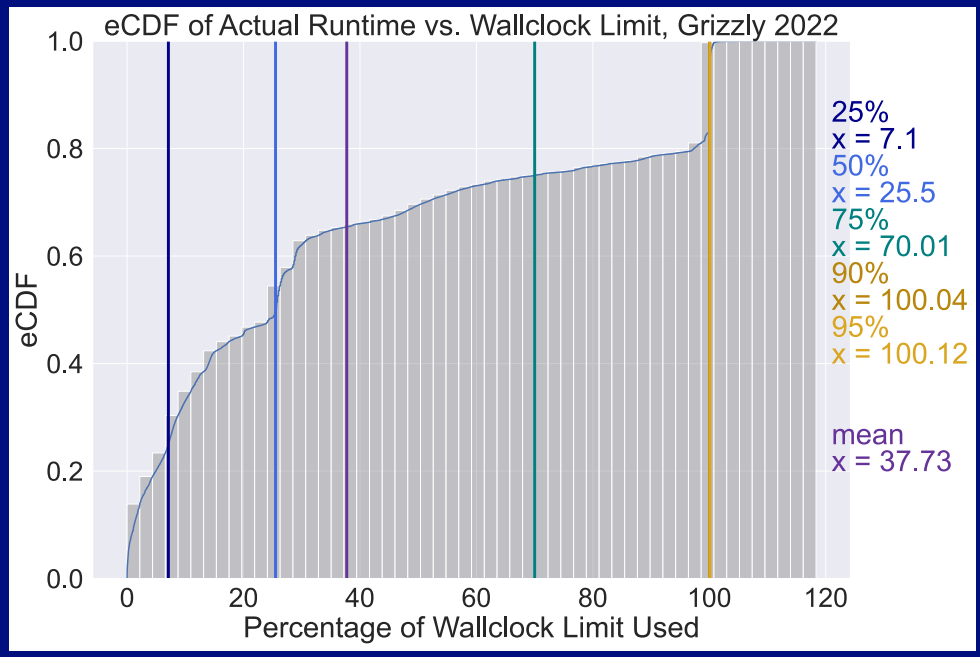
Christin Whitton | HPC-DES

Mentors: Nathan DeBardeleben and
Vanessa Job | HPC-DES



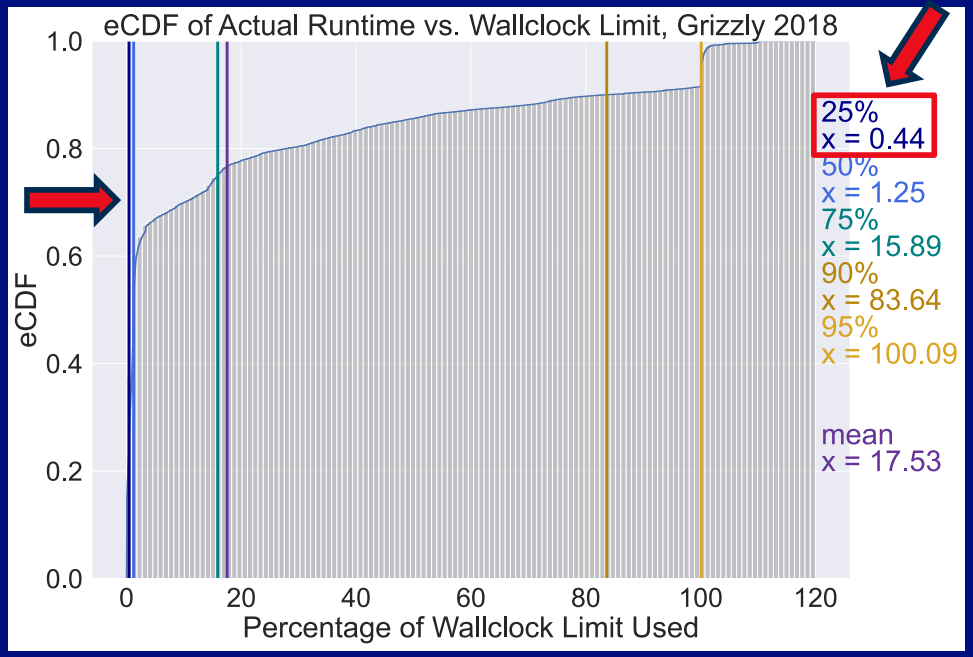
Motivation

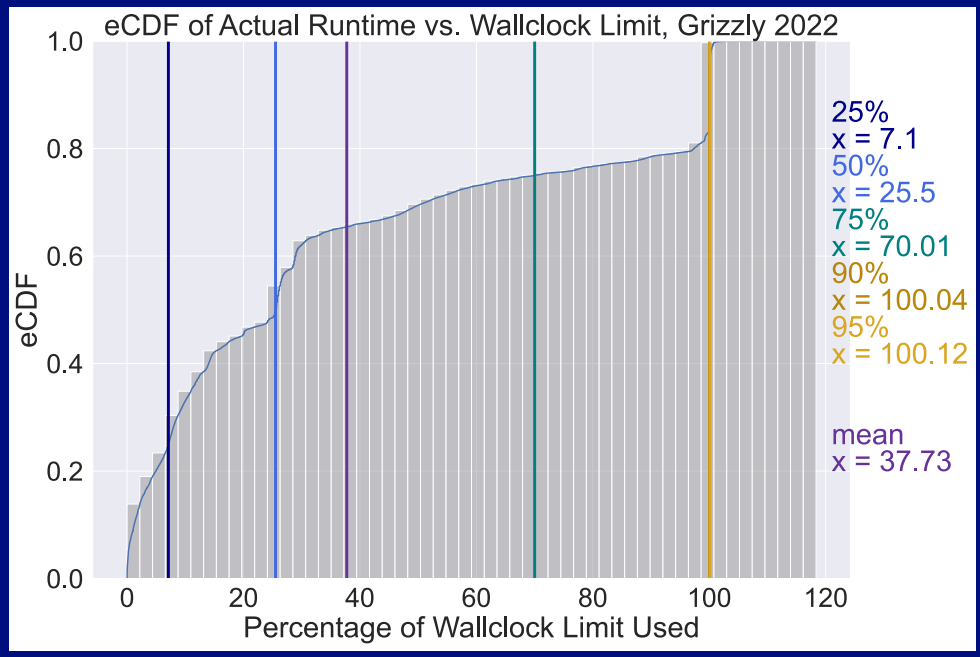




Grizzly Data

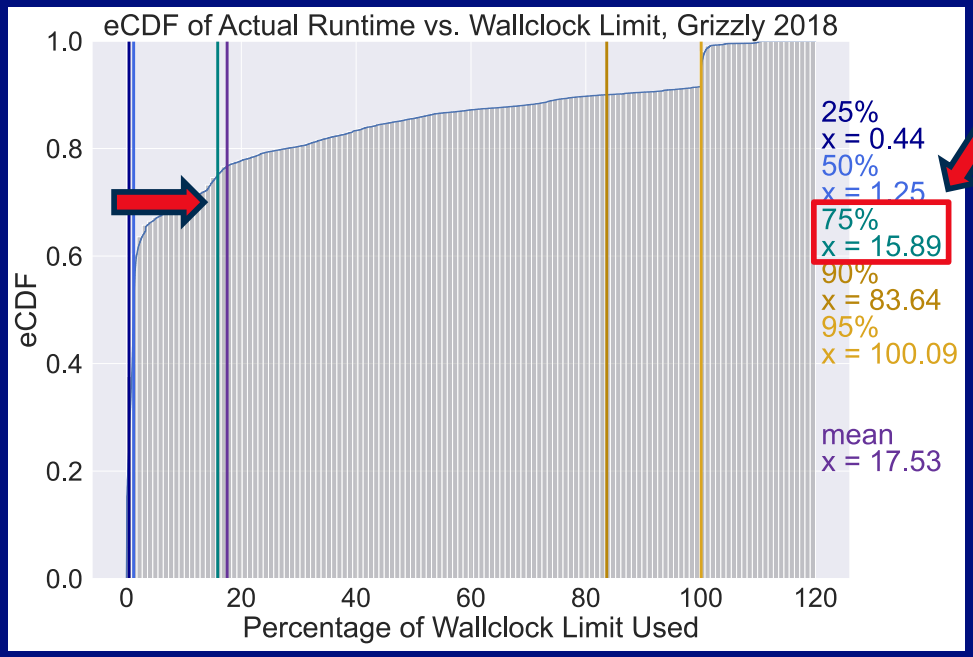
Motivation

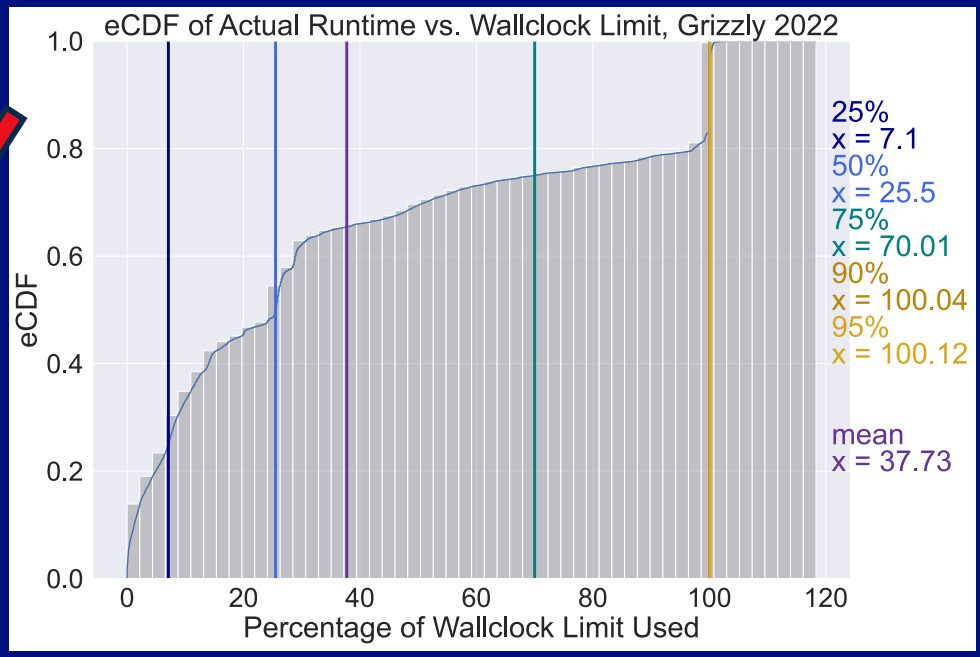




Grizzly Data

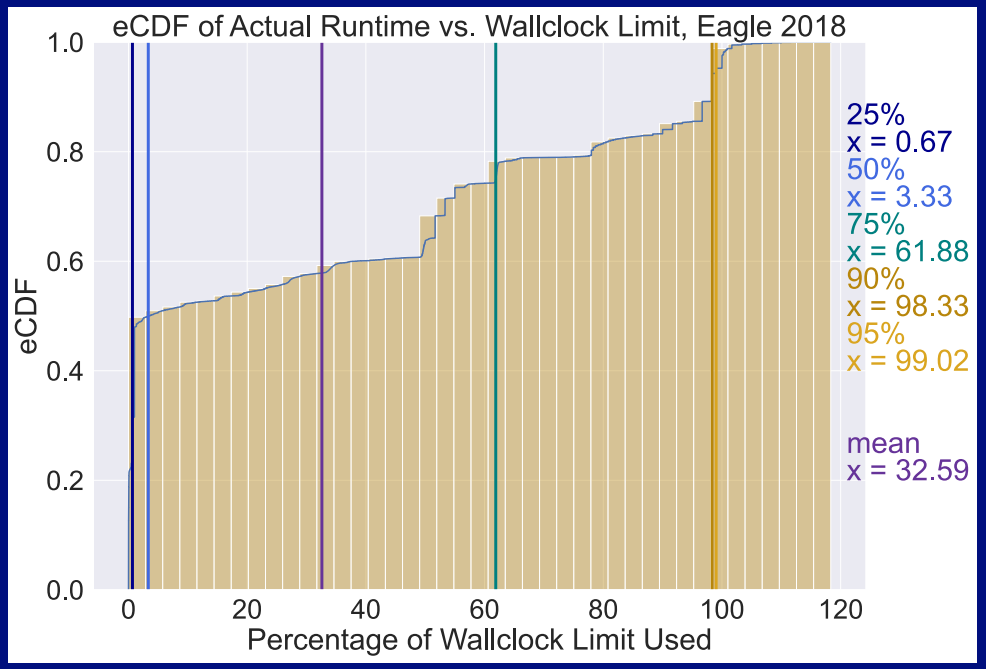
Motivation

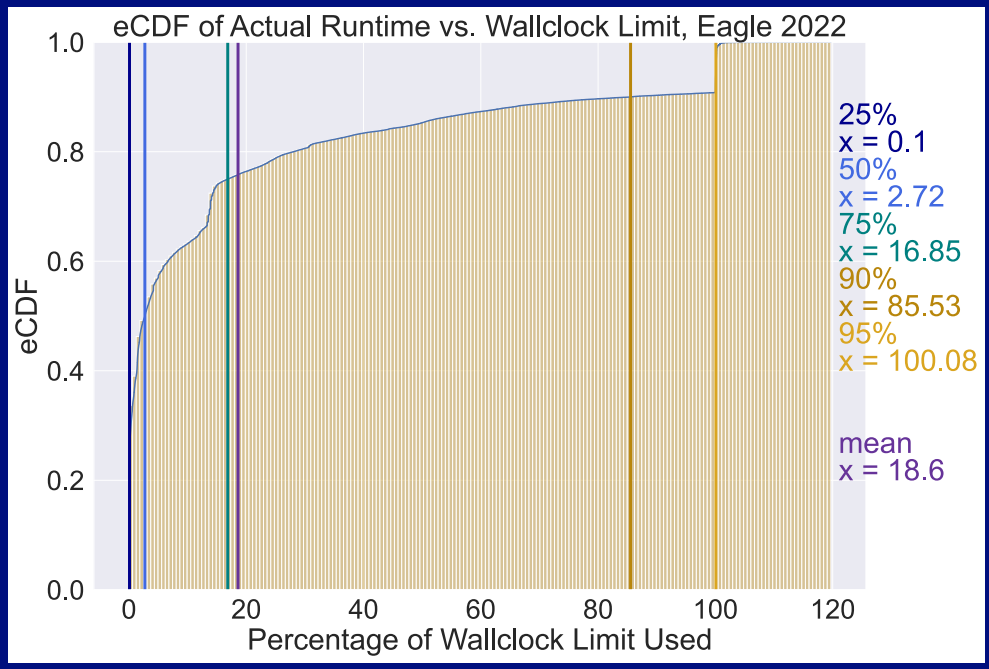




Grizzly Data

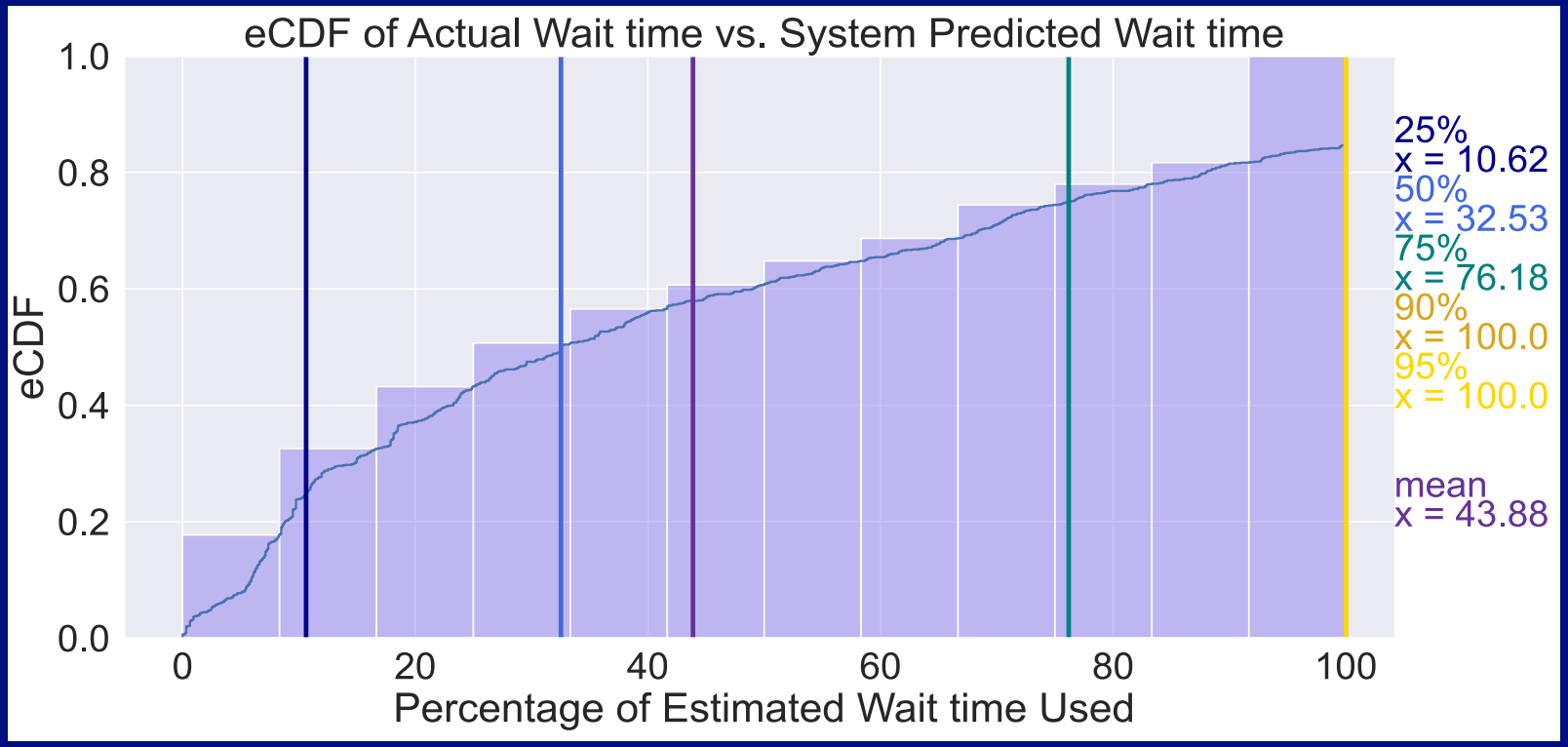
Motivation



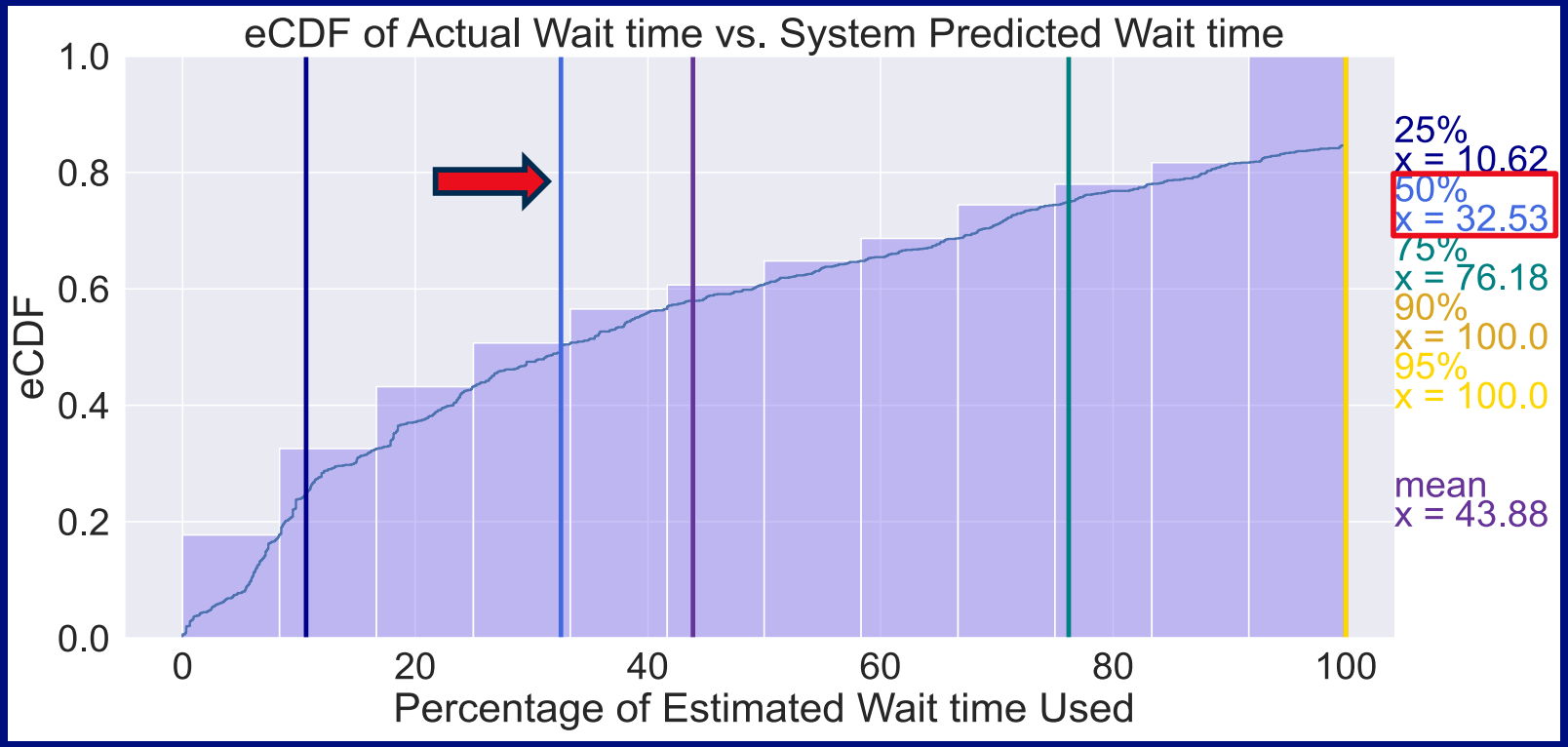


NREL Eagle Data

Job Scheduler Sim Data



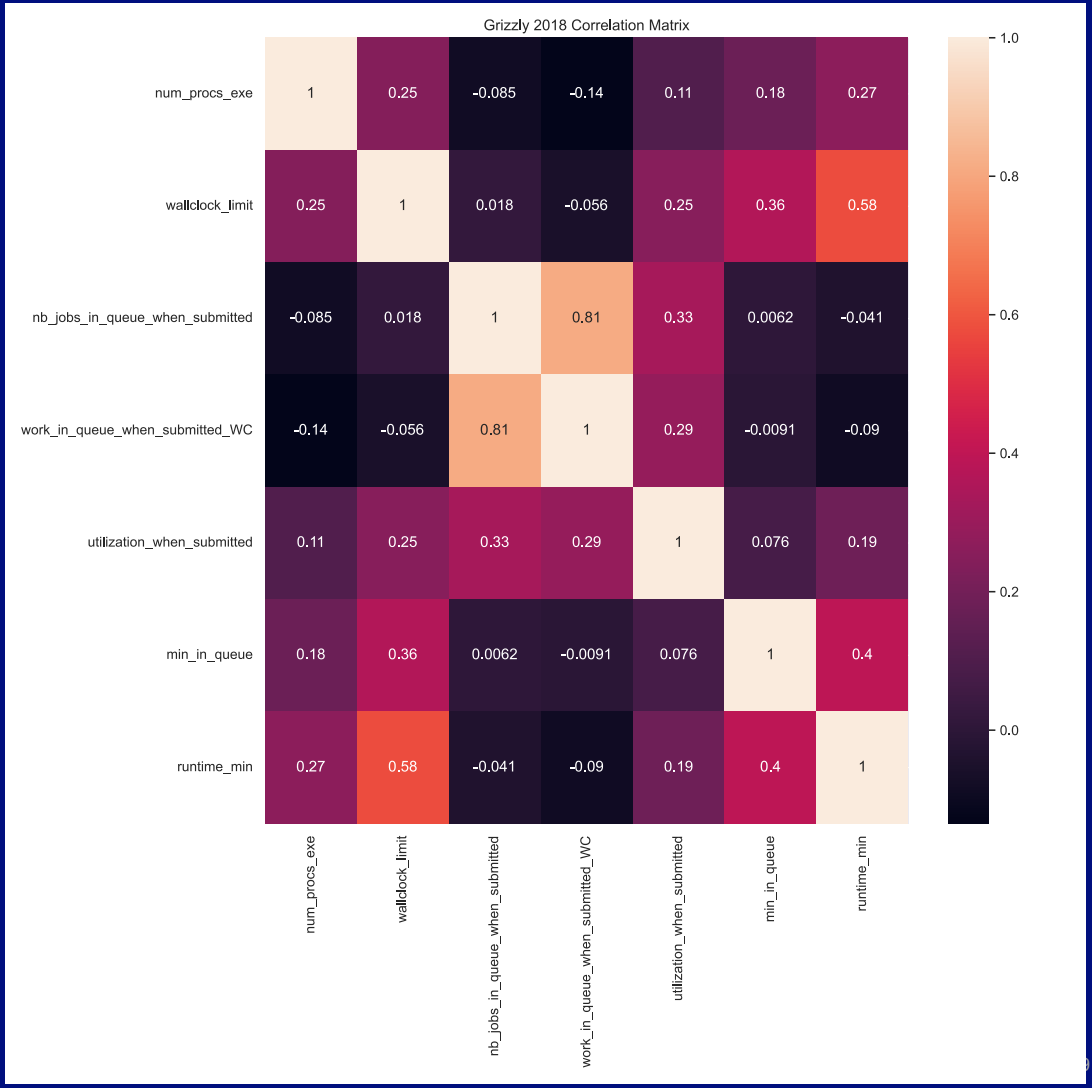
Job Scheduler Sim Data



Why?

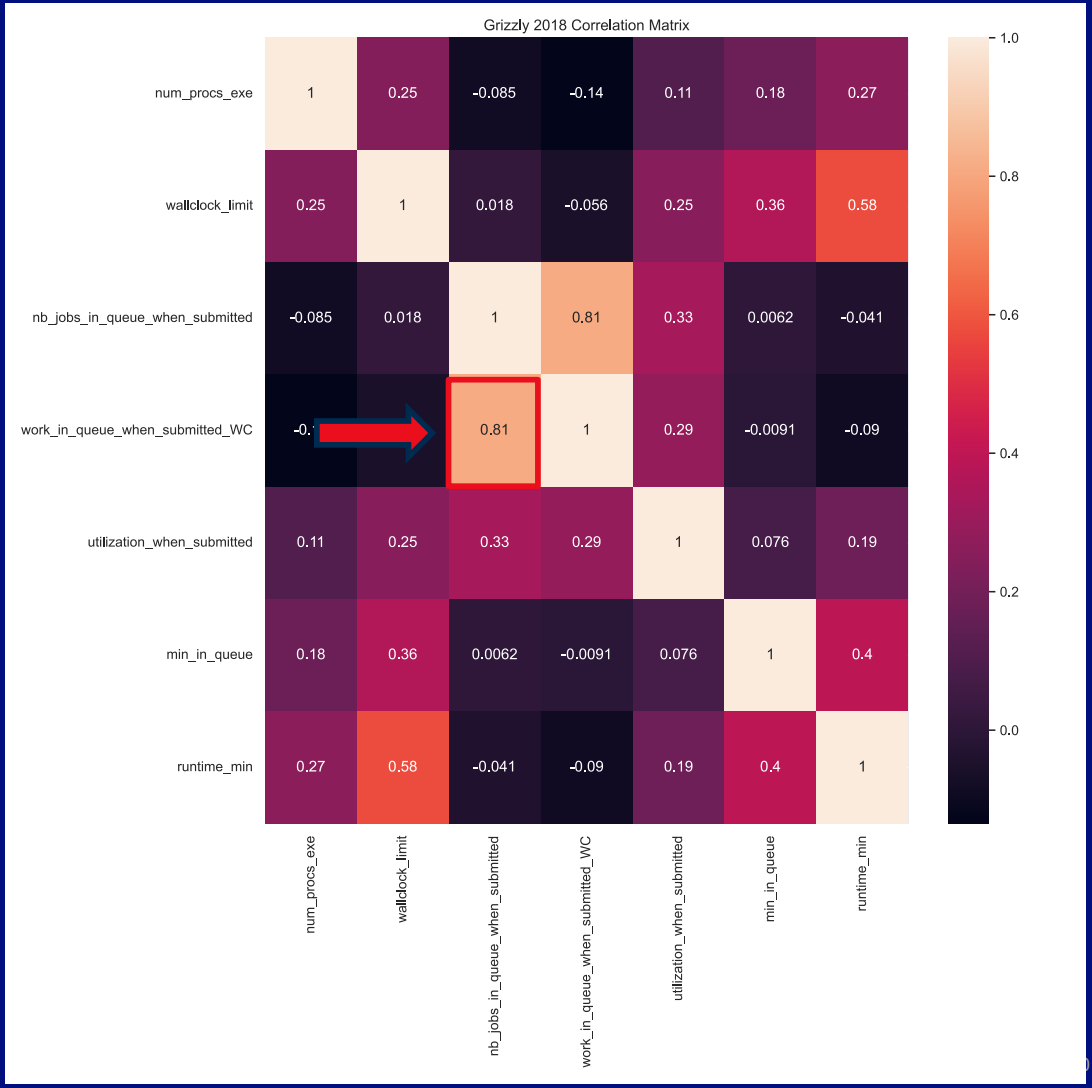
Data

Correlation matrix – how high is too high?



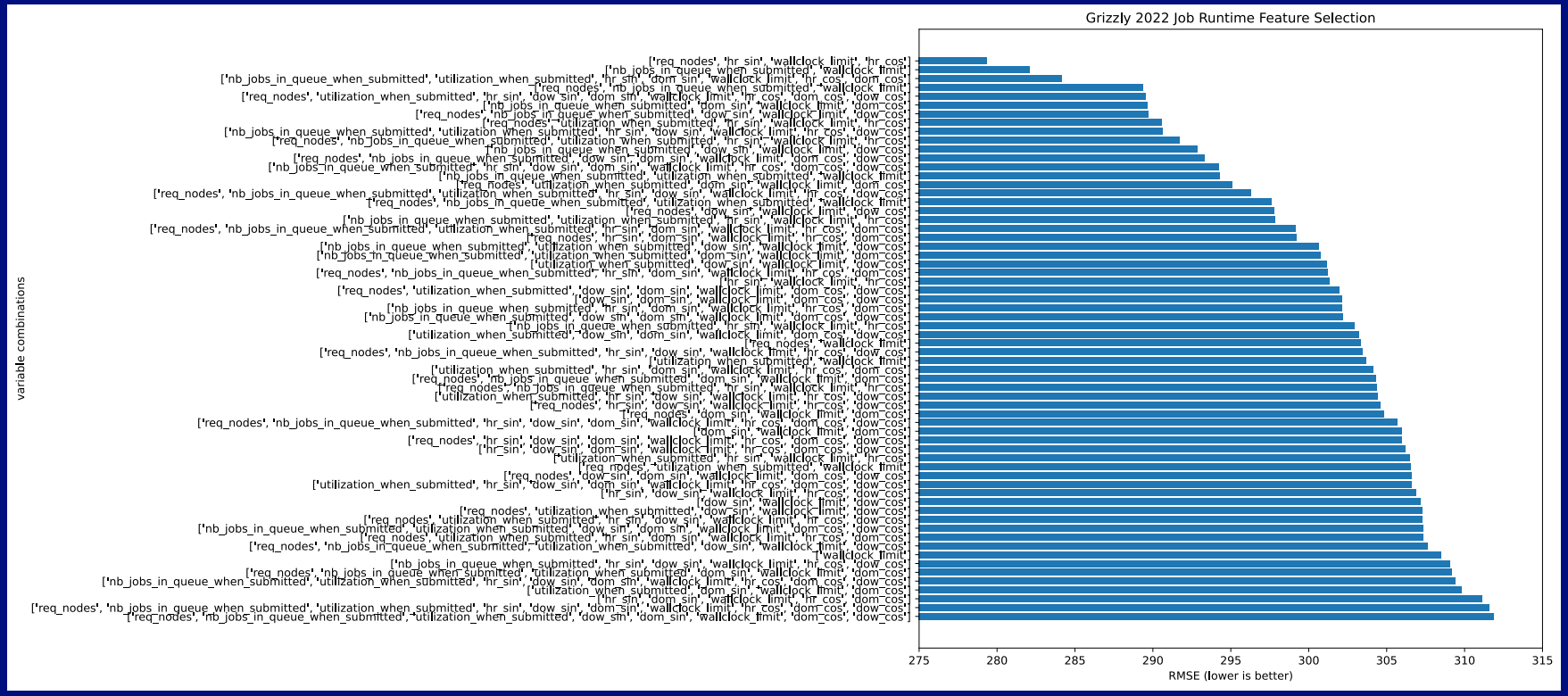
Data

Correlation matrix – how high is too high?

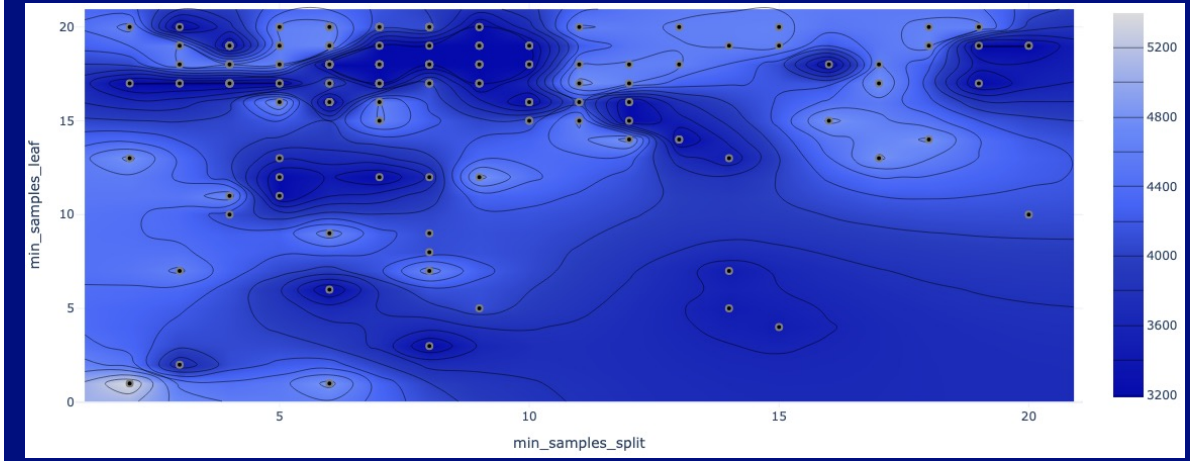


Los Alamos NATIONAL LABORATORY

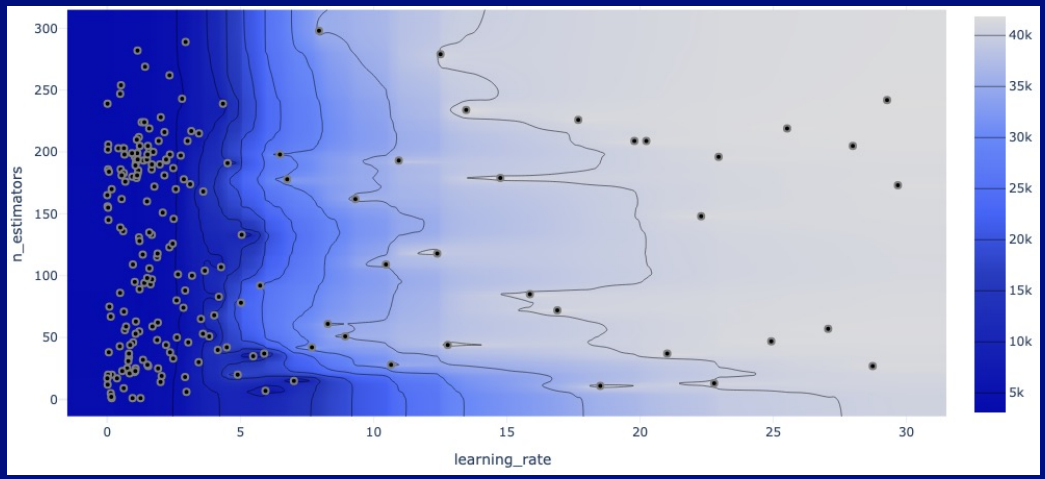
Data


Grizzly 2022 Job Runtime Feature Selection
variable combinations
['req_nodes', 'hr_sin', 'wallclock_limit', 'hr_cos']
['nb_jobs_in_queue_when_submitted', 'wallclock_limit']
['nb_jobs_in_queue_when_submitted', 'utilization_when_submitted', 'hr_sin', 'dom_sin', 'wallclock_limit', 'hr_cos', 'dom_cos']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'wallclock_limit']
['req_nodes', 'utilization_when_submitted', 'hr_sin', 'dow_sin', 'dom_sin', 'wallclock_limit', 'hr_cos', 'dom_cos', 'dow_cos']
['nb_jobs_in_queue_when_submitted', 'dom_sin', 'wallclock_limit', 'dom_cos']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'dow_sin', 'wallclock_limit', 'dow_cos']
['req_nodes', 'utilization_when_submitted', 'hr_sin', 'wallclock_limit', 'hr_cos']
['nb_jobs_in_queue_when_submitted', 'utilization_when_submitted', 'hr_sin', 'dow_sin', 'wallclock_limit', 'hr_cos', 'dow_cos']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'utilization_when_submitted', 'hr_sin', 'wallclock_limit', 'hr_cos']
['nb_jobs_in_queue_when_submitted', 'dow_sin', 'wallclock_limit', 'dow_cos']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'dow_sin', 'dom_sin', 'wallclock_limit', 'dom_cos', 'dow_cos']
['nb_jobs_in_queue_when_submitted', 'hr_sin', 'dow_sin', 'dom_sin', 'wallclock_limit', 'hr_cos', 'dow_cos']
['nb_jobs_in_queue_when_submitted', 'utilization_when_submitted', 'wallclock_limit']
['req_nodes', 'utilization_when_submitted', 'dom_sin', 'wallclock_limit', 'dom_cos']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'utilization_when_submitted', 'hr_sin', 'dow_sin', 'wallclock_limit', 'hr_cos', 'dow_cos']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'utilization_when_submitted', 'wallclock_limit']
['req_nodes', 'dow_sin', 'wallclock_limit', 'dow_cos']
['nb_jobs_in_queue_when_submitted', 'utilization_when_submitted', 'hr_sin', 'wallclock_limit', 'hr_cos']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'hr_sin', 'dom_sin', 'wallclock_limit', 'hr_cos', 'dom_cos']
['req_nodes', 'hr_sin', 'dom_sin', 'wallclock_limit', 'hr_cos', 'dom_cos']
['nb_jobs_in_queue_when_submitted', 'utilization_when_submitted', 'dow_sin', 'wallclock_limit', 'dow_cos']
['nb_jobs_in_queue_when_submitted', 'utilization_when_submitted', 'dom_sin', 'wallclock_limit', 'dom_cos']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'hr_sin', 'dom_sin', 'wallclock_limit', 'hr_cos', 'dom_cos', 'dow_cos']
['hr_sin', 'wallclock_limit', 'hr_cos']
['req_nodes', 'utilization_when_submitted', 'dow_sin', 'dom_sin', 'wallclock_limit', 'dom_cos', 'dow_cos']
['dow_sin', 'dom_sin', 'wallclock_limit', 'dom_cos', 'dow_cos']
['nb_jobs_in_queue_when_submitted', 'hr_sin', 'dom_sin', 'wallclock_limit', 'hr_cos', 'dom_cos']
['nb_jobs_in_queue_when_submitted', 'dow_sin', 'dom_sin', 'wallclock_limit', 'dom_cos', 'dow_cos']
['nb_jobs_in_queue_when_submitted', 'hr_sin', 'wallclock_limit', 'hr_cos']
['utilization_when_submitted', 'dow_sin', 'dom_sin', 'wallclock_limit', 'dom_cos', 'dow_cos']
['req_nodes', 'wallclock_limit']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'hr_sin', 'dow_sin', 'wallclock_limit', 'hr_cos', 'dow_cos']
['utilization_when_submitted', 'wallclock_limit']
['utilization_when_submitted', 'hr_sin', 'dom_sin', 'wallclock_limit', 'hr_cos', 'dom_cos']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'dom_sin', 'wallclock_limit', 'dom_cos']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'hr_sin', 'wallclock_limit', 'hr_cos']
['utilization_when_submitted', 'hr_sin', 'dow_sin', 'wallclock_limit', 'hr_cos', 'dow_cos']
['req_nodes', 'hr_sin', 'dow_sin', 'wallclock_limit', 'hr_cos', 'dow_cos']
['req_nodes', 'dom_sin', 'wallclock_limit', 'dom_cos']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'dow_sin', 'dom_sin', 'wallclock_limit', 'dom_cos', 'dow_cos']
['dom_sin', 'wallclock_limit', 'dom_cos']
['req_nodes', 'hr_sin', 'dow_sin', 'dom_sin', 'wallclock_limit', 'hr_cos', 'dom_cos', 'dow_cos']
['hr_sin', 'dow_sin', 'dom_sin', 'wallclock_limit', 'hr_cos', 'dom_cos', 'dow_cos']
['utilization_when_submitted', 'hr_sin', 'wallclock_limit', 'hr_cos']
['req_nodes', 'utilization_when_submitted', 'wallclock_limit']
['req_nodes', 'dow_sin', 'dom_sin', 'wallclock_limit', 'dom_cos', 'dow_cos']
['utilization_when_submitted', 'hr_sin', 'dow_sin', 'dom_sin', 'wallclock_limit', 'hr_cos', 'dom_cos', 'dow_cos']
['hr_sin', 'dow_sin', 'wallclock_limit', 'hr_cos', 'dow_cos']
['dow_sin', 'wallclock_limit', 'dow_cos']
['req_nodes', 'utilization_when_submitted', 'dow_sin', 'wallclock_limit', 'dow_cos']
['req_nodes', 'utilization_when_submitted', 'hr_sin', 'dow_sin', 'wallclock_limit', 'hr_cos', 'dow_cos']
['nb_jobs_in_queue_when_submitted', 'utilization_when_submitted', 'dow_sin', 'dom_sin', 'wallclock_limit', 'dom_cos', 'dow_cos']
['req_nodes', 'utilization_when_submitted', 'hr_sin', 'dom_sin', 'wallclock_limit', 'hr_cos', 'dom_cos']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'utilization_when_submitted', 'dow_sin', 'wallclock_limit', 'dow_cos']
['wallclock_limit']
['nb_jobs_in_queue_when_submitted', 'hr_sin', 'dow_sin', 'wallclock_limit', 'hr_cos', 'dow_cos']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'utilization_when_submitted', 'dom_sin', 'wallclock_limit', 'dom_cos']
['nb_jobs_in_queue_when_submitted', 'utilization_when_submitted', 'hr_sin', 'dow_sin', 'dom_sin', 'wallclock_limit', 'hr_cos', 'dom_cos', 'dow_cos']
['utilization_when_submitted', 'dom_sin', 'wallclock_limit', 'dom_cos']
['hr_sin', 'dom_sin', 'wallclock_limit', 'hr_cos', 'dom_cos']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'utilization_when_submitted', 'hr_sin', 'dow_sin', 'dom_sin', 'wallclock_limit', 'hr_cos', 'dom_cos', 'dow_cos']
['req_nodes', 'nb_jobs_in_queue_when_submitted', 'utilization_when_submitted', 'dow_sin', 'dom_sin', 'wallclock_limit', 'dom_cos', 'dow_cos']
275
280
285
290
295
300
305
310
315
RMSE (lower is better)

Model Tuning



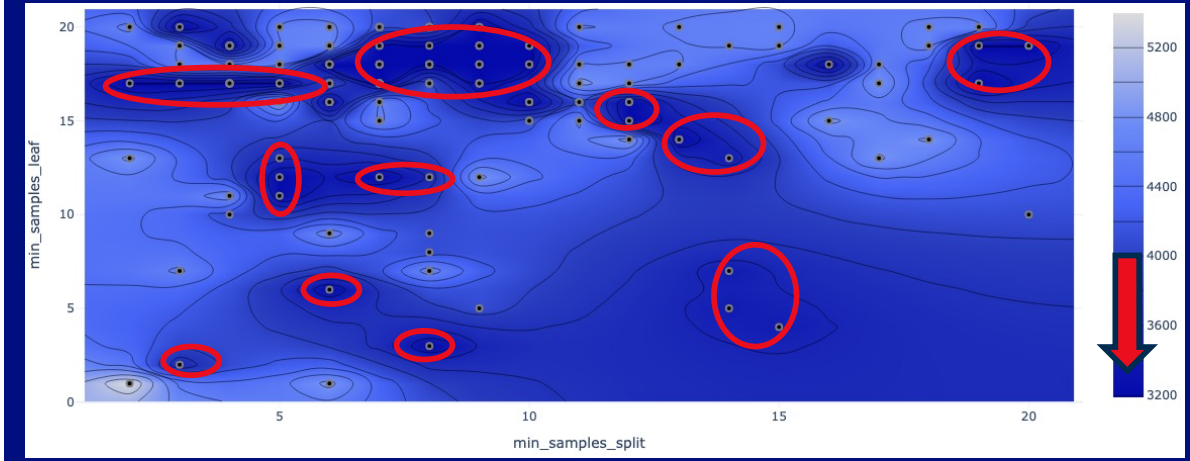
Decision Tree Optuna optimization



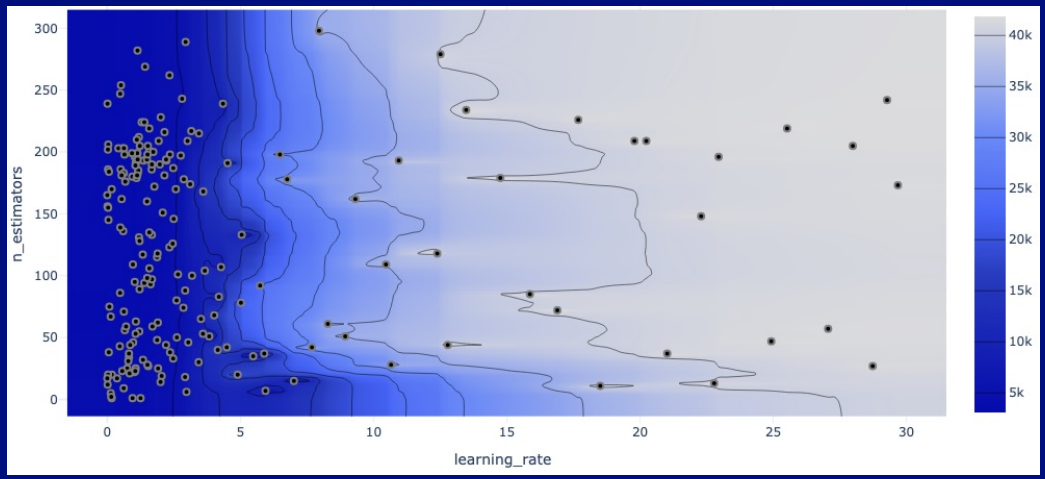
Adaboost Optuna Optimization

Los Alamos
NATIONAL LABORATORY

Model Tuning



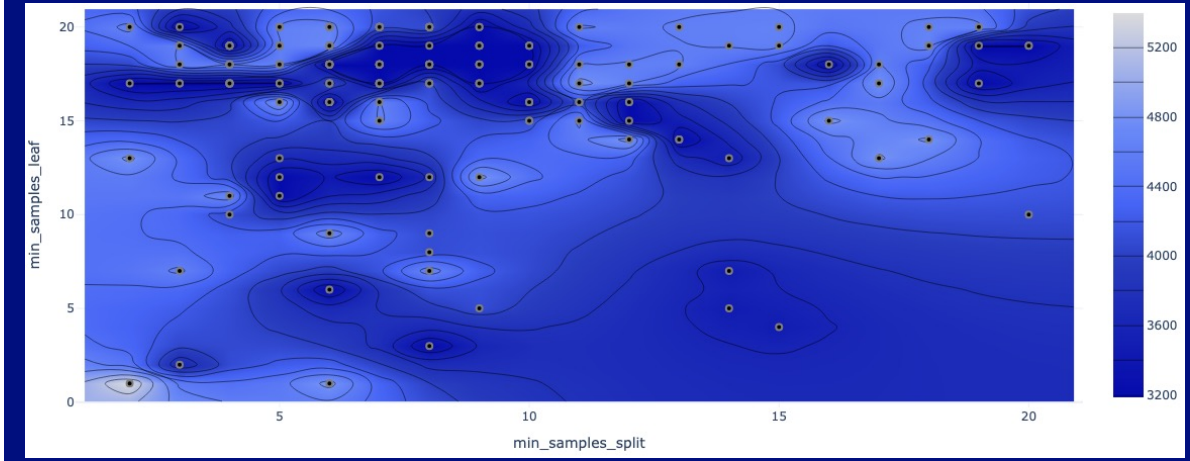
Decision Tree Optuna optimization



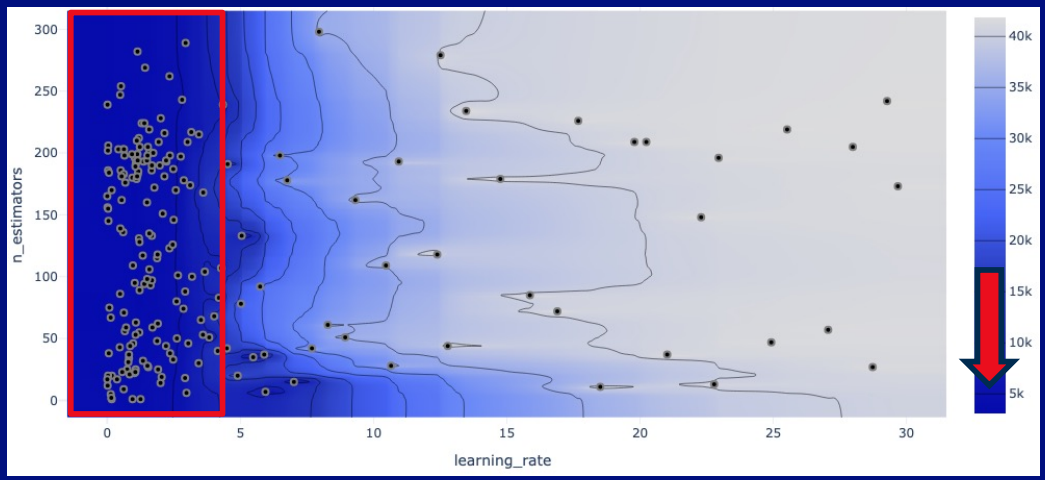
Adaboost Optuna Optimization

Los Alamos
NATIONAL LABORATORY

Model Tuning



Decision Tree Optuna optimization



Adaboost Optuna Optimization

Los Alamos
NATIONAL LABORATORY

Results

RMSE by Data, Target Variable and Machine Learning Method		DATA/TARGET VARIABLE			
		Grizzly 2018 Minutes in Queue	Grizzly 2022 Minutes in Queue	Grizzly 2018 Runtime Minutes	Grizzly 2022 Runtime Minutes
Machine Learning Method	Decision Tree	3185.97	1260.51	284.54	249.19
	Random Forest	3207.85	1315.86	275.47	240.76
	SVR (Linear)	3605.69	1505.42	280.38	310.46
	SVR (RBF)	4417.22	1536.7	284.64	318.69
	Adaboost	3104.81	1262.46	277.30	228.01
	XGBoost	3132.14	1298.98	272.43	248.63

Future Work

- Look at ensembles of models
- Job Scheduler simulation data – model at each event
- More data!
- Why did SVR perform poorly? Other ML methods?
- (eventually) User Interface

Questions?