PNNL-36729

Wholesale Electricity Analysis via Simulation & Learning Experiments (WEASLE)

Platform Development and Pilot Competition

September 2024

1 Brent Eldridge
2 Matt Cornachione
3 Liping Li
4 Jesse Holzer
5 Arun Veeramany

U.S. DEPARTMENT OF ENERGY
Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from
the Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062
www.osti.gov
ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: http://www.ntis.gov

PNNL-36729

Wholesale Electricity Analysis via Simulation & Learning Experiments (WEASLE)

Platform Development and Pilot Competition

September 2024

1 Brent Eldridge
2 Matt Cornachione
3 Liping Li
4 Jesse Holzer
5 Arun Veeramany

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Abstract

This document reports the development of the Wholesale Electricity Analysis via Simulation and Learning Experiments (WEASLE) platform and the pilot competition that was conducted to test the platform. Due to the increasing reliance on variable renewable energy resources for bulk power, the pilot competition, called the Energy Storage Participation Algorithm Competition (ESPA-Comp), was used to test the effect of various market designs on storage utilization and market efficiency. Basic details of the platform are provided, including an overview of the market clearing engine, the battery dispatch and degradation models, electric grid topology and resource mix, and software architecture. Two market designs were tested: a two-settlement market analogous to typical ISO design today, and a multi-settlement market that allows additional forward-trading periods during the real-time market. Results from the pilot competition show that the storage bidding problem is nontrivial and is well suited for future challenges. We find that: 1) all four teams utilized different approaches to the bidding problem, 2) different methodological approaches led to substantially different offer behaviors, 3) resource profits are clustered by team and methodological approach, 4) simulated offers reduced market surplus by about 0.5%, 5) substantially different prices between two-settlement and multi-settlement markets albeit minimal difference in overall market surplus.

Acknowledgments

This research was supported by the Energy Mission Seed Investment under the Laboratory Directed Research and Development (LDRD) Program at Pacific Northwest National Laboratory (PNNL). PNNL is a multi-program national laboratory operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under Contract No. DE-AC05-76RL01830.

We also thank additional contributors to the project. PNNL staff that contributed to the project included Brittany Tarufelli, Kostas Oikonomou, Li He, Abhishek Somani, Bryan Huang, Mark Weimar, and Hayden Reeve. Our external advisors included Anthony Giacomoni (PJM), Richard O'Neill, Benjamin Hobbs (Johns Hopkins University), Jacob Mays (Cornell University), and Robin Hytowitz (EPRI). Finally, external participants of the pilot competition included Yanchao Liu (Wayne State University), Zhenyu Zhou (Wayne State University), Rajni Bansal (University of California, San Diego), and Zhirui Liang (Johns Hopkins University).

Contents

Figures

Tables

1.0 Introduction

Market design has emerged as an important tool to ensure reliable, efficient, and secure transition to a high-renewable electric grid (Milligan *et al.* 2016, Ela *et al.*, Perez-Arriaga, Jenkins, and Batlle 2017, Litvinov, Zhao, and Zheng 2019, Eldridge and Somani 2022a). In the US, state policies continue to push wholesale energy markets to integrate higher proportions of clean and renewable energy (Barbose 2023). While regulatory and policy mandates can push the grid integration of new resources, efficient market design can ensure that it is done at least cost and with minimal disruptions to grid resilience and reliability.

Battery energy storage systems (BESS) are likely to play a key role, as it allows meeting higher regulatory renewable energy targets with less renewable investment and less renewable curtailment (Arbabzadeh *et al.* 2019). However, BESS integration with bulk grid market operations is still developing (Haas *et al.* 2022). To support rigorous assessment of new market design proposals, this paper describes a new simulation platform called Wholesale Analysis via Simulation and Learning Experiments (WEASLE) and an associated pilot competition called the Energy Storage Participation Algorithm Competition (ESPA-Comp).

Batteries profit in wholesale energy markets through three main channels. First, batteries ca provide arbitrage, i.e., purchasing energy for charge at a lower cost than the revenue from energy discharge (Salles *et al.* 2017). Today, few batteries are profitable through energy arbitrage alone. A second revenue stream comes from providing ancillary services (Chen *et al.* 2010). Ancillary services can differ from market to market but are typically reserve products, which provides reliable backup service in case of system contingencies, or regulation service, which is used to balance the second-to-second fluctuations (Cramton 2017). Although many BESS profit primarily through ancillary services today, continued investment in BESS has the potential to collapse ancillary service prices (Salles *et al.* 2017, Bhatnagar *et al.* 2013). Energy arbitrage may therefore become a crucial part of BESS operation and grid stability in high-renewable electric grids (Arbabzadeh *et al.* 2019).

Today's electricity markets use a two-settlement (TS) market design, in which generation schedules are scheduled and receive a forward financial settlement at the day-ahead price for their planned production schedule in the day-ahead market (Crampton 2019). The difference between a resource's day-ahead position and their actual production is subsequently settled in the real-time market at the real-time price. This results in exactly two settlements per resource: first in the day-ahead market and once again in the real-time market.

The TS market was designed to best suit the needs of conventional generators (Eldridge and Somani 2022b). Many of these generators have long start up times and must commit their plants 8 to 24 hours before power is needed. Day-ahead markets are designed for this purpose. Storage systems have different constraints, however. They are typically much faster to respond to changes in power demand, but many BESS are limited to a maximum of about 4 hours of discharge time when fully charged. Alternative market structures may be better suited to planning battery operations, but high quality market simulations and studies are required to make any more concrete recommendations.

1.1 Contribution

We propose a novel electricity market modeling platform that allows multiple users to interact through realistic market processes in a competitive environment. Market modeling methods typically fall into one of three categories: optimization, equilibrium, or simulation (Haugen *et al.* 2024). Each method has strengths and weaknesses which need to be appropriately harnessed or mitigated depending on the scope of research. Optimization-based methods attempt to solve system dispatch decisions to optimality. However, this method has difficulty including strategic decisions that market participants might make, and it therefore might overestimate the benefits of newly proposed market designs. Equilibrium approaches model the strategic incentives of market participants, and therefore holds promise to analyze market design issues. However, because equilibrium problems can often be orders of magnitude more complex than optimization problems, there is a limited scope of problems that can analyzed via equilibrium methods. Instead, equilibrium methods often rely on simplifications of the underlying system, for example, simplified technology constraints, representative agents, reduced decision space, or relaxed equilibrium conditions. In contrast, simulation-based methods allow a high degree of modeling detail regarding both technological constraints and strategic incentives.

Overall, the simulation platform that we've developed, called WEASLE, mainly falls into the simulation category. However, it is composed of modules that each utilize a different method. The first main component is the market clearing engine. We developed an optimization-based market clearing model in GAMS (GAMS 2022). As shown in Figure 1, the market clearing engine takes inputs from the electric topology, generation supply, and consumer demand from the physical system as well as market design features such as the energy and ancillary service product definitions and the assumed trade frequency and horizon. This aspect of the platform models the standard security-constrained unit commitment problem, a mixed-integer linear programming (MILP) problem that is solved in typical production cost modeling software found throughout industry and the research community (Hobbs 2001, Holzer *et al.* 2024).

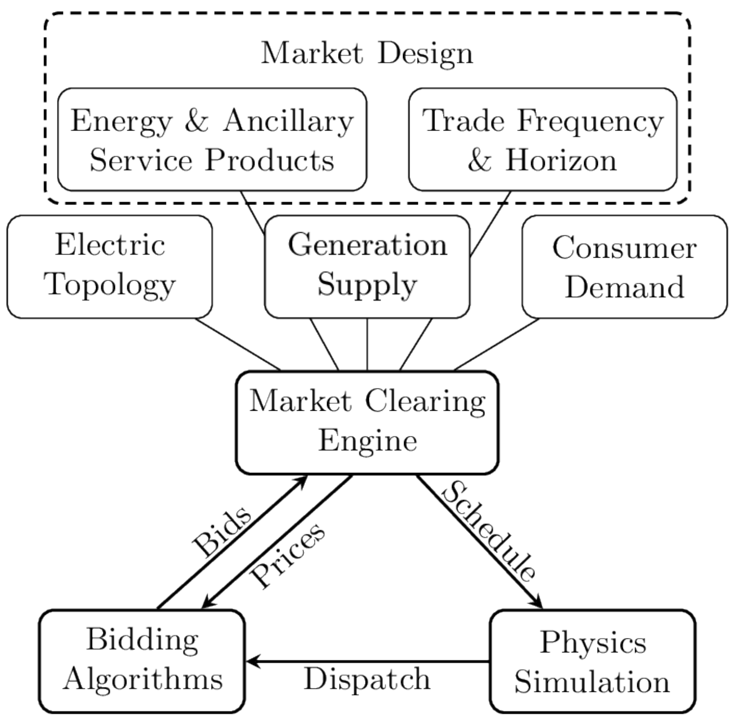

Figure 1: WEASLE Platform Overview

Second, the platform includes a physics simulation of BESS dispatch and degradation cost for a lithium-ion battery system. Our model implements the zero-order equivalent circuit model to simulate nonlinear charging and discharging efficiencies and a degradation cost calculation, both based on formulations in Rosewater (2019). We include these facets of BESS operations separately from the market clearing engine; because the simulated efficiency and degradation costs are nonlinear, they cannot be cast within the MILP problem formulation framework used for solving SCUC. These features of BESS dispatch must therefore be approximated and entered into the market clearing engine through storage bidding parameters.

Last, the platform includes a module for bidding algorithms for resources that participate in the simulated market. The bidding problem is an equilibrium problem: the goal is to maximize resource profits subject to the physical constraints of the resources, the market clearing process, and the strategic incentives of all other market participants. The WEASLE platform's bidding module allows multiple distinct users to submit separate bidding algorithms for each resource. This enables the platform to operate a competition with human participants who submit bidding algorithms that respond to realistic market incentives.

As mentioned above, participants in such a competition are tasked with solving an equilibrium problem. Often, this problem will be too complex to be solved exactly, and we anticipate that a participant's bidding algorithm must make some tradeoffs regarding which physical constraints, strategic incentives, or other aspects to model. Uncertainty may also play a key role since not all market inputs (e.g., load and weather forecasts) can be known with certainty. However, by rewarding the bidding algorithms based on their profits in the market simulation, our hope is that the competition provides a testbed for generating resource bids that reflect similar incentives as faced by real-world market participants.

To this end, we developed the Wholesale Electricity Analysis via Simulation and Learning Experiments (WEASLE). The WEASLE platform is built around a market clearing engine designed to easily accommodate a wide range of market structures. WEASLE also uses a physical battery dispatch and degradation model to effectively model storage performance outside of the market. This platform is also designed to call external algorithms and accept offers produced. This enables the platform to operate a competition with realistic human bidding strategies. WEASLE is the engine behind the Energy Storage Participation Algorithm Competition (ESPA-Comp).

In this paper we describe the WEASLE architecture and summarize results from the ESPA-Comp Pilot. Section 2.0 describes the WEASLE SCUC formulation, our high-renewable grid topology, and our battery model. Section 3.0 describes the website and platform designed to run the ESPA-Comp and how our pilot competition was run. Section 4.0 presents the results for the pilot competition, and Section 5.0 concludes with discussion of lessons learned and steps for future competitions.

2.0 Models

Complete formulations for all models are given in Eldridge *et al.* (2024). The general market clearing optimization model is a security constrained unit commitment (SCUC) model. Battery dispatch and degradation are formulated as two separate models: one dispatch model that is formulated as a nonlinear program (NLP), and a separate degradation cost calculation.

2.1 Market Clearing Engine

The WEASLE market clearing engine is a formulated as a surplus maximization problem. It includes system constraints on bus voltage angles, line flow, and power balance. Network power flows are modeled using DC power flow and do not include losses. The market clearing engine is written in GAMS (GAMS 2022) and solved using CPLEX.

The inputs of the market clearing engine include data on grid characteristics and a general discrete time horizon as well as bids and offers provided by demand and supply resources along with technical resource capabilities. The outputs include schedules of commitment status and dispatch of energy and ancillary products for each resource as well as market clearing prices of energy and ancillary products. The output variable values are chosen so as to maximize the total market surplus subject to constraints on individual resource operation and on the system as a whole.

Five different resource types are modeled with distinct sets of constraints: conventional generation, renewable generation, storage units, flexible demand, and virtual offers. Conventional generators and storage units have associated binary unit commitment variables. All other constraints are linear, resulting in a MILP formulation.

Resource operation constraints include maximum and minimum real power output levels, ramping limits, must-run and planned outage, minimum uptime and downtime, and others specific to certain types of resources, such as state of charge management for storage. System-wide constraints include supply and demand balance for energy and ancillary products and security constraints preventing the power flows along lines from exceeding their limits. The market surplus maximization objective consists of the value to consumers derived from consuming energy, minus the cost to producers incurred by producing energy, minus penalties on violations of certain constraints that are treated as soft constraints. The value and cost of producing energy is modeled as typical economic supply and demand functions, with diminishing returns to scale. In addition to these convex cost and value features, the objective includes nonconvex startup, shutdown, and fixed operating (no load) costs. After the model is solved a first time to determine resource dispatch schedules, the discrete variables (which generally consist of the resource commitment variables) are fixed to their optimal values, and then the model is solved again. This second solve is a convex optimization problem and therefore produces Lagrange multipliers that are used as market clearing prices for energy and ancillary products.

The market clearing engine is deliberately agnostic to the specific market structure, allowing it to solve the market under a wide range of structures. Unit commitment variables are fixed during the real-time market. Four ancillary services, regulation up, regulation down, spinning reserve (up), and non-spinning reserve (up) are added as reliability constraints. The capacity that a resource provides for each service is cumulative depending on the up or down direction of the reserve. That is, a resource that provides some of its unused capacity for regulation up service cannot also use that capacity to provide spinning or non-spinning reserve.

2.2 Market Design Specification

One of the features of the WEASLE platform is a flexible market design specification. Each market type can be given multiple specifications for different market clearing models that might run at different frequencies, at different times of the day, or with different settlement rules. An example of different market types within a single market specification would be a day-ahead and real-time market.

Each interval within each market model must be specified as one of the following settlement types:

- **Physical Delivery Interval (PHYS):** a time interval that represents physical delivery of all products cleared by the market optimization.
- **Forward Interval (FWD):** a time interval in the market clearing optimization model that does not require physical delivery but is cleared and results in a financially binding schedule.
- **Advisory Interval (ADVS):** a time interval in the market clearing optimization model that does not require physical delivery and does not result in any financially binding schedules. These intervals are typically included to avoid "end-of-horizon" effects in the market clearing solution.

To specify each market type, a user provides four input parameters.

- **Starting Period**: The first time interval included in the market model.
- **Market Clearing Period**: The clock when the market model is solved, which may be in advance of the starting period.
- **Interval Durations**: A list of the number of intervals and the duration of each interval, which may be different for each interval.
- **Interval Types**: A list of the type of each interval, which may be either physical, forward, or advisory. Only physical and forward periods are associated with a settlement.

We developed three distinct market structures for the platform.

- **Two-Settlement (TS):** This is the conventional market design. A day-ahead market (DAM) clears once a day at 9:00am, covering 36 hours starting at midnight on the upcoming day. The first 24 hours contain forward settlements while the last 12 are advisory. A real-time market (RTM) clears every five minutes of each day for a total of 288 RTM per day. The RTM clears 3 hours ahead on a five minute interval. The first interval is a physical interval while the remaining 35 are advisory.
- **Multi-Settlement (MS):** This includes a DAM identical to the TS market. The real-time market (RTM) clears every five minutes of with a 3-hour lookahead time. The first interval is a physical interval, then the MS market adds an additional 23 forward intervals to the real time market. The last 12 intervals are advisory.

- **Rolling Horizon Forward (RHF):** This market features 3 distinct clearing intervals, all with one physical interval and the remainder forward. The first is a 36 hour lookahead cleared at the top of every hour. This has a non-uniform time interval with 24 5-minute intervals, 40 15-minutes intervals, and 24 60-minute intervals. The second is a 12 hour lookahead cleared every intermediate 15 minutes with 24 5-minute intervals and the remainder 15. The last is a 2 hour lookahead cleared every intermediate 5 minutes with 5-minute intervals only.

The WEASLE platform can readily simulate these three market structures. The interface allows the easy addition of further market structures.

2.3 Battery Dispatch and Degradation

The physical battery model is based on Rosewater *et al.* (2019). The characteristics for each battery system are shown in Table 2.1. Following Rosewater, we limit the state of charge to vary between 20% and 95% of capacity, restricting to a range in which the model is valid. We selected scales such that the battery has roughly four hours of discharge capacity when discharging from 95% to 20%. Internal resistive losses are calculated by a zero-order circuit model. The model's open circuit voltage model is calculated by a cubic best fit model based on the state-of-charge (SoC).

Table 2.1 : Battery Dispatch Model Parameters

Parameter	Value
Charge Capacity	3413.33
Coulombic Efficiency	0.946
Inverter Efficiency Coefficient, 0	0
Inverter Efficiency Coefficient, 1	0.99531
Inverter Efficiency Coefficient, 2	-0.00027348
Battery Internal Resistance	365.333 μΩ
Maximum Discharge Power	125 MW
Maximum Charge Power	125 MW
SoC Capacity	640 MWh
Maximum SoC	95%
Initial SoC	60%
Minimum SoC	20%
Minimum Battery Voltage	680 V
Maximum Battery Voltage	820 V
Maximum Current Discharge	-1000 A
Maximum Current Charge	1000 A
Voltage Cubic Polynomial Fit, 0	669.282
Voltage Cubic Polynomial Fit, 1	201.004
Voltage Cubic Polynomial Fit, 2	-368.742
Voltage Cubic Polynomial Fit, 3	320.377
Cell Count	250

We include a temperature model with resistive heating and passive cooling of the battery coupled to the environment. We impose a maximum temperature of 45 Celsius, above which the battery will not be able to dispatch.

Battery degradation costs are based on a state of health calculation from Rosewater *et al.* (2019). The calculation considers four components to degradation: battery lifetime, SoC history, depth-of-discharge (DoD) history, and temperature. DoD is computed using a rainflow counting algorithm (Downing, 1982) with the SoC history. Degradation increases with time, average temperature, average SoC, and average depth of discharge. Costs are then computed by determining an end-of-life cost, which can be interpreted as the battery replacement cost due to utilization of the asset. Storage units are scaled such that under typical operations battery lifetime is approximately 15 years.

2.4 Grid Topology

The grid's electrical topology is based on an aggregate model of the Western Electricity Coordinating Council (WECC) 2030 planned system . All generation and demand within each balancing authority are aggregated, and lines connecting balancing authorities are combined. This gives the system shown in Figure 2 (Sourced from Hart and Mileva 2022), which includes a total of 40 buses and 118 transmission lines.

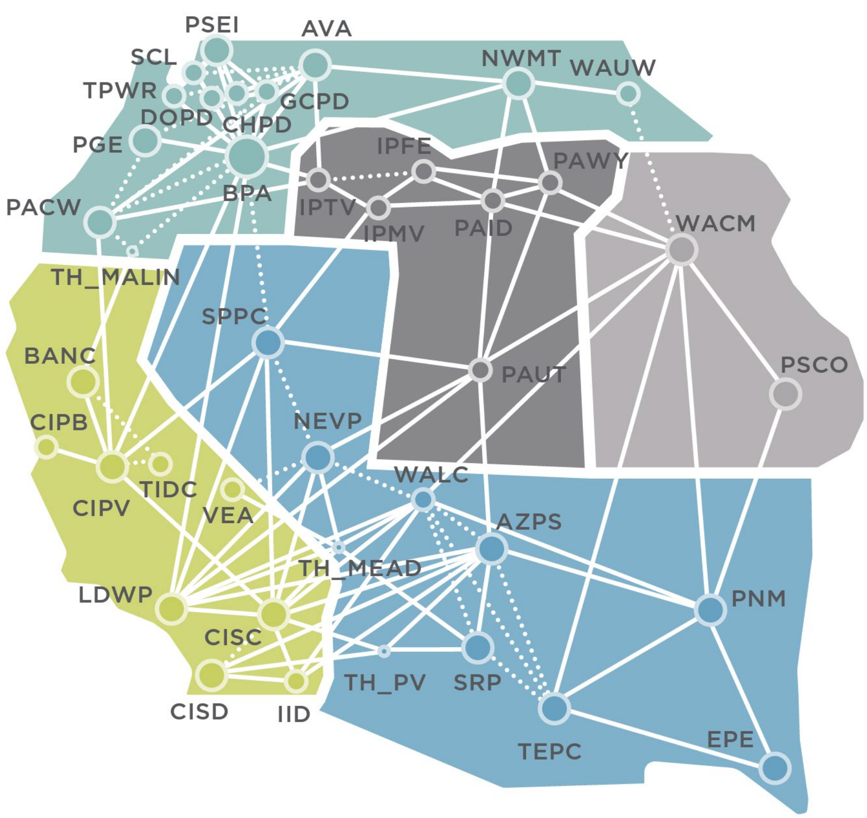

Figure 2 : Western Interconnection Case Study Zonal Topology

Generators within each balancing authority are aggregated by type including coal, natural gas, solar, wind, hydroelectric, and nuclear to provide a single representative resource of each type of resource in each zone. We then scale all of the generation to a high-renewable mix scenario detailed in Table.

We set an overall system generation capacity of 300GW with a peak demand of 190GW. For wind resources we adopt a capacity factor of 0.25, and for solar resources we adopt a capacity factor of 0.15. For these renewables, the contribution to system capacity is computed as the maximum capacity time the capacity factor. The system has twelve identical storage resources. These are distributed across six different buses with 1-3 storage units at each bus.

Table 2.2 : Benchmark Resource Capacity for Scenario Design

Resource	**NREL LA100 2045 SB100 Stress Scenario**	**CEC 2045 SB100 Core Scenario**	**CA / WECC in ADS 2030**	**ESPA-Comp Scenarios[d]**
Source	Cochran *et al.* (2021)	CEC (2021)	WECC (2024)	Tarufelli *et al.* (2024)
	Percentage of Total Capacity			
Battery Storage	7%	18%	2% / 1%	15%
Wind	26%	14%	6% / 12%	15%
Solar	31%	46%	17% / 13%	25%
NG-Steam/Combined Cycle	9%	9%	32% / 31%	20%
Coal	0%	0%	0% / 5%	2%
Demand Response	1%	1%	5% / 3%	3%
Other Renewables	26%	11%	38% / 35%	20%

Coal and natural gas offer curves are based on average monthly prices within each balancing authority. Hydroelectric generation availability is based on historic data. Wind and solar resource availability is based on hourly profiles by balancing authority from the EIA, and five-minute profiles from publicly available historic data for Bonneville Power Administration (BPA), and California ISO (CAISO). Correlated noise profiles were created to generate renewable energy forecasts.

3.0 Competition Platform

ESPA-Comp participants are able to access the WEASLE platform as external users through the ESPA-Comp website, (espa-competition.pnnl.gov. The website allows participants to access competition resources, submit their offer algorithms in a sandbox test environment, and make their final algorithm submissions for competition evaluation. To run market simulations, algorithm submission are handled on the back-end via a virtual machine (VM) and high performance cluster (HPC). The VM runs the market clearing engine while the HPC runs participant offer algorithms.

3.1 Submission Workflow

The overall platform workflow is depicted below in Figure 3. This diagram includes two different modes of operation: sandbox or competition. Sandbox runs are intended to aid in the testing and development of participant algorithms. Competition runs are initiated by the ESPA-Comp admins are used to perform the final market simulation and evaluation of results.

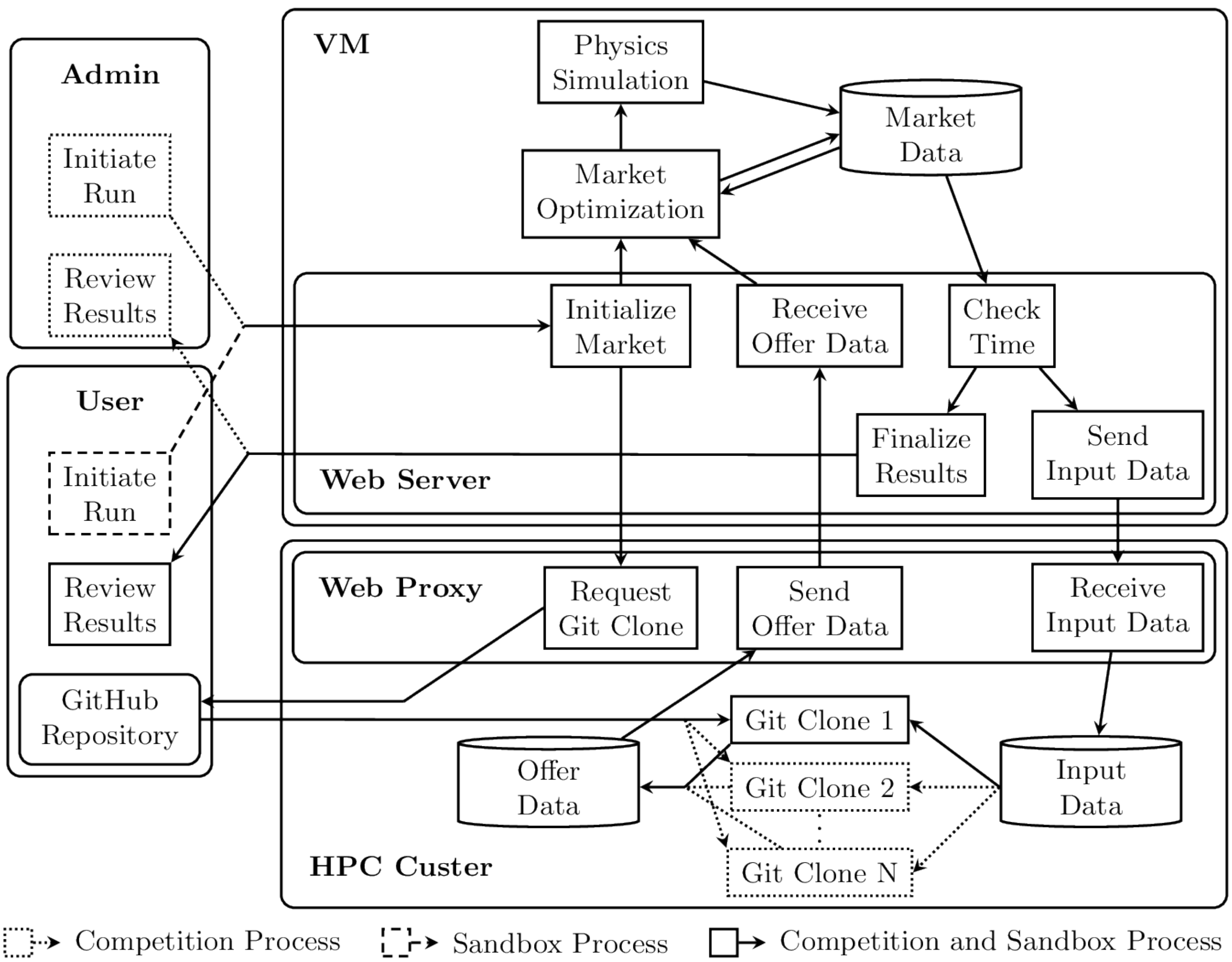

Figure 3: WEASLE Platform Backend Architecture

In both cases, the market clearing engine aggregates data at each interval to send to participant algorithms. This data includes renewable and demand forecasts, previous market clearing results, historic renewable, demand, and prices, and resource-specific dispatch and settlement information. The participant algorithm must generate an offer, formatted such that the market clearing engine can accept it. Once the market clearing engine has received an offer, it clears the market and sends information back to participant algorithms. This continues until the end of the specified simulation time horizon.

3.2 Sandbox

Before entering the Sandbox, participants first must save their code to a GitHub repository linked through ssh keys to the WEASLE platform. Once an algorithm is ready for testing, participants use the ESPA-Comp website to initialize a submission. This submission first clones the participant code onto the HPC. Then it signals the VM to begin a market clearing simulation. During the submission, participants may specify the duration of the simulation, which can range from 5 minutes to 30 days. Participants may also select a market configuration from the options detailed in the Market Formulations document (Eldridge *et al.* 2024), a starting date, and a node location for their storage resource. All other storage units bid into the market with a default offer of $0/MWh for charging and discharging.

The submission will run until completion or until an error is encountered. Upon termination of a simulation, results are made available to the participant through the ESPA-Comp website. These results include any offers generated, the latest profit summary, and any errors encountered.

3.3 Competition

ESPA-Comp participants must also submit their algorithm for competition evaluation. Competition submission are immediately cloned to the HPC but is not executed until the competition submission window closes. The competition admin selects the start date and duration of the simulation, the market structure, and the location of the storage units, which are all communicated to participants in advance.

When ready, the admin initiates a competition simulation. The competition simulation calls all participant algorithms simultaneously. The market clearing engine will wait until all algorithms have completed execution and returned an offer. In the case of empty offers, the market clearing engine will use the last submitted offer.

Like the sandbox, the competition runs until the specified end date is reached. Each participant may review results from their storage units as soon as the simulation completes. Additional data is saved for admin review, both for competition scoring and for market analysis.

4.0 Results

For the ESPA-Comp pilot we first provided a two-month long Sandbox window, during which participants could develop and test their algorithms. For the competition, we ran two market structures, the TS and the MS market. Although we were able to develop the RFH market, it was removed from competition evaluation to allow more time for teams to devote to the TS and MS market designs and to allow more simulation time for each simulation.

The month of August was selected to run the competition simulation. This month contained the summer peak conditions, when the electric system is closest to its maximum capacity, and it contained several days with LMP spikes during energy shortages. After an initial run for 25 days in August, additional time was given for participants to their review results and submit updated algorithms if desired. Both markets were then re-ran for three simulated days from August 17-19.

A baseline case was ran using default $0/MWh bids for all storage devices, and this baseline was used as a comparative benchmark for the competition simulation. In this section, we examine overall storage unit profits, the offers generated by the four different algorithms, and compare the market performance between the competition and the baseline case.

4.1 Storage Unit Profits

ESPA-Comp scores and rankings were determined by overall storage unit profits. Rankings for each category are shown in Tables 4.1 and 4.2. Rankings are based on the final 3-day simulation. Across both market designs, resources from Team PNNL-A earned an average of $11.7k per day and WayneSt resources earned $2.8k/day. PNNL-B and JHU-UCSD resources both lost money on average.

Table 4.1: Team Rankings, Two-Settlement Market

Rank	Team	Resource ID	Average Daily Profit ($1000s, 25-day simulation)	Average Daily Profit ($1000s, 3-day simulation)
1	PNNL-A	R00231	315.08	27.66
2	PNNL-A	R00234	315.80	23.44
3	WayneSt	R00230	85.92	20.65
4	PNNL-A	R00232	312.92	19.20
5	WayneSt	R00240	85.00	15.21
6	WayneSt	R00239	83.44	8.56
7	PNNL-B	R00229	-39.16	8.54
8	PNNL-B	R00237	13.36	-7.00
9	PNNL-B	R00238	43.36	-10.97
10	JHU-UCSD	R00236	-40.84	-32.08
11	JHU-UCSD	R00235	-35.48	-33.18
12	JHU-UCSD	R00233	-45.32	-33.79

Table 4.2: Team Rankings, Multi-Settlement Market

Rank	Team	Resource ID	Average Daily Profit ($1000s, 25-day simulation)	Average Daily Profit ($1000s, 3-day simulation)
1	PNNL-A	R00231	230.32	2.90
2	WayneSt	R00230	-11.4	1.98
3	PNNL-A	R00232	106.92	-0.21
4	WayneSt	R00240	-82.4	-2.20
5	PNNL-A	R00234	219.96	-2.97
6	PNNL-B	R00237	108.92	-4.58
7	PNNL-B	R00238	106.52	-9.28
8	PNNL-B	R00229	-38.08	-16.39
9	WayneSt	R00239	-100.68	-27.18
10	JHU-UCSD	R00236	24.12	-29.591
11	JHU-UCSD	R00235	23.88	-32.92
12	JHU-UCSD	R00233	26.92	-56.08

Several points are apparent from the rankings. First, algorithm scores tend to cluster by team. Team WayneSt had the largest spread, yet scores for other teams are separated by at most one rank. This indicates that the approaches used by different teams lead to significantly different outcomes. In other words, the storage bidding problem is apparently not trivial and requires careful consideration of the problem space.

Second, many storage units posted a loss over the competition. In the MS market, only two resources were profitable during the final 3-day simulation. This outcome is a result of degradation costs outstripping market revenue. As stated in the previous paragraph, this reiterates that the problem is not trivial. Furthermore, these results stand in contrast to results from traditional production cost modeling, which typically show positive profits for all resources. The inclusion of degradation costs is an important aspect of systems with significant BESS capacity.

Third, we clearly see that the TS market was more profitable for storage units than the MS market. Although this may be considered good from a BESS owner perspective, the driver of the result is shortage pricing caused by insufficient generation capacity. The MS market avoids much of the shortage due to better management of storage state-of-charge, but unfortunately, it is difficult to ensure that storage units are compensated for helping avoid the tight conditions. This is likely to be a reoccurring issue in electricity market design. That is, if high prices and high resource profits are often driven by short term shortage conditions, and the task of market design is (partially) to avoid shortage conditions, then how can resources be rewarded in the absence of system shortages?

Next, Tables 4.3 and 4.4 divide the storage unit profits from each team into components of day ahead market (DAM) settlements, real-time market (RTM) settlements, and degradation cost. The aggregate results show that degradation costs were significant in both markets. In the TS market, storage unit profits on average exceeded degradation cost, while in the MS market degradation costs significantly outweighed profits. The degradation costs and DAM profits are similar between both markets. The difference was in the RTM, which makes sense given that the RTM clearing intervals differ between TS and MS. In our competition, most storage units posted a loss in the MS RTM, in contrast to strong earnings in the TS RTM.

Table 4.3: DAM, RTM, and Degradation Profit Components, TS Market

	DAM (K$)	RTM (K$)	Degradation (K$)	Total (K$)
PNNL-A	65.65	75.93	-71.28	70.30
WayneSt	74.60	38.86	-69.05	44.41
PNNL-B	-1.89	50.54	-58.08	-9.43
JHU-UCSD	-12.04	-1.56	-85.44	-99.04

Table 4.4 DAM, RTM, and Degradation Profit Components, MS Market

	DAM (K$)	RTM (K$)	Degradation (K$)	Total (K$)
PNNL-A	64.37	13.65	-78.31	-0.28
WayneSt	73.39	-27.97	-72.81	-27.39
PNNL-B	16.82	15.71	-62.78	-30.25
JHU-UCSD	-20.80	-19.38	-78.42	-118.59

4.2 Storage Unit Offers

All competitors used the same algorithm for each resource and for both TS and MS markets. Each competitor algorithm produced distinct offer values. Offers for the TS market are illustrative and are shown in Figures 4-7. Offers are aggregated by team across all three units. The median value is shown with a solid line. A band encompassing the 16th to 84th percentile is shown in a lighter shade. For teams that submitted block offers, each block is included in the calculation, weighted by the MW quantity offered.

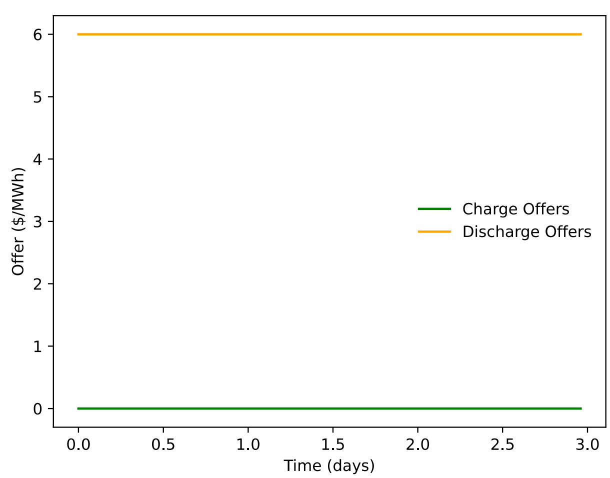

Figure 4: DAM Offers, PNNL-A, TS Market

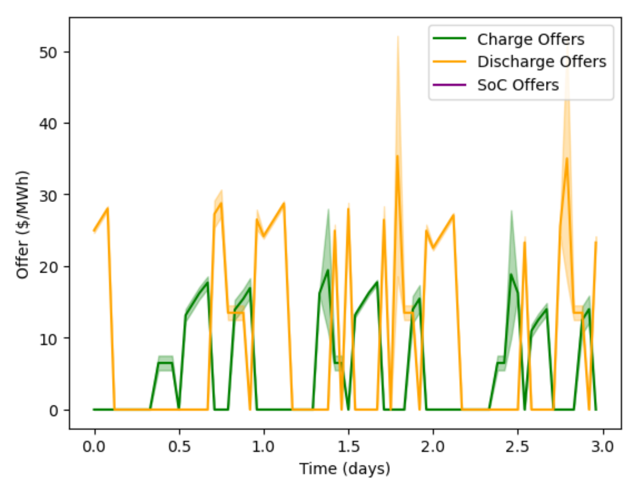

Figure 5: DAM Offers, WayneSt, TS Market

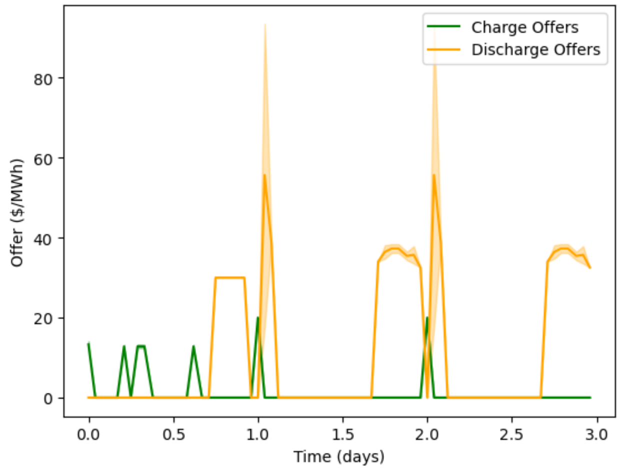

Figure 6: DAM Offers, PNNL-B, TS Market

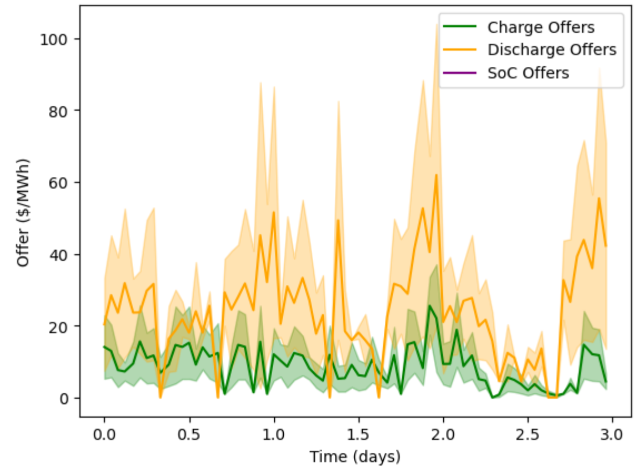

Figure 7: DAM Offers, JHU-UCSD, TS Market

Teams showed a variety of bidding strategies in the DAM. The winning team, PNNL-A, used a offer strategy with a constant $6 spread in charge and discharge prices. Team JHU-UCSD employed a dynamic algorithm with charge and discharge offers varying over time. Teams WayneSt and PNNL-B used a $0 base offer with occasional increases in offer values.

In the RTM, all participants provided SoC offers. The SoC offer is a single bid curve for energy at the last offer period in the market horizon, which influences whether the storage unit charges or discharges depending on whether the energy can be charged at less than the bid value. As in the DAM, the participant algorithms all returned a variety of different offers. RTM offers are shown in Figure REF. Team PNNL-A offered with a $0 baseline and occasional increases throughout the competition. Team JHU-UCSD offered in predominantly negative offers. These were rarely accepted by the market, limiting this team's RTM profits. Team WayneSt offered at a constant $0/MWh. Team PNNL-B offered in at two levels, both with a very high offer cost.

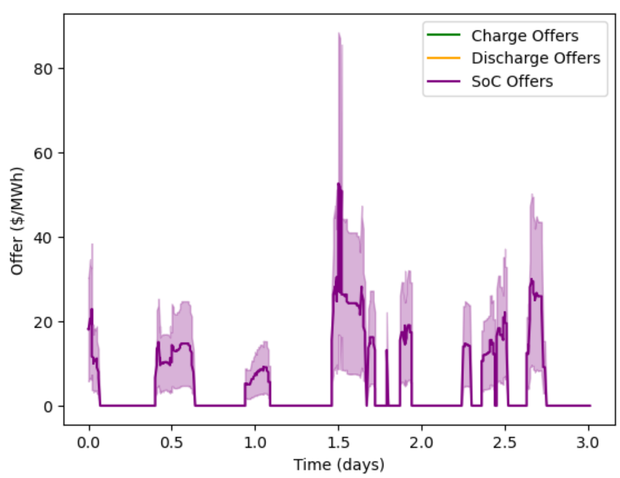

Figure 8: RTM Offers, PNNL-A, TS Market

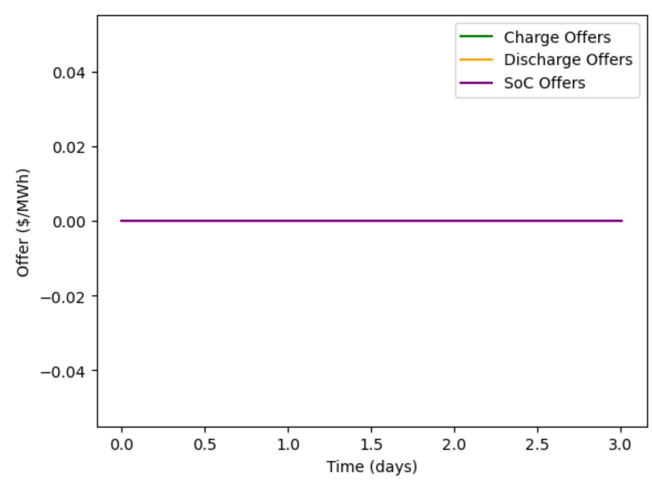

Figure 9: RTM Offers, WayneSt, TS Market

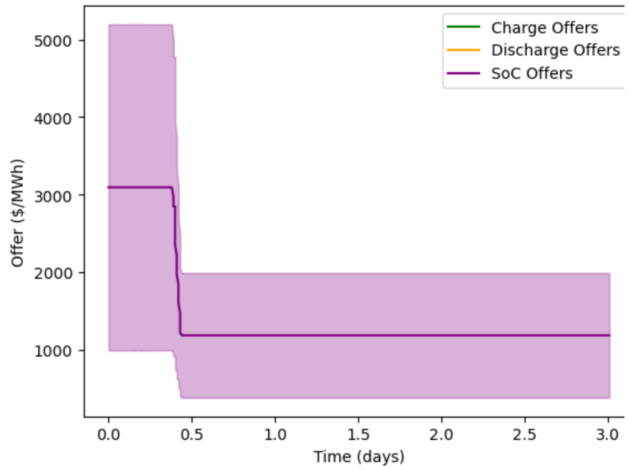

Figure 10: RTM Offers, PNNL-B, TS Market

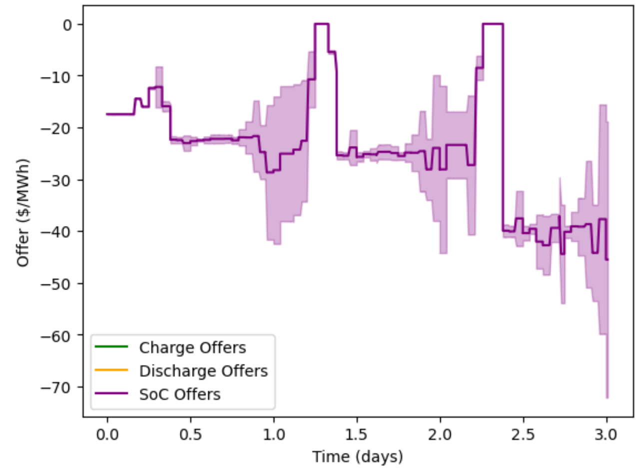

Figure 11: RTM Offers, JHU-UCSD, TS Market

4.3 Market Performance

The competition simulates a realistic, imperfect market, in comparison to the baseline simulation. When the TS and MS market designs result in changes compared to baseline, the differences can be attributed to 1) strategic incentives from the market, and/or 2) effects from the device dispatch and degradation simulation.

LMPs from the simulation are shown in Figures 12-14. For simplicity, only the RTM LMP is shown. LMPs from the baseline simulation were identical for both TS and MS markets. The baseline simulation shows a price spike around $700/MWh on the first day, and then two smaller price spikes around $200/MWh on the second and third day. Prices in the TS market also reach fairly close to $700/MWh on the first day but are somewhat higher, around $400-$500/MWh on days two and three. For the MS market, price spikes are much more restrained, around $200-$300/MWh on days one and three, and no noticeable price spike on day two.

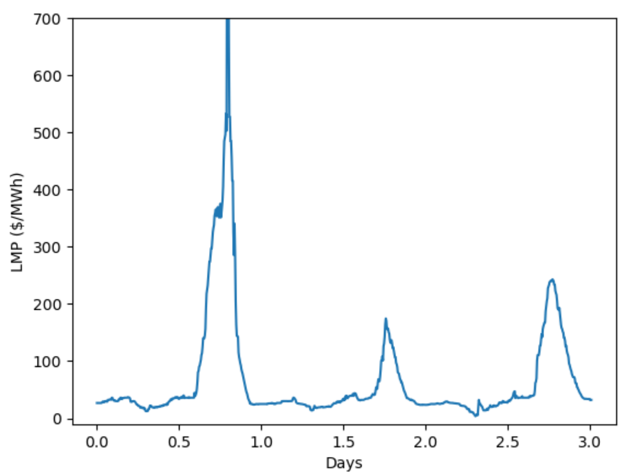

Figure 12: Real-time LMPs, TS and MS Baseline Results

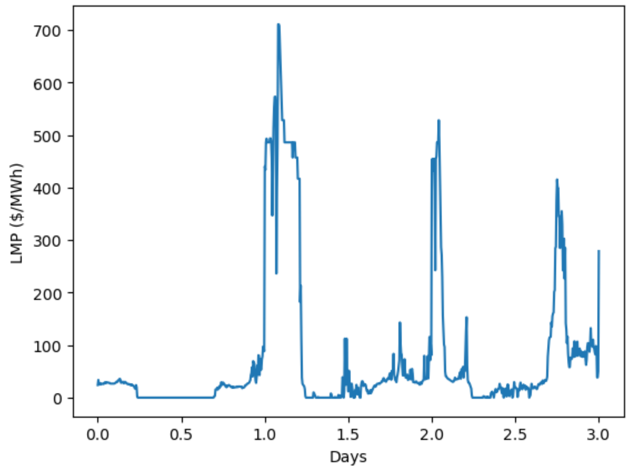

Figure 13: Real-time LMPs, TS Competition Results

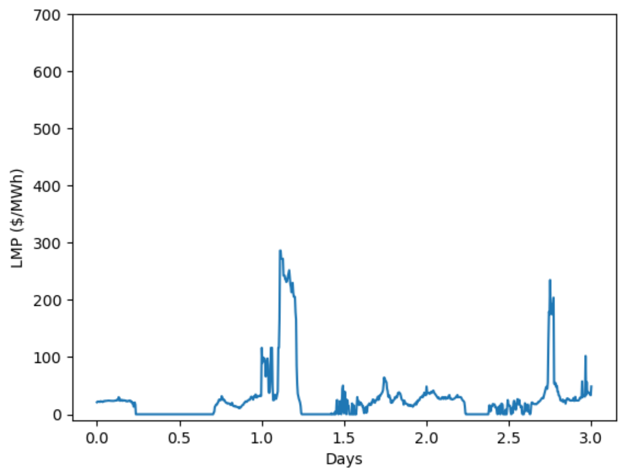

Figure 14: Real-time LMPs, MS Competition Results

The implication from the differences in LMP is that our 3-day window includes very tight system capacity conditions. The MS market is evidently much better at avoiding these tight conditions. However, although this likely contributes a reliability benefit, it is not clear if the MS is overall more efficient than TS. For a clearer comparison of market efficiency, we must look at changes in the simulated market surplus with appropriate adjustments for storage bidding values and degradation costs.

Table 4.5 shows the overall market surplus and degradation costs for the competition. This compares the competition to baseline for TS and MS markets and for DAM and RTM. For the RTM surplus, we used the physically dispatched values for both TS and MS. We also adjusted the RTM surplus to incorporate degradation. Storage units offer in with opportunity costs, which can be interpreted as a proxy for the degradation cost, so we removed the RTM storage unit offer surplus and added degradation costs. In all cases, we find that the competition surplus is below the baseline surplus by roughly half a percent. We found nearly identical surplus values in the TS market and the MS market. This result held for both the baseline case and the competition, suggesting that market efficiency is approximately the same for both market configurations.

Table 4.5: Market Surplus Comparison

Case	TS DAM ($M/day)	TS RTM ($M/day)	MS DAM ($M/day)	MS RTM ($M/day)
Baseline	183	189	183	189
Competition	182	188	182	188
Difference (%)	-0.54	-0.53	-0.54	-0.53

Table 4.6, below, compares load curtailment and penalty violations between the baseline and competition runs. Curtailed loads are shown as a percentage of the peak load, meaning that the baseline case curtailed load equivalent to 1.24% of the system's peak. Both TS and MS markets curtailed less load, with the MS market curtailment nearly half of the baseline case. This suggests that under these offer strategies, the system was better able to meet demand in the competition than the baseline. However, the competition results also show a small amount of transmission violation for the TS market, and very modest transmission violation in the MS market.

Table 4.6: Reliability Comparison

Case	TS Curtailed Load (%)	MS Curtailed Load (%)	TS Violation ($k)	MS Violation ($k)
Baseline	1.24	1.24	0	0
Competition	0.94	0.64	14.1	0.1
Difference (%)	-24.19	-48.39	N/A	N/A

5.0 Conclusion and Discussion

The WEASLE simulation platform and ESPA-Comp Pilot represent a significant new capability to assess electricity market design. Approaches based on traditional production cost modeling tools or game-theoretic analysis are limited by assumptions that don't fully capture real-world markets. For example production cost models must assume perfect competition and cannot fully reflect inefficiencies caused by strategic actions taken by market participants. While such strategic behavior is possible in game-theoretic models, they often make other simplifications such as perfect information, deterministic outcomes, or otherwise simplified technology constraints. Many existing models are generally constrained by selective focus on specific market failures, or are geared towards other applications, such as long term planning, where market design details may be less influential.

The WEASLE platform aims to overcome these limitations through the use of large, open competitions in the style of the ARPA-E GO Competition. Competitions involve real participants who are incentivized through monetary rewards to mimic market behavior found in the real world. This approach allows high-fidelity modeling of electricity market dynamics and the underlying physics of the resources that produce or consume electricity. Unlike agent-based simulations, our competition-based approach maintains the integrity of participant behavior, which is not biased towards specific outcomes by the modeler. Therefore, we believe that the WEASLE platform is capable of providing a more accurate market design evaluation and a consistent platform for evaluating diverse approaches to the market bidder problem.

One of the key innovations of the WEALSE platform is its ability to support a wide variety of bidding algorithm methods. Competitors are free to submit code that leverages advances in artificial intelligence, machine learning, optimization, or any other methods that can be compiled in code. Unlike traditional economic experiments that use human subjects to manually bid into simulated markets, WEASLE allows automated bid generation to drastically decrease the cost of simulating the market and ed markets, WEASLE allows automated bid generation to drastically expedite market clearing iterations and decrease the overall cost of the experiment. As well, the platform's detailed market clearing engine allows simulating much more precise, complex, and realistic market settlements than earlier economic experiments.

Key results from the ESPA-Comp pilot showed that markets like the two-settlement (TS) system resulted in larger transmission violation penalties, higher LMPs, and higher resource profits than the multi-settlement system. Although market surplus did not change significantly between the two market designs, this effect on transmission violations and system reliability provides one possible metric for market designers to support changes to market policies.

The main outcome from the ESPA-Comp pilot was the successful demonstration of the WEASLE platform's capability. The platform effectively supported participation of multiple competition teams, each submitting different offer algorithms, and it facilitated results and analysis of the competition market simulation. At this proof-of-concept stage, the WEASLE platform is now read for additional development to allow simulations of broader technologies and market policies.

In conclusion, the WEASLE platform addresses many limitations of traditional production cost model simulations by providing high-fidelity simulations of electric system resources and distributed offer generation by competitive participants. It offers a realistic, flexible, and scalable framework for evaluating electricity market design reforms, and it allows novel insights into strategic behaviors of market participants. With further development, we hope that the WEASLE platform demonstration can lead to future competitions that may help influence future market design reforms to help integrate an efficient, reliable, and resilient high renewable electric system.

6.0 References

Arbabzadeh, Maryam, Ramteen Sioshansi, Jeremiah X. Johnson, and Gregory A. Keoleian. "The role of energy storage in deep decarbonization of electricity production." Nature communications 10, no. 1 (2019): 3413.

Barbose, Galen. "US State Renewables Portfolio & Clean Electricity Standards: 2024 Status Update." (2024).

Bhatnagar, Dhruv, Aileen B. Currier, Jacquelynne Hernandez, Ookie Ma, and Brendan Kirby. Market and policy barriers to energy storage deployment. No. SAND2013-7606. Sandia National Lab.(SNL-NM), Albuquerque, NM (United States); United States., Washington, DC, 2013.

California Energy Commission (CEC). "S.B. 100 Joint Agency Report." Link: https://www.energy.ca.gov/sb100. (2021).

Chen, Yonghong, Marc Keyser, Matthew H. Tackett, and Xingwang Ma. "Incorporating short-term stored energy resource into Midwest ISO energy and ancillary service market." IEEE Transactions on Power Systems 26, no. 2 (2010): 829-838.

Cramton, Peter. "Electricity market design." Oxford Review of Economic Policy 33, no. 4 (2017): 589-612.

Downing, Stephen D., and D. F. Socie. "Simple rainflow counting algorithms." International journal of fatigue 4, no. 1 (1982): 31-40.

Ela, Erik, Michael Milligan, Aaron Bloom, Audun Botterud, Aaron Townsend, Todd Levin, and Bethany A. Frew. "Wholesale electricity market design with increasing levels of renewable generation: Incentivizing flexibility in system operations." The Electricity Journal 29, no. 4 (2016): 51-60.

Eldridge, Brent C., and Abhishek Somani. Impact of FERC Order 2222 on DER Participation Rules in US Electricity Markets. No. PNNL-33383. Pacific Northwest National Laboratory (PNNL), Richland, WA (United States), 2022a.

Eldridge, Brent C., and Abhishek Somani. Modeling and Market Design Considerations for Conventional and Decarbonized Resources. No. PNNL-33792. Pacific Northwest National Laboratory (PNNL), Richland, WA (United States), 2022b.

Eldridge, Brent, Jesse Holzer, Kostas Oikonomou, Brittany Tarufelli, Li He, and Abhishek Somani. Model Formulations: Market clearing models, market design specifications, and dispatch simulation. Pacific Northwest National Laboratory (PNNL), Richland, WA (United States), 2024.

GAMS Development Corporation. General Algebraic Modeling System (GAMS) Release 41.3.0, Fairfax, VA, USA. (2022).

Haas, Reinhard, Claudia Kemfert, Hans Auer, Amela Ajanovic, Marlene Sayer, and Albert Hiesl. "On the economics of storage for electricity: Current state and future market design prospects." Wiley Interdisciplinary Reviews: Energy and Environment 11, no. 3 (2022): e431.

Hart, Elaine, and Ana Mileva. "Advancing resource adequacy analysis with the GridPath RA Toolkit: A case study of the Western US." Energy Systems Integration Groups ESIG) Webinar Presentation 6 (2022).

Haugen, Mari, Paris L. Blaisdell-Pijuan, Audun Botterud, Todd Levin, Zhi Zhou, Michael Belsnes, Magnus Korpås, and Abhishek Somani. "Power market models for the clean energy transition: State of the art and future research needs." Applied Energy 357 (2024): 122495.

Hobbs, Benjamin F., ed. The next generation of electric power unit commitment models. Vol. 36. Springer Science & Business Media, 2001.

Holzer, Jesse T., Carleton J. Coffrin, Christopher DeMarco, Ray Duthu, Stephen T. Elbert, Brent C. Eldridge, Tarek Elgindy et al. Grid optimization competition challenge 3 problem formulation. No. PNNL-35792. Pacific Northwest National Laboratory (PNNL), Richland, WA (United States), 2024.

Litvinov, Eugene, Feng Zhao, and Tongxin Zheng. "Electricity markets in the united states: Power industry restructuring processes for the present and future." IEEE Power and Energy Magazine 17, no. 1 (2019): 32-42.

Milligan, Michael, Bethany A. Frew, Aaron Bloom, Erik Ela, Audun Botterud, Aaron Townsend, and Todd Levin. "Wholesale electricity market design with increasing levels of renewable generation: Revenue sufficiency and long-term reliability." The Electricity Journal 29, no. 2 (2016): 26-38.

Pérez-Arriaga, Ignacio J., Jesse D. Jenkins, and Carlos Batlle. "A regulatory framework for an evolving electricity sector: Highlights of the MIT utility of the future study." Economics of Energy & Environmental Policy 6, no. 1 (2017): 71-92.

Plott, Charles R., and Vernon L. Smith, eds. Handbook of experimental economics results. Vol. 1. Elsevier, 2008.

Rosewater, David M., David A. Copp, Tu A. Nguyen, Raymond H. Byrne, and Surya Santoso. "Battery energy storage models for optimal control." IEEE access 7 (2019): 178357-178391.

Salles, Mauricio BC, Junling Huang, Michael J. Aziz, and William W. Hogan. "Potential arbitrage revenue of energy storage systems in PJM." Energies 10, no. 8 (2017): 1100.

Tarufelli, Brittany, Brent Eldridge, Li He, Konstantinos Oikonomou, Jesse Holzer, Matt Cornachione, Abhishek Somani. Energy Storage Participation Algorithm Competition Overview: Market simulation, stages, storage attributes, and scoring. Pacific Northwest National Laboratory (PNNL), Richland, WA (United States), 2024.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354

1-888-375-PNNL (7665)

www.pnnl.gov