



# Machine Learning Applications to Forecast Operations of Energy Storage Resources and Crypto Loads

Pengwei Du and Jian Ma

ERCOT  
August 31, 2022

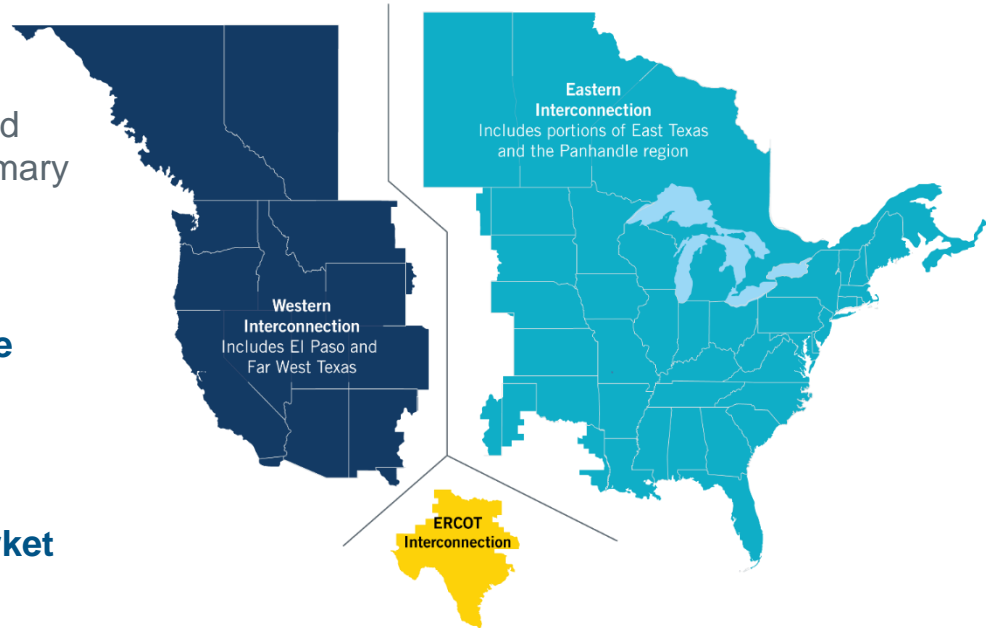
# Overview

- Grid Operations at ERCOT
- Challenges for Machine Learning Applications Implemented for Future Power Grid
- A Case Study of Machine Learning Applications at ERCOT: Forecasting Energy Storage Resources

# WHAT IS ERCOT?

The Texas Legislature restructured the Texas electric market in 1999 and assigned ERCOT four primary responsibilities:

- **System reliability**
- **Competitive wholesale market**
- **Open access to transmission**
- **Competitive retail market**



ERCOT is a nonprofit organization that is regulated by the Public Utility Commission of Texas, with oversight by the Texas Legislature.

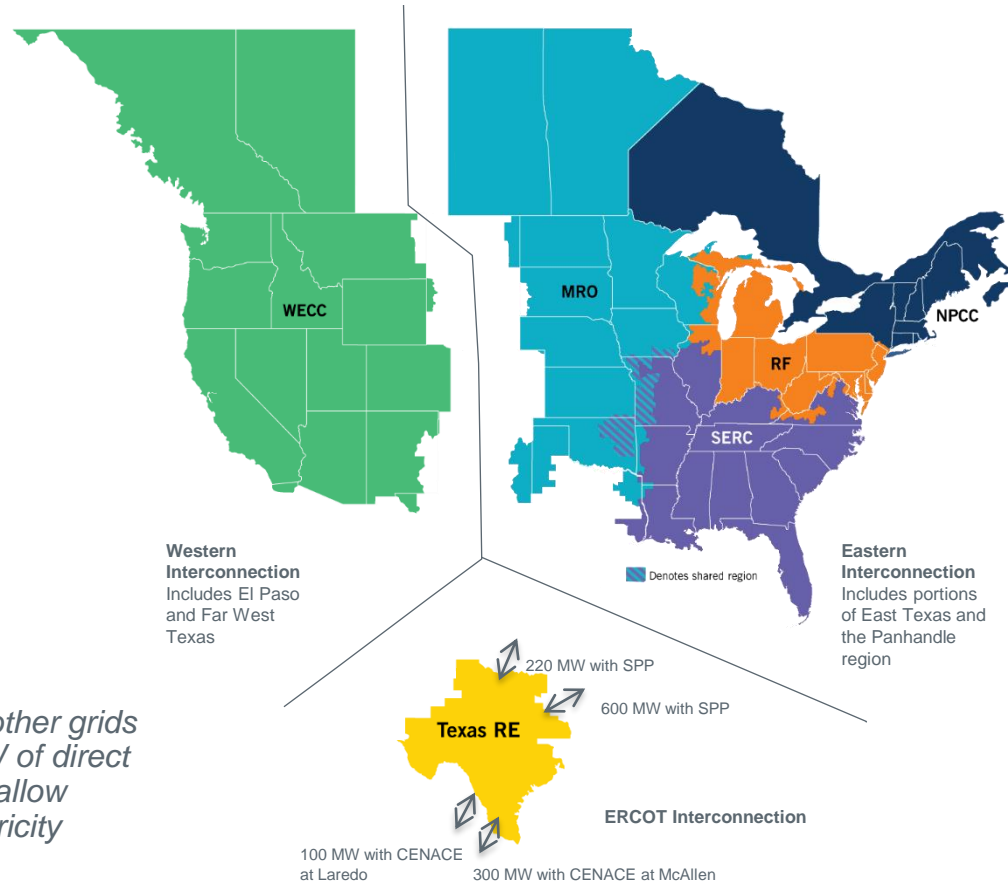
ERCOT is not a market participant and does not own generation or transmission/distribution wires.

# THE ERCOT REGION

The interconnected electrical system serving most of Texas, with limited external connections

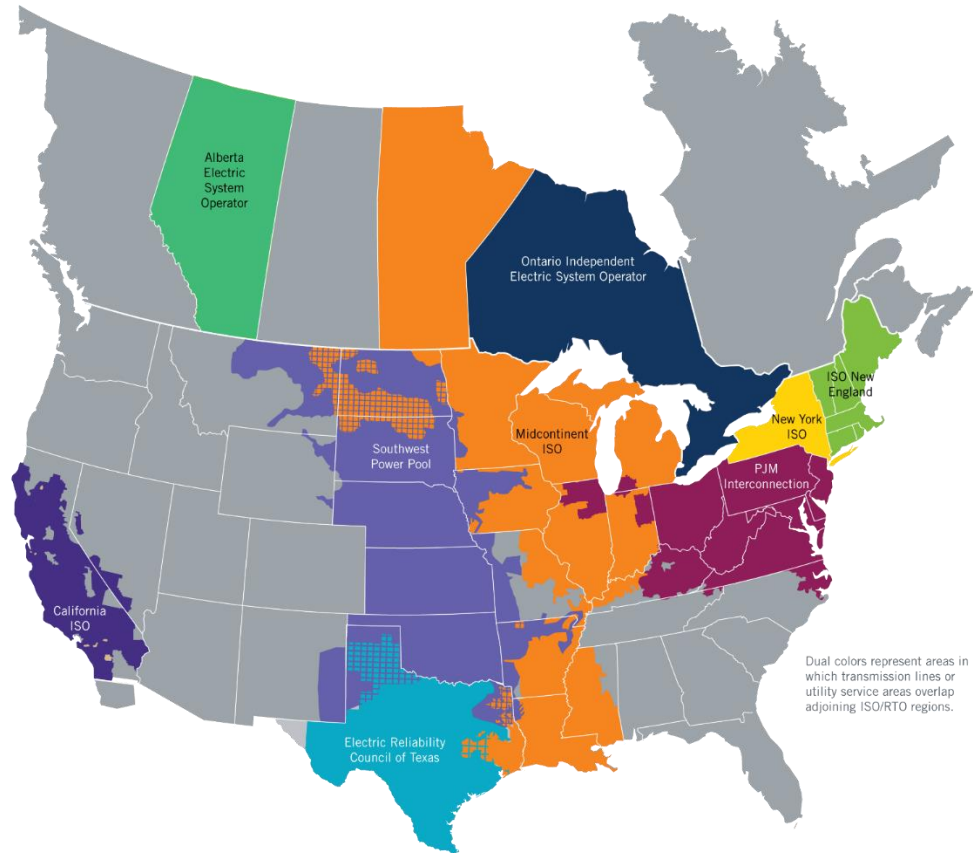
- 90% of Texas electric load; 75% of Texas land
- 74,820 MW peak, August 12, 2019
- More than 46,500 miles of transmission lines
- 710+ generation units (excluding PUNs)

*ERCOT connections to other grids are limited to ~1,220 MW of direct current (DC) ties, which allow control over flow of electricity*



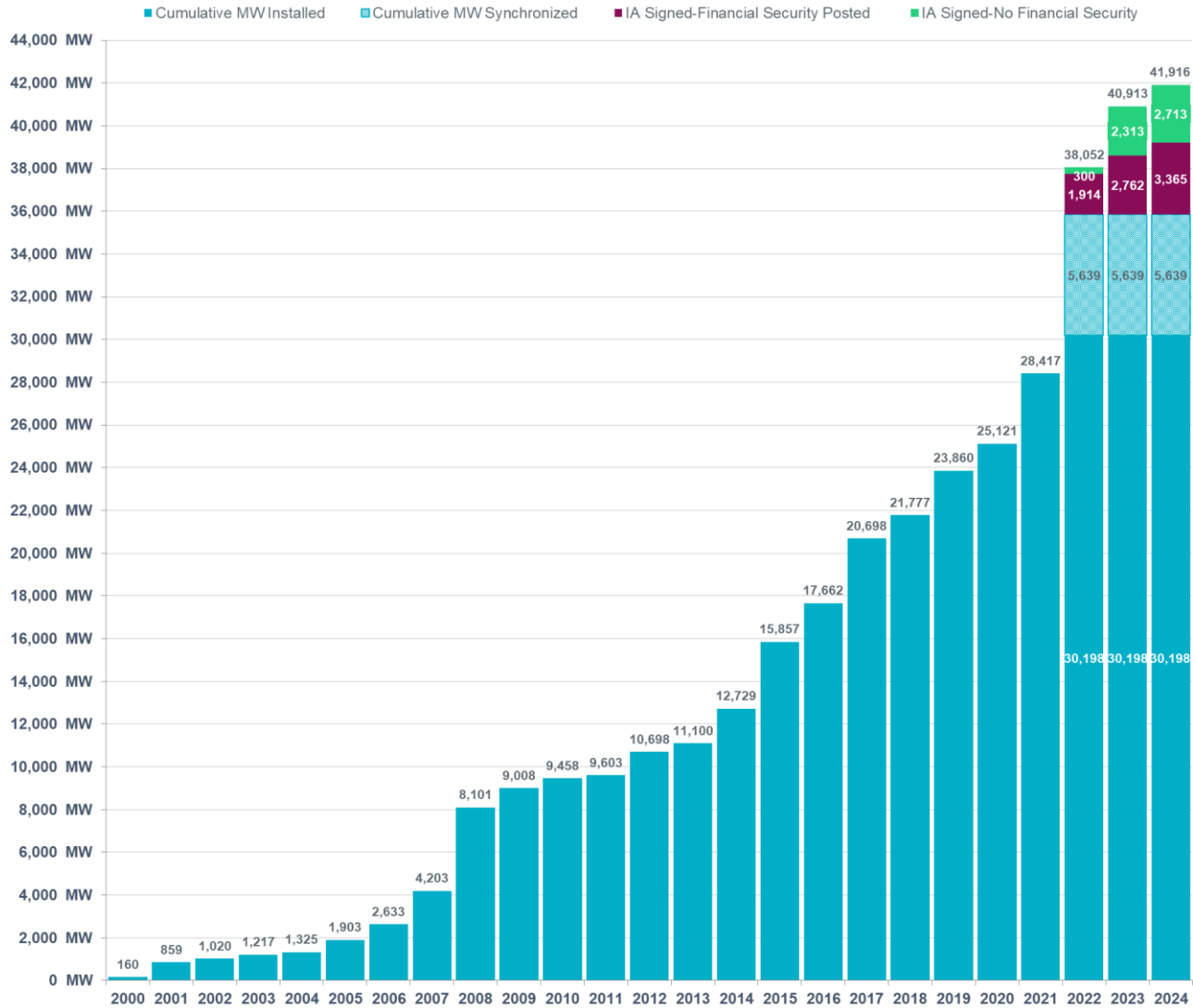
# ISOs and RTOs

- ERCOT is one of nine **independent system operators** and **regional transmission organizations** in the U.S. and Canada.
- Together, ISO/RTOs serve about two-thirds of electric consumers in the U.S. and more than half of consumers in Canada.



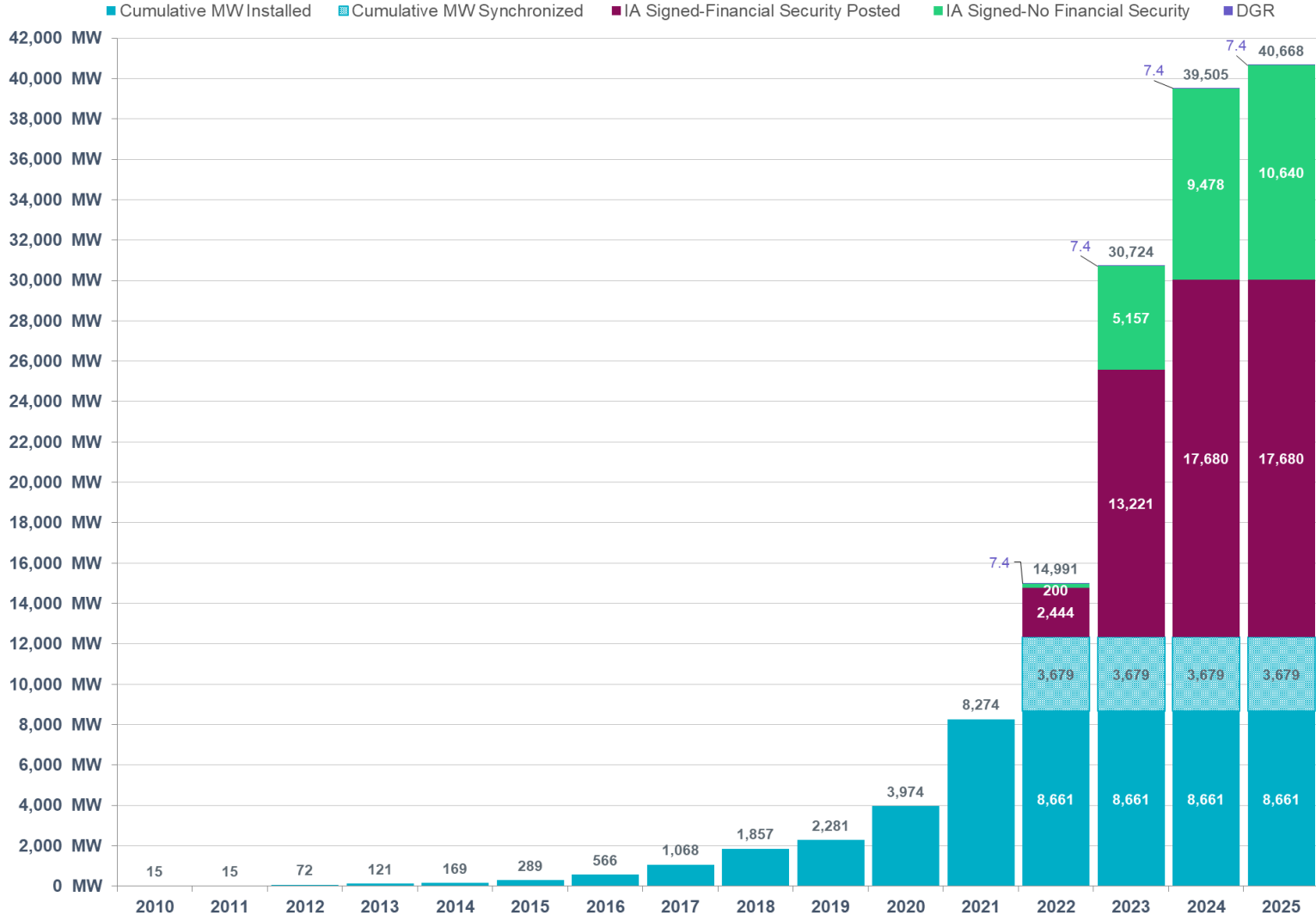
# GROWTH IN WIND INSTALLED CAPACITY

ERCOT Wind Additions by Year (as of Jul 31, 2022)



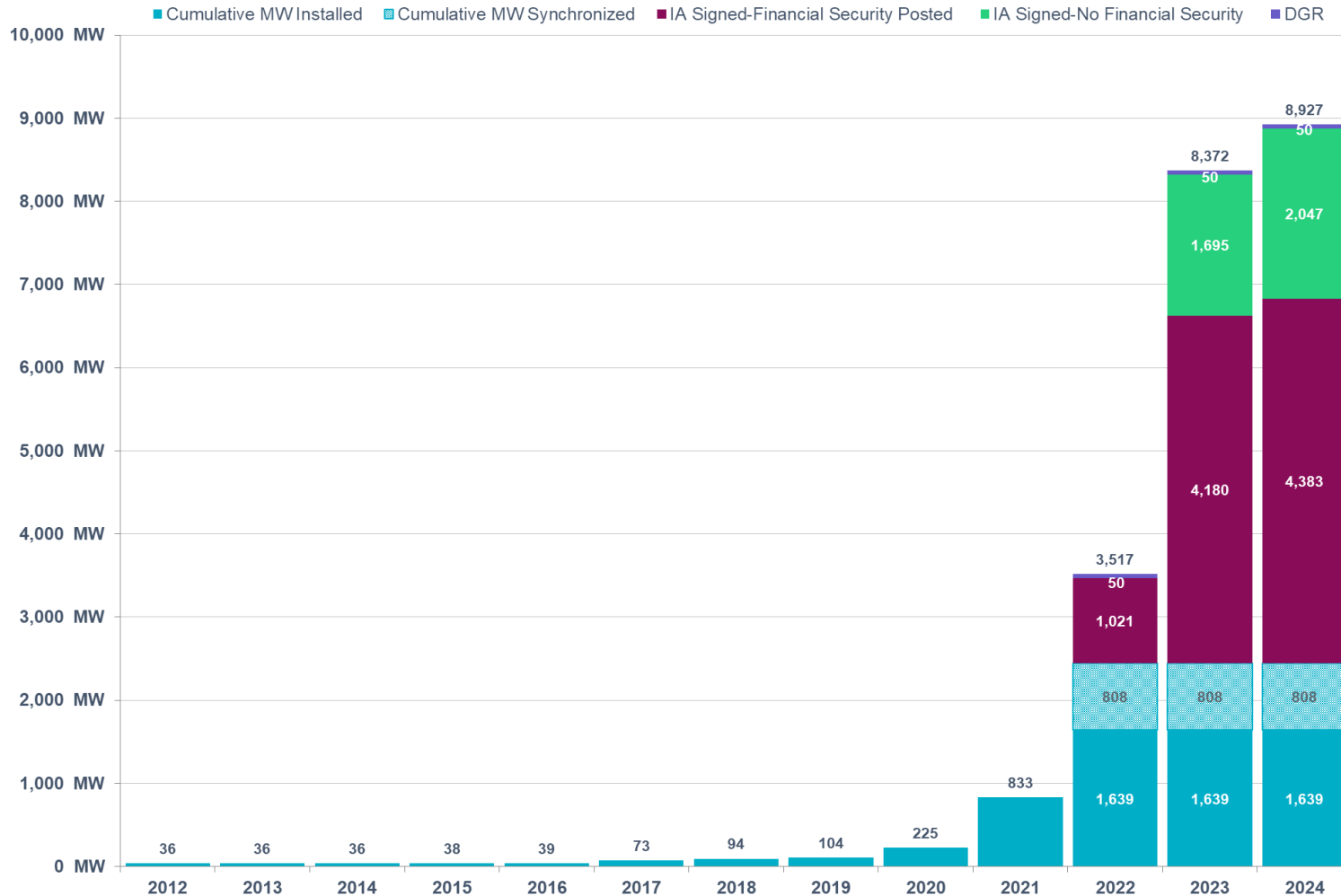
# GROWTH IN SOLAR INSTALLED CAPACITY

ERCOT Solar Additions by Year (as of Jul 31, 2022)



# BATTERY STORAGE CAPACITY

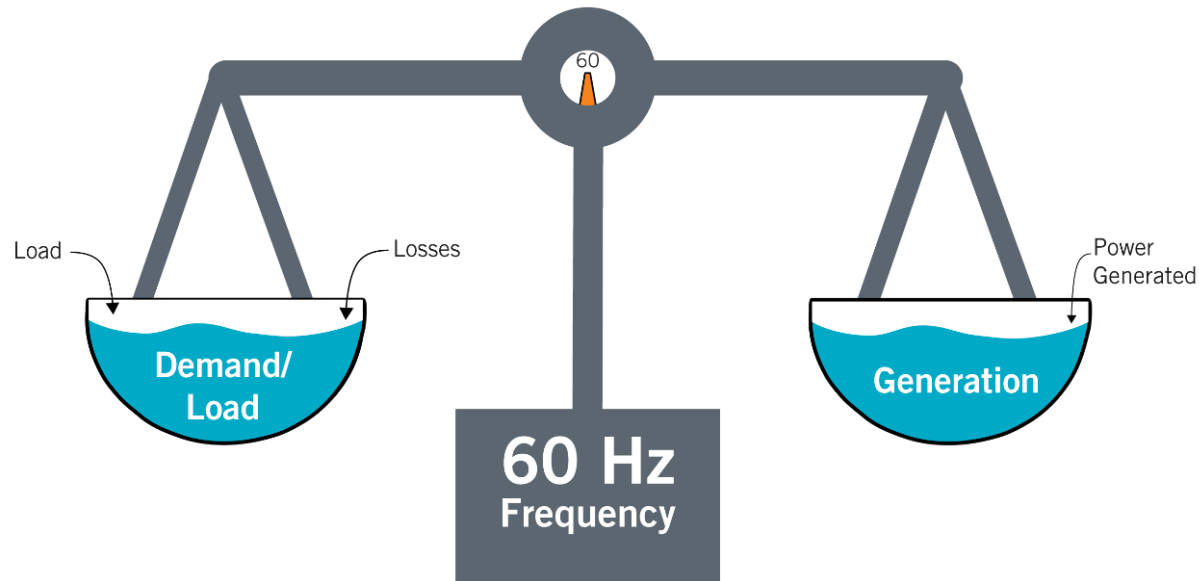
ERCOT Battery Additions by Year (as of Jul 31, 2022)





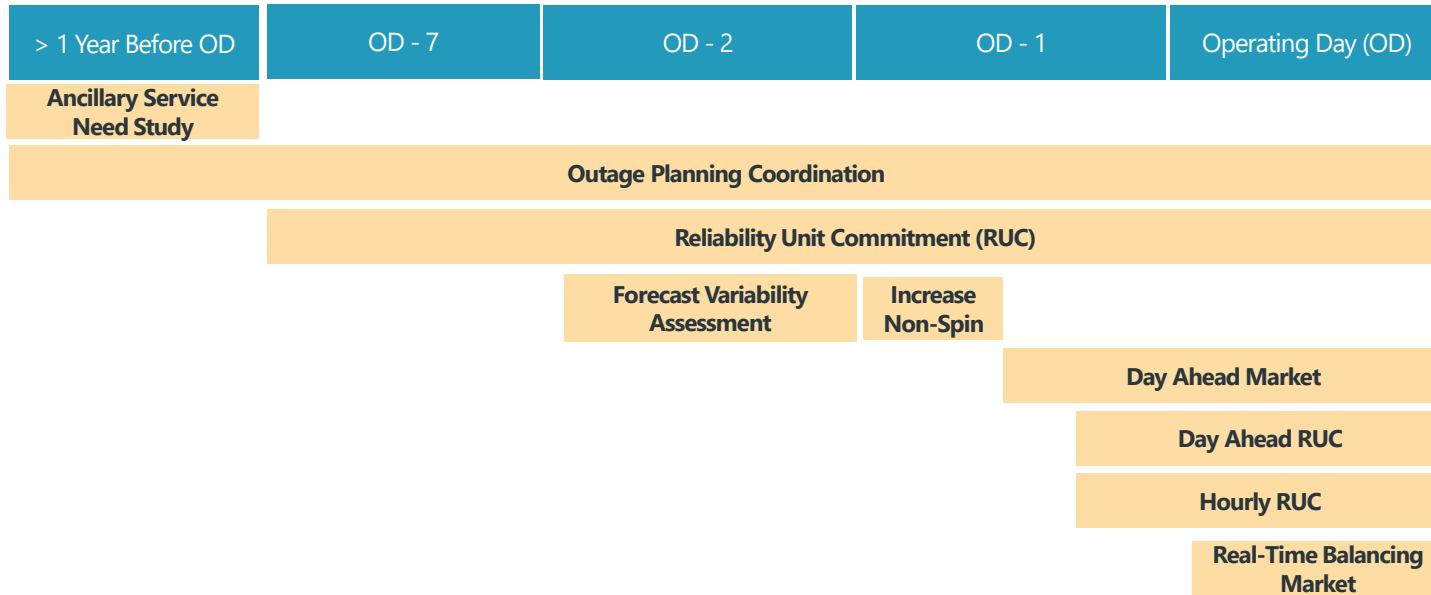
# POWER SUPPLY (GENERATION) MUST MATCH LOAD (DEMAND)

- The fundamental concept behind ERCOT operations is that generation has to match load at all times



- In other words, a 1 MW reduction in load has the same effect on the grid as a 1 MW increase in generation.

# OPERATIONAL STUDIES & ANALYSIS



# ANCILLARY SERVICE (AS)

## REGULATION

71 MW to 921 MW\*

- Generators or Controllable Load Resources (CLRs) respond within 5 seconds to ERCOT LFC instructions
- FRRS providers respond within 60 cycles of either its receipt of ERCOT instruction or at 59.91 Hz; 65 MW maximum for FRRS-Up, 35 MW maximum for FRRS-Down
- Capacity that is used to provide conventional Regulation must be capable of being sustained for 1 hour

## RESPONSIVE RESERVE SERVICE (RRS)

### Fast Frequency Response (FFR)

Load Resources on Under Frequency Relay (UFR)

### Primary Frequency Response (PFR)

2,300 to 3,534 MW\*

### FFR

- Triggered at 59.85 Hz and full response in 15 cycles
- Once deployed, sustain for up to 15 mins. Once recalled, restore within 15 mins
- Maximum 450 MW of RRS may be provided by FFR Resources

### PFR

- PFR capable capacity reserved on generators or Controllable Load Resources (CLR)
- Minimum 1,150 MW must be provided by resources capable of PFR
- Capacity that is used to provide RRS-PFR must be capable of being sustained for 1 hour

### Load Resources with under frequency relay (UFR)

- Triggered at 59.70 Hz and full response in 30 cycles
- Sustain until recalled. Once recalled, restore within 3 hours
- Beyond the minimum PFR, up to 60% of total RRS can come from Load Resources on UFR or FFR

## ERCOT CONTINGENCY RESERVE SERVICE (ECRS)

10-minute ramp

Load Resources may or may not be on UFR

1,093 MW to 3,039 MW\*\*

### Generation

- Online or offline capacity that can be converted to energy within 10 minutes
- Dispatched by SCED
- Capacity that is used to provide ECRS must be capable of being sustained for 2 hours

### Load Resources (UFR not required)

- Up to 50% of ECRS capacity can come from Load Resources with or without UFR
- Once deployed, must respond within 10 minutes. Restoration within 3 hours

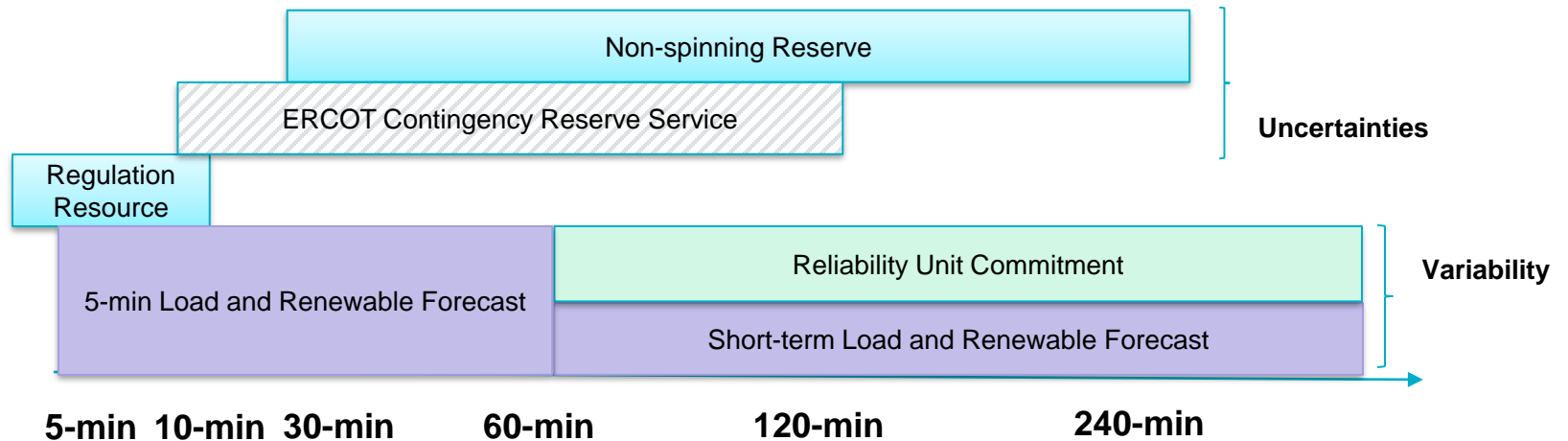
## NON-SPINNING RESERVE SERVICE

2,384 MW to 5,276 MW\*

- Online or offline capacity from generators and Load Resources that can be deployed within 30 minutes
- Online capacity is dispatched by SCED; Offline capacity is dispatched by XML instruction
- Minimum 1,430 MW must be provided by SCED dispatchable Resources
- Non-Controllable Load Resources may provide Non-Spin; UFR if available, must be disarmed in Real Time
- Capacity that is used to provide Non-Spin must be capable of being sustained for 4 hours

# MARKET AND OPERATIONS

- ERCOT ensures that there are enough resources and resource flexibility available on the system to meet net load, net load changes, and uncertainties by using the capacity available from Ancillary Services reserved and Reliability Unit Commitment.



\* ERCOT Contingency Reserve Service will be implemented in May/June, 2023.

# CHALLENGES

- The increase in uncertainties and variabilities leads to the difficulty in maintaining a balance between load and resources, and the grid reliability is of a particular concern for “tail events”.
- A looking-forward grid operation is desired to provide more flexibility to schedule and commit the dispatchable resources, which requires accurate and robust forecast for wind/solar/load and non-traditional resources like Energy Storage Resources and Crypto Loads.
- **Machine learning** is a promising solution to predict and mitigate the increasing reliability risk for a future grid.

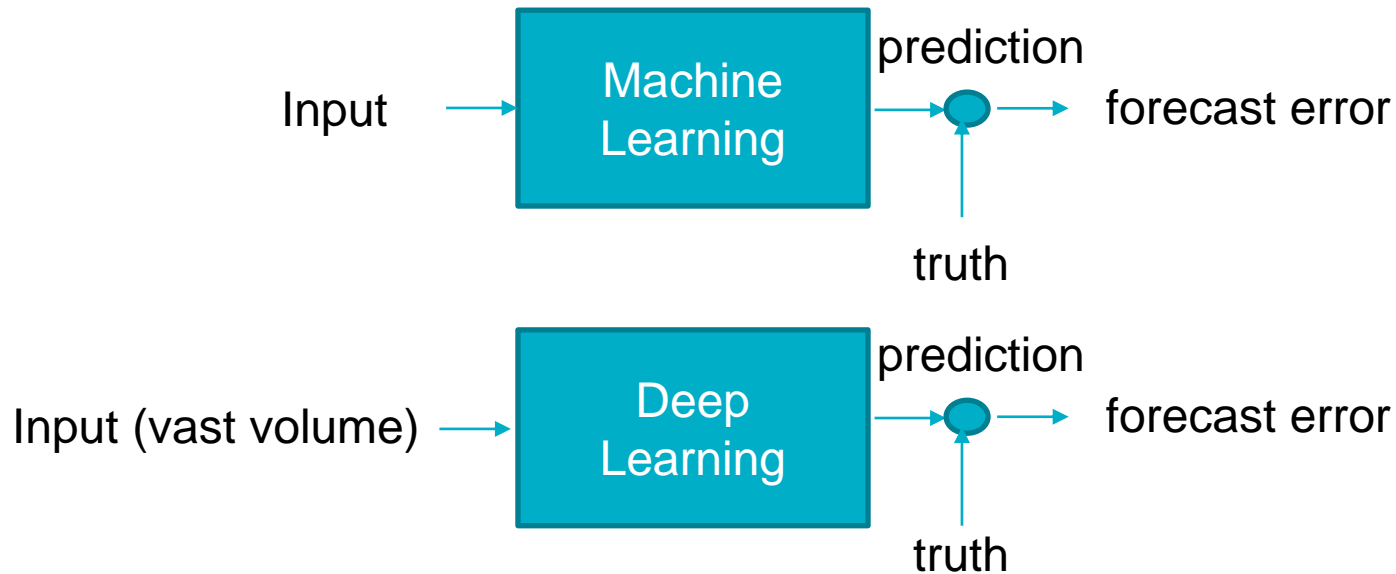
# MACHINE LEARNING

- The term “Machine Learning” (“ML”) refers to the use of sophisticated computer algorithms to automatically improve decision making through learned experience and the use of, typically large, data sets.
  - ML concepts have been discussed in academia and industry since the 1950s, but have only seen real-world application during the last few decades.
  - Recent improvements in enabling technologies –computing speeds, data storage costs, and algorithm design-- have made ML applications more relevant.
  - Today, ML techniques are applied in nearly every industry to perform classification (e.g. computer vision) and forecasting (e.g. financial metrics, wind, etc.), or to inform human decision making.

# CHALLENGES AND BARRIERS TO MACHINE LEARNING IMPLEMENTATION

Challenge	Description
Organization direction, intentionality, OKRs.	Mos of utilities do not have a comprehensive ML strategy or any corporate goals (OKRs) that are related to ML initiatives (other than those related to specific ML use cases).
Budget	In some cases, ML initiatives may fit within the existing scope of work and budget of a department. However, other applications could require a significant amount of effort.
Data access and distributed ownership	To train a successful ML model requires a large amount of representative data, relevant to the scenarios trained. Four issues need to be addressed with regards to data access: <ol style="list-style-type: none"> <li>1) Segregation of data located on different enterprise systems,</li> <li>2) Access to data by users in a different department from the department that owns the database,</li> <li>3) Data governance, and</li> <li>4) Data availability to meet granularity, latency, and data quality requirements.</li> </ol>
Creation and tuning of models	The success of ML needs a good understanding of data itself and how the process is trained. Rather than treating ML as a black-box approach, a significant amount of effort is required to create and tune the ML models, even on a continuous basis.
Limited ability to use cloud services	To properly accommodate BES Cyber Assets and Protected Cyber Assets in cloud computing, existing definitions in NERC CIP standards may need to be revised.
ML Infrastructure	ML implementation requires consideration of the location of data, network communication bandwidth and speed, processing power requirements, and security of data.
Security	Some data used by a new ML application may not have been considered sensitive previously. However, to the extent that the data is being used to make critical operational decisions, the security and redundancy associated with it may need to be revised.

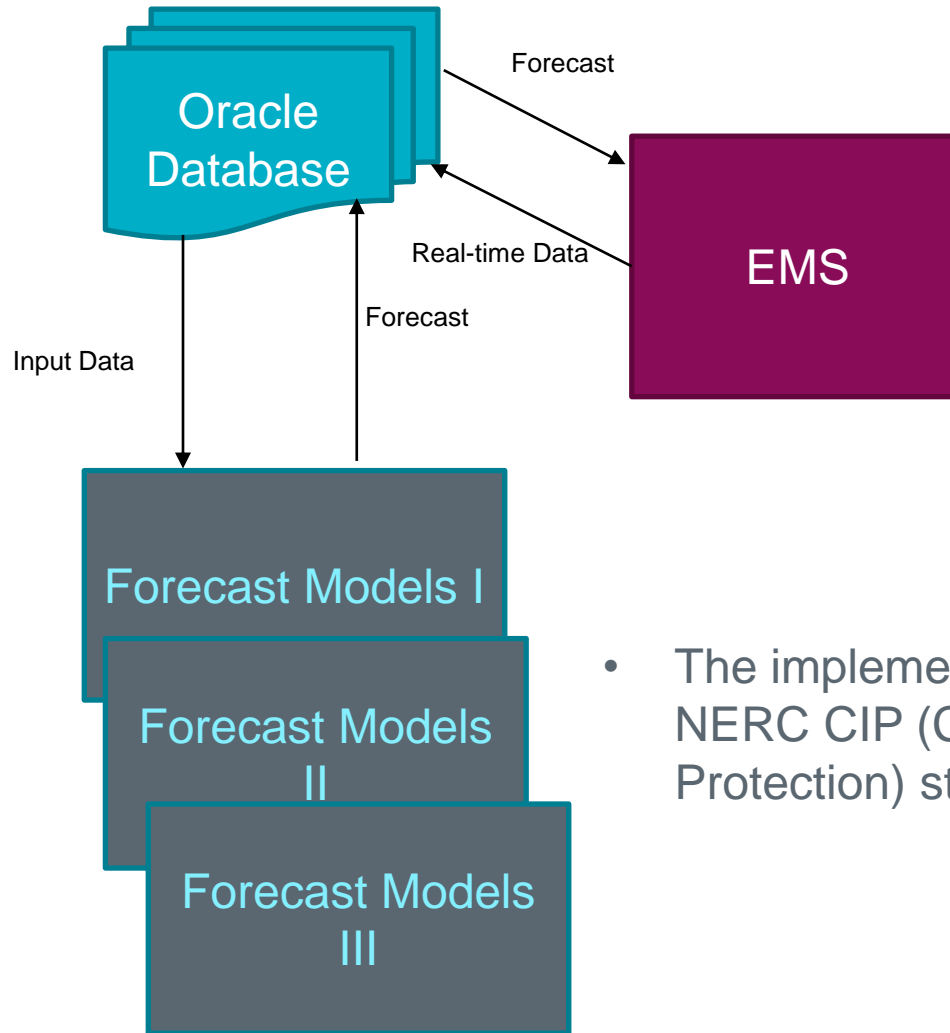
# ADVANCED MACHINE LEARNING MODELS



- Machine learning searches for a repeated pattern driven by inferred variables.
- Recent advancement in deep learning models and tools makes tuning of models less time-consuming.
- Deep learning has found many successful applications in other industries nowadays.



# IT SETUP

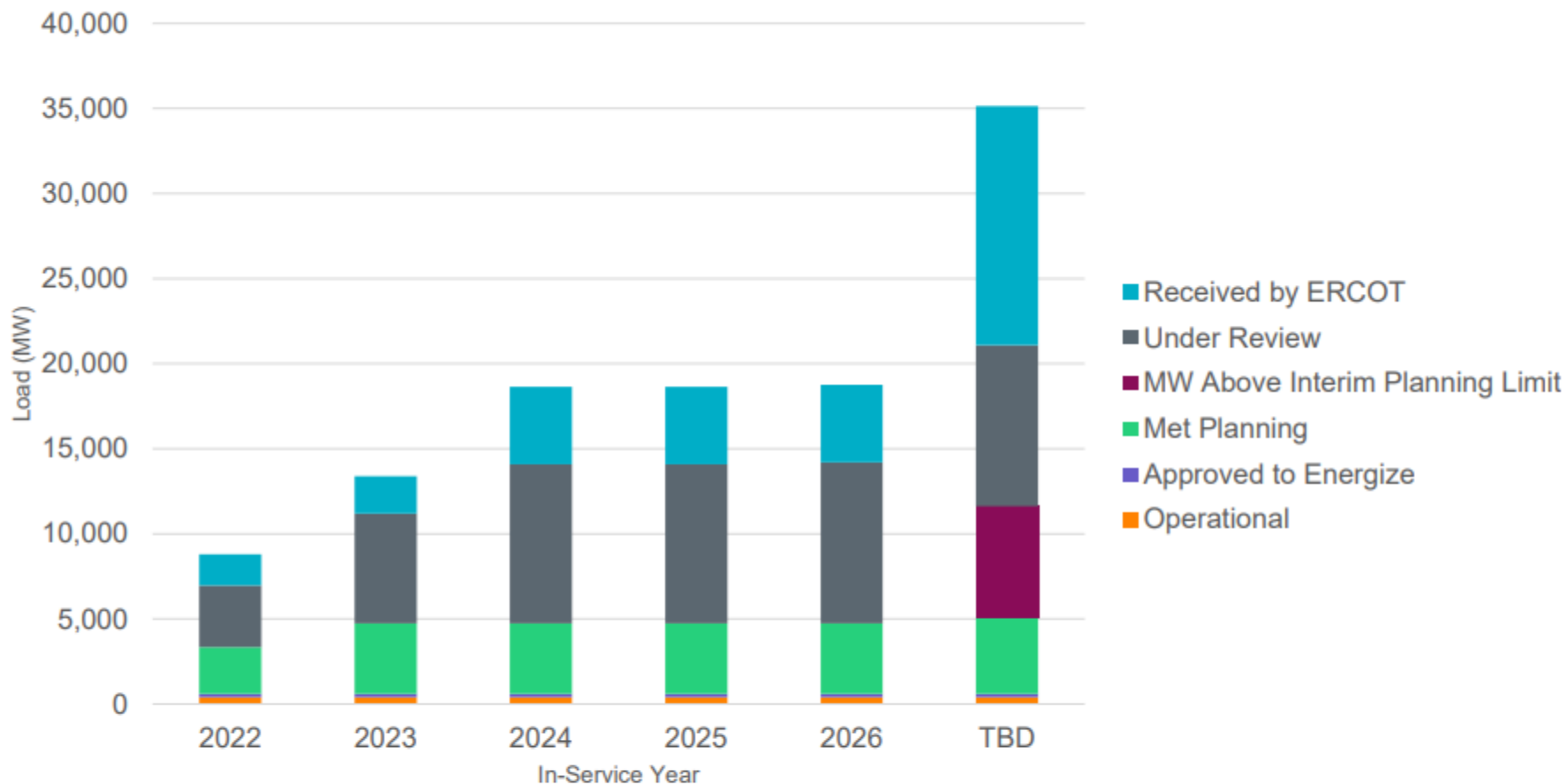


- The implementation needs to meet NERC CIP (Critical Infrastructure Protection) standards.

# LARGE FLEXIBLE LOAD (LFL)

- Historically, ERCOT has treated all loads as firm.
- Large, traditional datacenters or crypto loads can curtail their operations when the price is high a) to minimize their costs and b) to assist ERCOT by returning electricity to the grid during times of scarcity.
- A variety of LFL exist currently, which can be characterized as controllable, non-controllable, Private Use Networks (PUN), or co-located.
- Extra telemetry data point(s) will need to be set up to track some of LFL.
- Large Flexible Load Task Force (LFLTF) was created to work through grid code and market issues related to the integration of Large Flexible Loads (<https://www.ercot.com/committees/tac/lfltf>).

# Current LLI Interconnection Queue



- Chart is cumulative, with 35,142 MW currently being tracked in the queue
- The TBD column contains planned projects with no defined target in-service date, as well as remaining MWs after an interim limit has been applied to a proposed project

# LOAD TYPES & CHARACTERISTICS MATRIX

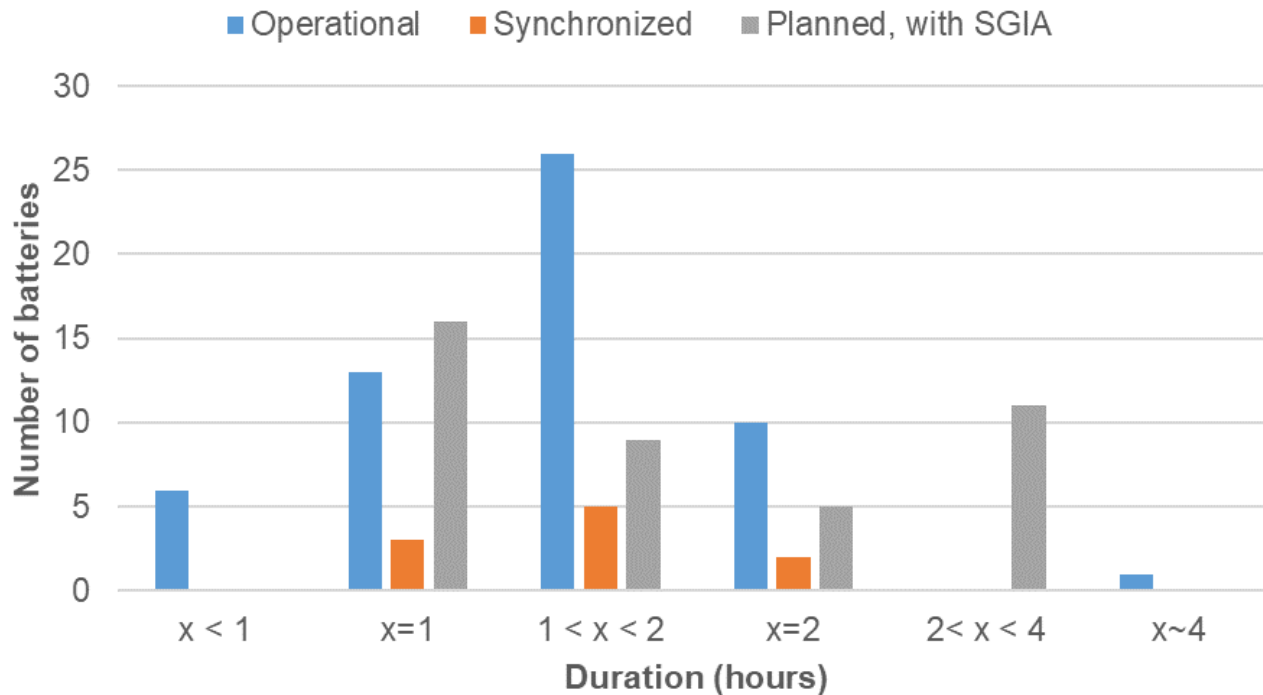
	Load Resources (Providing Energy and/or AS via DR; Registered)		Emergency Response Service (ERS) Load	Large Flexible Load ≥75 MW or => 20 MW (if co-located)		Transmission Firm Load
	Controllable Load Resource (CLR)	NCLR		SCED Flexible Load	Interruptible Load	
Characteristic	Loads meeting the size limitations and capable of controllably reducing or increasing consumption via SCED dispatch with new Base Points delivered every SCED Run (talk of SCED being process as 1 minute execution)	<p>If providing RRS must be Controlled by high-set UFRs @ 59.70 Hz and must be able to interrupt within 30 cycles.</p> <p>May be deployed with 10 or 30 minute notice.</p> <p>Must return to service withing 3 hours following ERCOT deployment,</p>	<p>Any load capable of interruption prior to EEA on instruction from ERCOT</p> <p>A load, or aggregation of loads, contracted to provide ERS</p> <p>May be deployed with 10 or 30 minute notice.</p> <p>Must return to service withing 3 hours following ERCOT deployment, Must maintain 95% Availability factor.</p>	<p>The portion of a Large Load that has bids and offers in SCED</p> <p>Nodal vs Zonal Shift Factor and Pricing to be determined</p>	<p>Large Load capable of interruption prior to EEA on instruction from ERCOT</p> <p>Blocky Response time = x minutes or less</p> <p>Duration time &gt; y minutes</p> <p>Customer expected to passively interrupt before ERCOT reserves drop below the MCL.</p> <p>ERCOT shall instruct customer to go off-line when reserves reach the MCL</p>	Firm Load may passively respond to ERCOT prices

Source: [DRAFT Load Type Characteristics Matrix Small Group 081622](#)



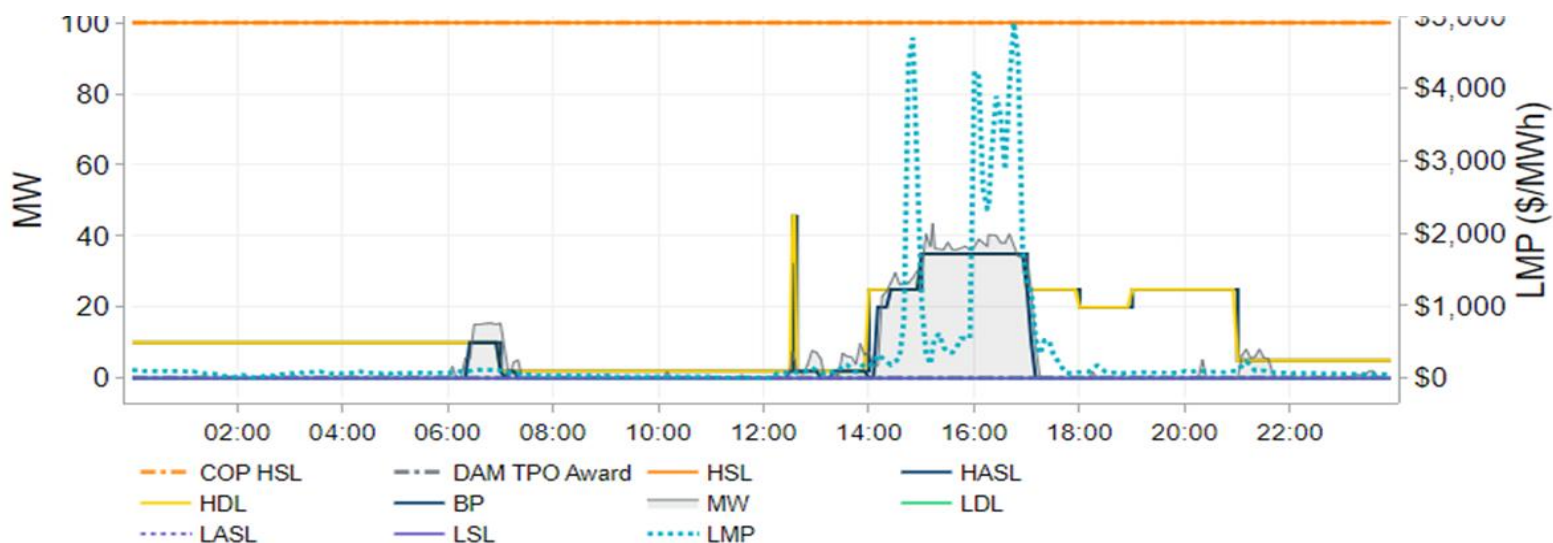
# ERCOT ESR OVERVIEW

- ERCOT is modifying its systems to help address the grid's changing resource mix, including energy storage technologies.
- The changes will allow these emerging technologies to expand their participation in ERCOT's wholesale electricity markets.
- 1,732 MW of installed battery storage (as of April 2022) and 18,000 MW of new battery storage capacity in the interconnection queue.
- Most of ESRs have an energy duration less than 2 hours.



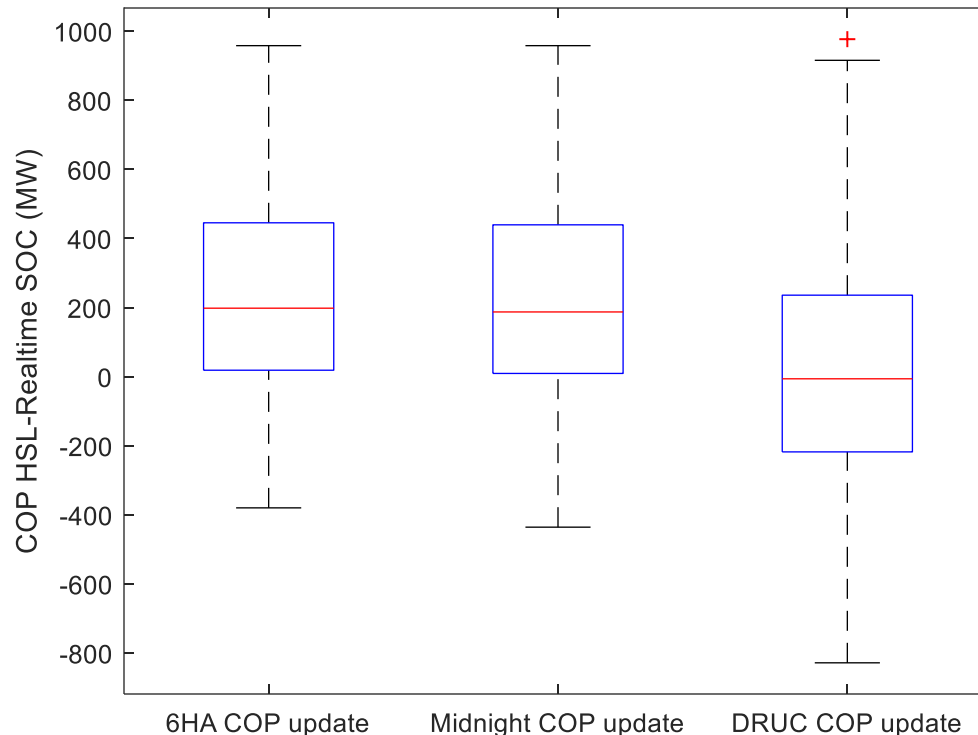
# ESRs

- ESRs inject or withdraw power as a result of following BPs and deployment of Ancillary Service
  - The large ramp rate allows ESRs to change power very quickly
  - The charging/discharging is also subject to State of Charge constraints
  - They may participate in SCED to obtain a BP



# (COP HSL – REAL-TIME SOC) FOR ESRS

- ERCOT expects each ESR QSE to submit a current operating plan (COP) that is based on the QSE's best estimate of the anticipated or expected operating conditions of each of ESR in each of the hours covered by the updated COP.
- If an extra capacity was estimated in COP HSL beyond Real-time SOC, it may over-count the capacity in the future hours.



# ERCOT ESR Modeling and Operation

- “**Combo Model**” – Current approach of representing a battery as a Generation Resource (GR) and a Controllable Load Resource (CLR)
- “Single Model” – Future approach of representing a battery as a single resource
- ESR can provide both Energy and AS if qualified
- **Ancillary Services** are procured to ensure sufficient resource capacity is on-line, or able to be brought on-line in a timely manner, to balance the variability that cannot be covered by the 5-minute energy market.
- Currently, there are three types of Ancillary Services in ERCOT, namely
  - ✓ Regulation (Reg-Up/Reg-Down) Service
  - ✓ Responsive Reserve Service (RRS)
  - ✓ Non-Spinning Reserve Service (Non-Spin)

**HSL** - High Sustained Limit

**HASL** - High Ancillary Service Limit

**HDL** - High Dispatch Limit

**LDL** - Low Dispatch Limit

**LASL** - Low Ancillary Service Limit

**LSL** - Low Sustained Limit

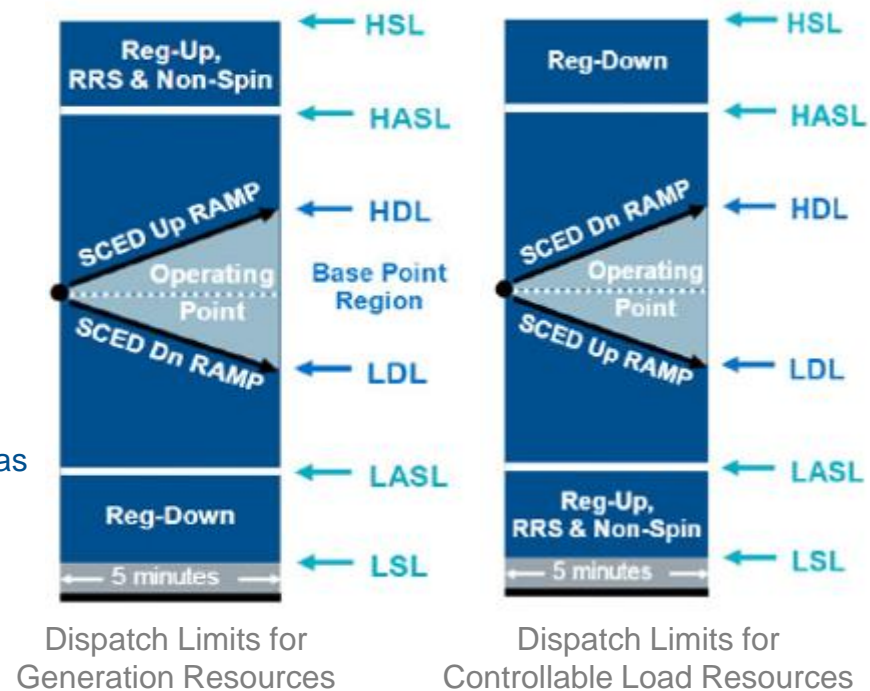
**Storage Resource Limit** (Only for storage resource modeled as Both Generation Resource and Controllable Load Resource):

**MXOS** - Maximum Operating State of Charge

**MNOS** - Minimum Operating State of Charge

**MXCP** - Maximum Operating Charge Power Limit

**MDCP** - Maximum Operating Discharge Power Limit





# SPLIT AS PART OF STATE OF CHARGE (SOC)

- Deployed Ancillary Services (AS) provided by **Generation Resources**:

$$\mathbf{GEN\_AS = GEN\_REG + GEN\_NSP + GEN\_RES}$$

where:

$$\mathbf{GEN\_REG = RGUASD\_UNIT - RGDASD\_UNIT \text{ (Regulation Service Deployed)}}$$

- RGUASD\_UNIT (MW): Regulation Up Ancillary Service Deployed
- RGDASD\_UNIT (MW): Regulation Down Ancillary Service Deployed

$$\mathbf{GEN\_NSP = NSDeployed10 + NSDeployed30 \text{ (Non-Spinning Service Deployed)}}$$

- NSDeployed10 (MW): 10-min Non-Spinning Service Deployed
- NSDeployed30 (MW): 30-min Non-Spinning Service Deployed

$$\mathbf{GEN\_RES = RRRS - RRSC \text{ (Responsive Reserve Service Deployed)}}$$

- RRRS (MW): Responsive Reserve Responsibility
- RRSC (MW): Responsive Reserve Schedule

- Deployed Ancillary Services (AS) provided by **Controllable Load Resources** :

$$\mathbf{LOD\_AS = LOD\_REG + LOD\_NSP + LOD\_RES}$$

where:

$$\mathbf{LOD\_REG = RDRS * RDPF - RURS * RUPF \text{ (Regulation Service Deployed)}}$$

- RURS (MW): Regulation Up Responsibility
- RDRS (MW): Regulation Down Responsibility
- RUPF (%): Regulation Up Participation Factor
- RDPF (%): Regulation Down Participation Factor

$$\mathbf{LOD\_NSP = NDPL \text{ (Non-Spinning Service Deployed)}}$$

- NDPL (MW): Non-Spinning Service Deployed

$$\mathbf{LOD\_RES = RRRS - RRSC \text{ (Responsive Reserve Service Deployed)}}$$

- RRRS (MW): Responsive Reserve Responsibility
- RRSC (MW): Responsive Reserve Schedule

# CHANGE OF STATE OF CHARGE (SOC) DUE TO ECONOMIC DISPATCH

- Adjusted SOC:  **$SOC\_ADJ = SOC - (LOD\_AS - GEN\_AS) * Hour$**

subject to SOC MAX Limit and MIN Limit:

$$SOC\_ADJ\_MAX = SOC\_MAX - GEN\_AS\_Resp * Hour$$

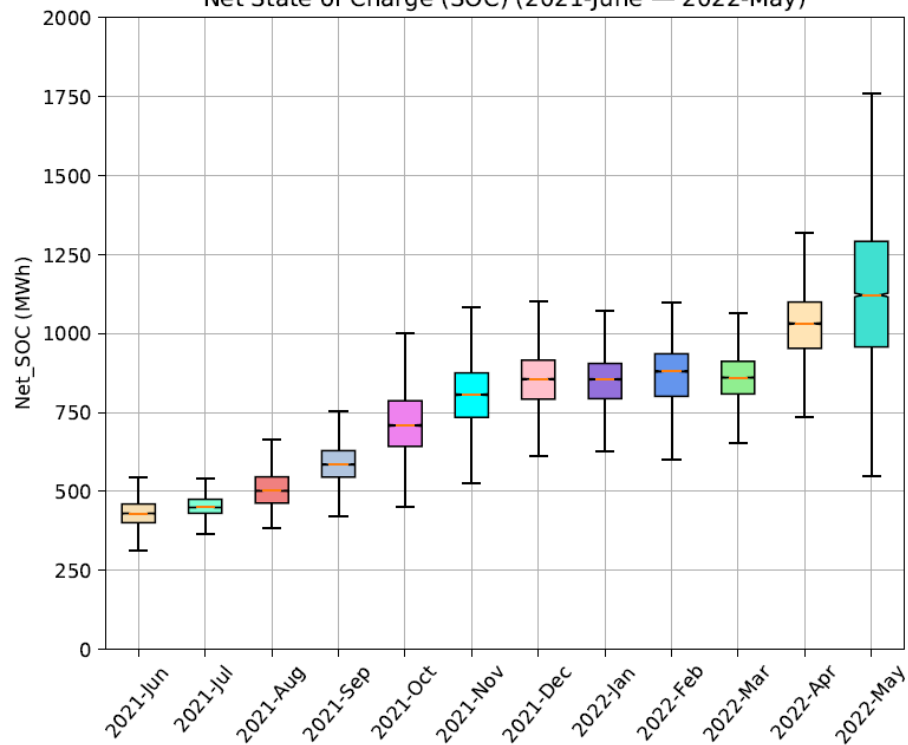
$$SOC\_ADJ\_MIN = SOC\_MIN + LOD\_AS\_Resp * Hour$$

where:

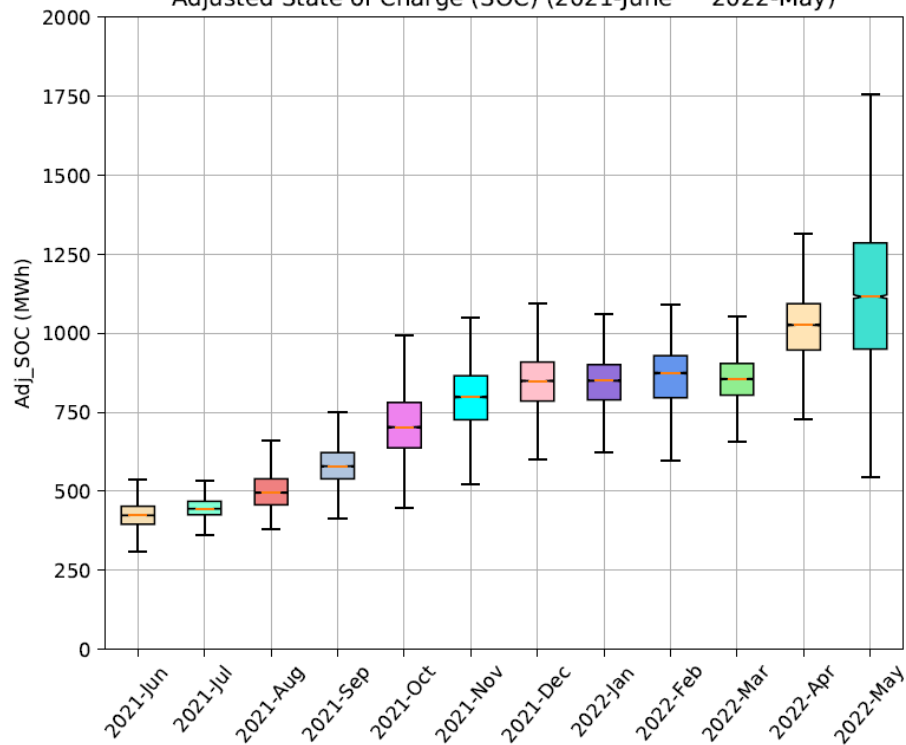
$$- GEN\_AS\_Resp = (ReUp\_Resp + ReDn\_Resp) + NonSpin\_Resp + RRS\_Resp$$

$$- LOD\_AS\_Resp = (ReUp\_Resp + ReDn\_Resp) + NonSpin\_Resp + RRS\_Resp$$

Net State of Charge (SOC) (2021-June — 2022-May)

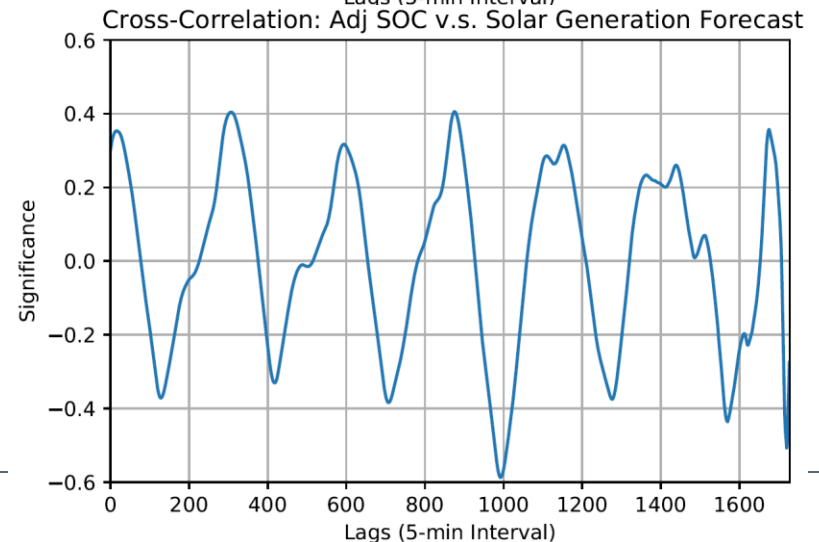
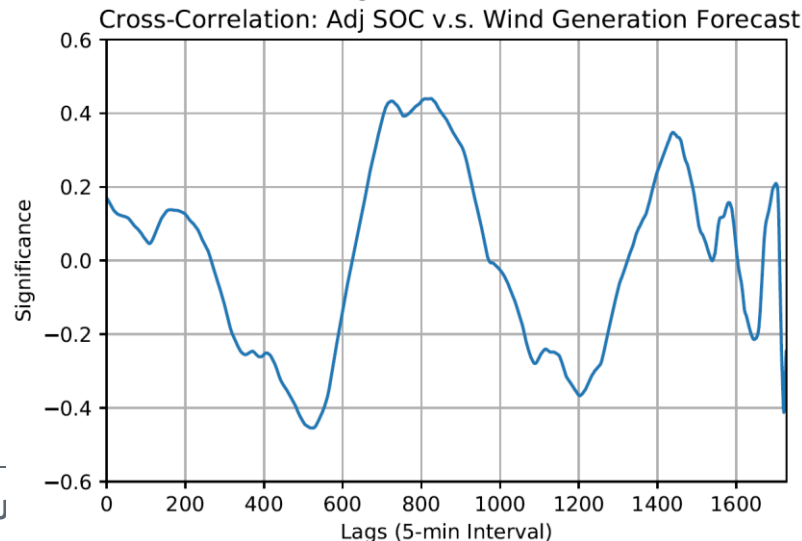
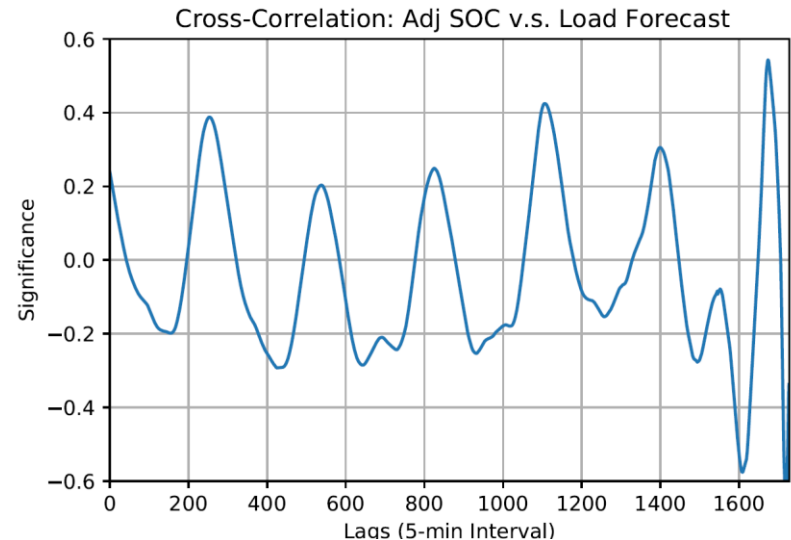
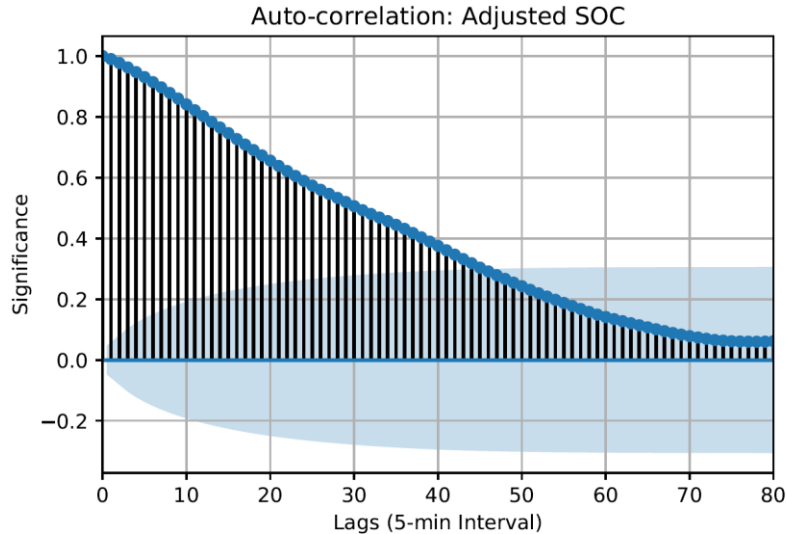


Adjusted State of Charge (SOC) (2021-June — 2022-May)



# CORRELATION ANALYSIS OF FEATURES

- SOC has strong autocorrelation up to 2 hours.
- SOC has strong cross-correlation with load forecast (LF), wind generation forecast (WF), and solar generation forecast (SF).



# FORECAST SETTINGS

- Developed a demonstration tool (Python/Scikit-Learn and SQL).
- Two modes: Off-line mode and on-line mode.
- Major objectives:
  - (Off-line mode) to evaluate performance of forecast models using historical data.
  - (On-line mode) to demonstrate operation of the tool using real-time data.
- Historical data used: June 2021 – May 2022 (12 months).
- After a screening study on several available forecast models, four models have been selected:

Models	Advantages
Decision Tree (DT)	Reliable performance for all types of data problems
Support Vector Regression (SVR)	
Huber Regression (HUB)	Robust linear regression and highly insensitive to outliers
Theil Sen Regression (TSN)	

- Four scenarios have been studied:

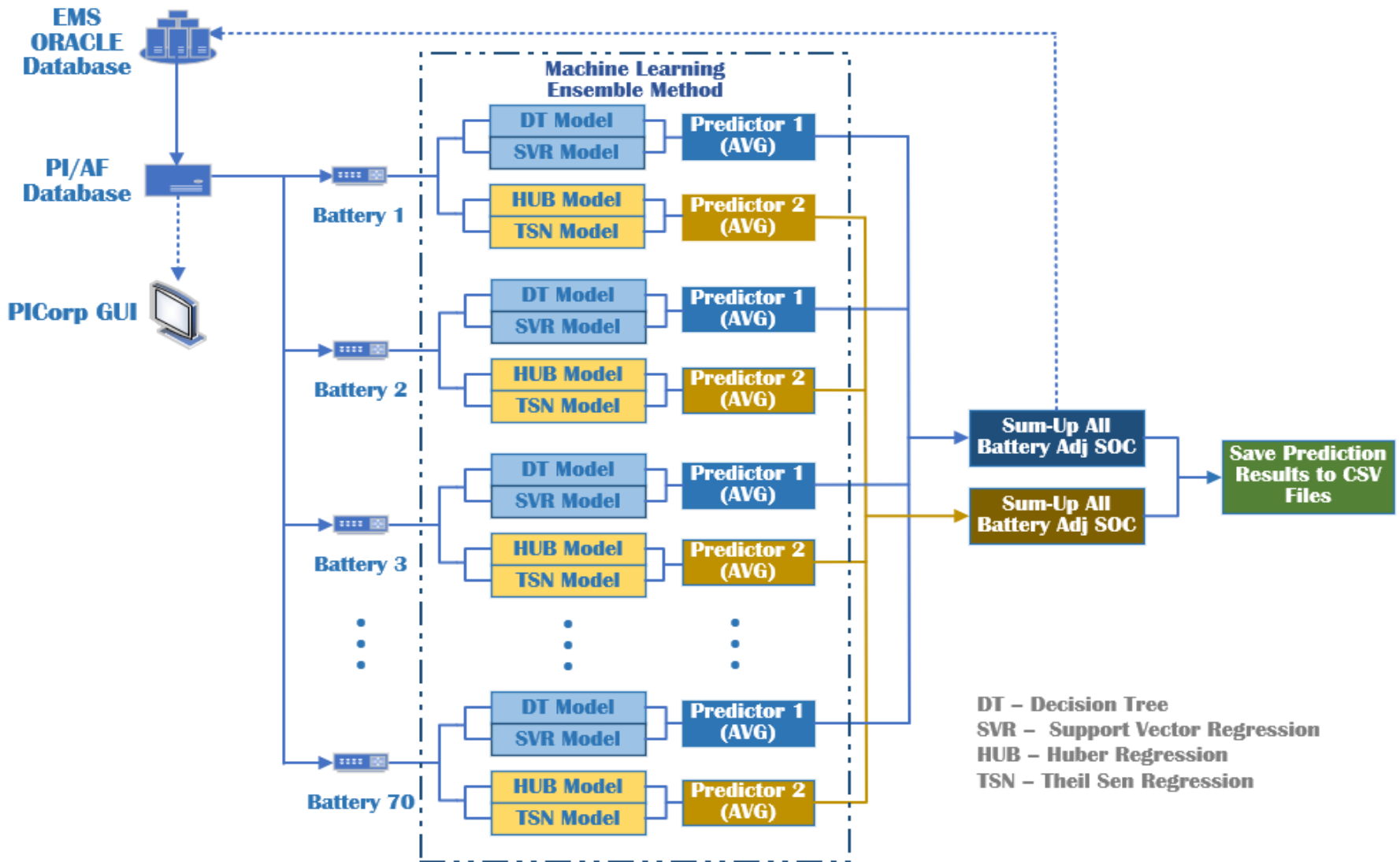
	Feature Group (FG) 1	Feature Group (FG) 2
Predictor 1 (DT+SVR)	Scenario 1	Scenario 2
Predictor 2 (HUB+TSN)	Scenario 3	Scenario 4

- Predictor 1 (DT+SVR): average of results of DT and SVR models as prediction output
- Predictor 2 (HUB+TSN): average of results of HUB and TSN models as prediction output
- Feature Group 1 (FG1): WF, SF, LF
- Feature Group 2 (FG2): WF, SF, LF, HSL, LSL, UDBP (Updated Desired Base Point)

# FORECAST MODEL PARAMETER SETTINGS

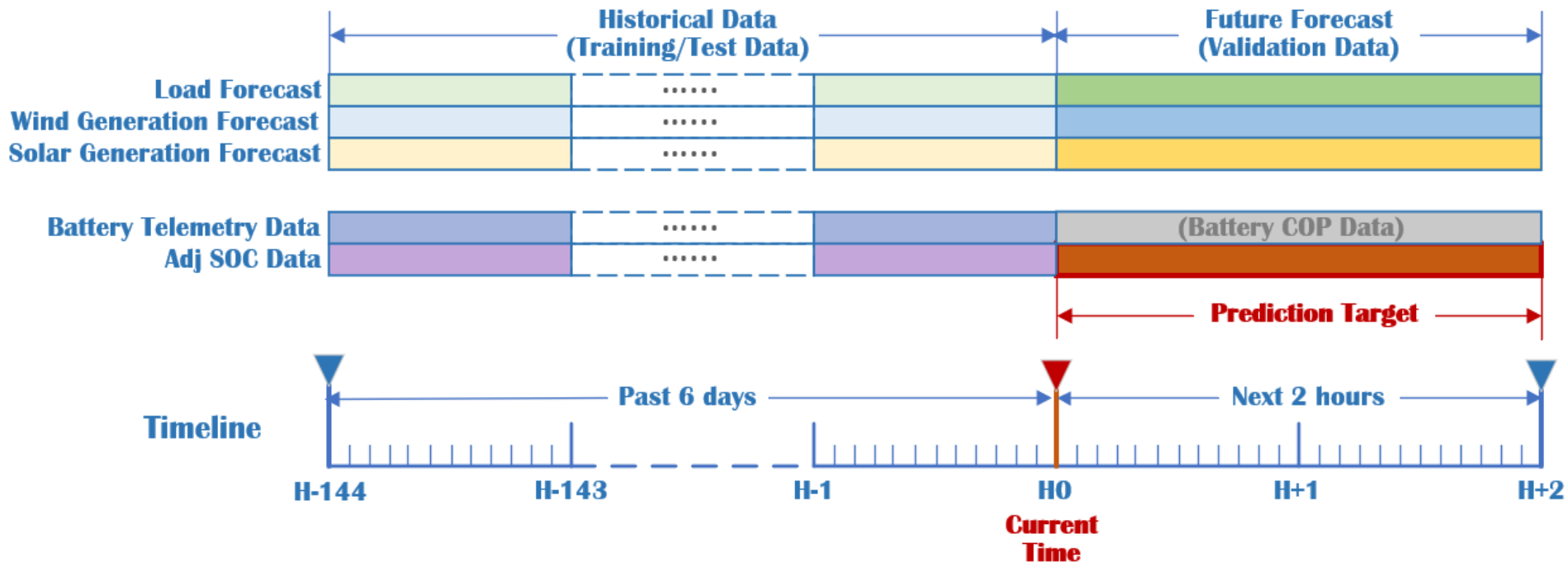
- Hyperparameters of models were tuned by **Grid Search** (a tuning technique that attempts to compute the optimum values of hyperparameters).
- Grid Search is an exhaustive search performed on specific parameter values of a model.
- Decision Tree (DT)
  - Maximum tree of depth (*max\_depth*)
  - Minimum samples for a node split (*min\_samples\_split*)
  - Minimum samples for a terminal node (*min\_samples\_leaf*)
  - Maximum number of terminal nodes (*max\_leaf\_nodes*)
  - Maximum features to consider for split (*max\_feature*)
- Support Vector Regression (SVR)
  - Kernel type (kernel)
  - Degree of the polynomial kernel function (*degree*)
  - Regularization parameter (*C*)
  - Epsilon in the epsilon-SVR model (*epsilon*)
  - Kernel coefficient for 'rbf', 'poly' and 'sigmoid' (*gamma*)
- Huber Regression (HUB)
  - Epsilon controls number of samples classified as outliers (*epsilon*)
  - Maximum number of iterations (*max\_iter*)
  - Strength of squared L2 regularization (*alpha*)
- Theil Sen Regression (TSN)
  - Maximum number of subsamples (*max\_subpopulation*)
  - Maximum number of iterations for calculation of spatial median (*max\_iter*)
  - Tolerance when calculating spatial media (*tol*)

# FORECAST DATAFLOW



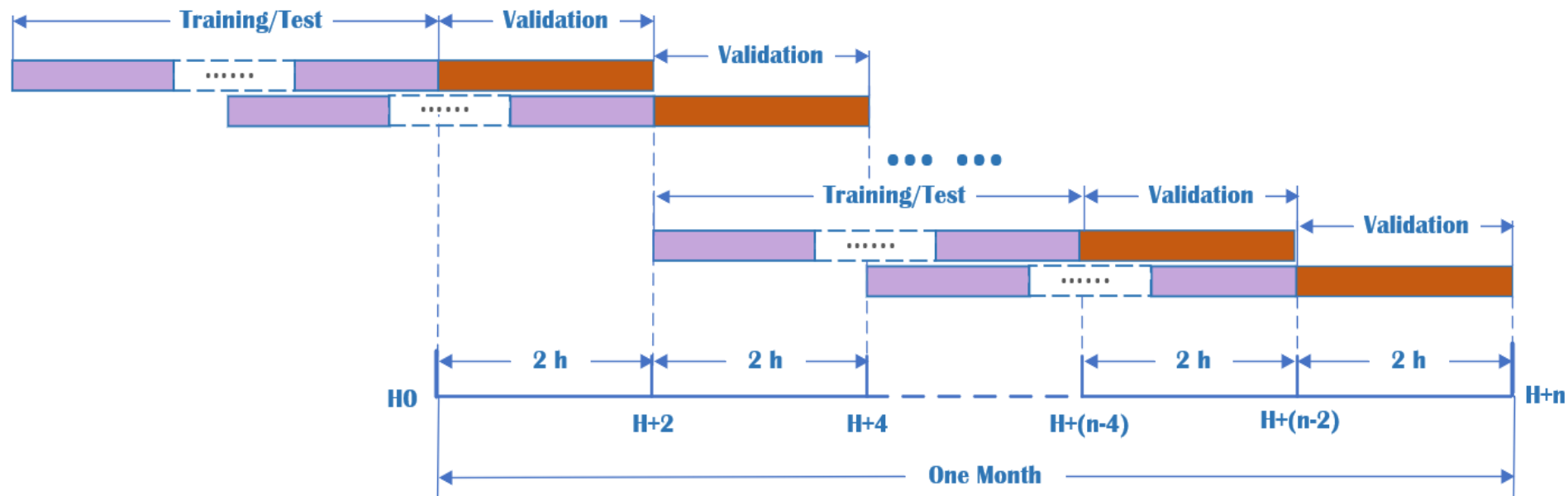
# FORECAST TIMELINE (GENERIC)

- Use past 6 days 5-minute interval historical data as training/test data (80/20 split)
- Input features: Load, wind, solar forecast data (and battery telemetry data and COP data)
- Predict adjusted SOC for next 2 hours in 5-minute interval (validation data)



# OFF-LINE MODE TIMELINE

- Predictor 1 (DT+SVR) & Predictor 2 (HUB+TSN) run on historical data at 2-hour increment steps
- Average of 24 5-min interval forecast errors on each step is calculated as the forecast error for that step
- Apply on 12 months (2021-June – 2022-May)
- Forecast errors are summarized on month by month





# FORECAST PERFORMANCE METRICS

- Forecast error metrics:
  - MAPE (Mean Absolute Percentage Error)
  - MAE (Mean Absolute Error)
  - MSE (Mean Squared Error)
  - RMSE (Root-Mean-Square Error)
- MAPE** has the advantage of being scale-independent.
- MAE** is an unambiguous measure of average error.
- RMSE and MSE are more sensitive to outliers than MAPE and MAE.
- This study uses MAPE and MAE to evaluate average model performance.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right|$$

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

$$MSE = \frac{\sum (y_t - \hat{y}_t)^2}{n}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$$

where:

$n$  is the number of fitted points,  
 $y_t$  is the actual value,  
 $\hat{y}_t$  is the forecast value.

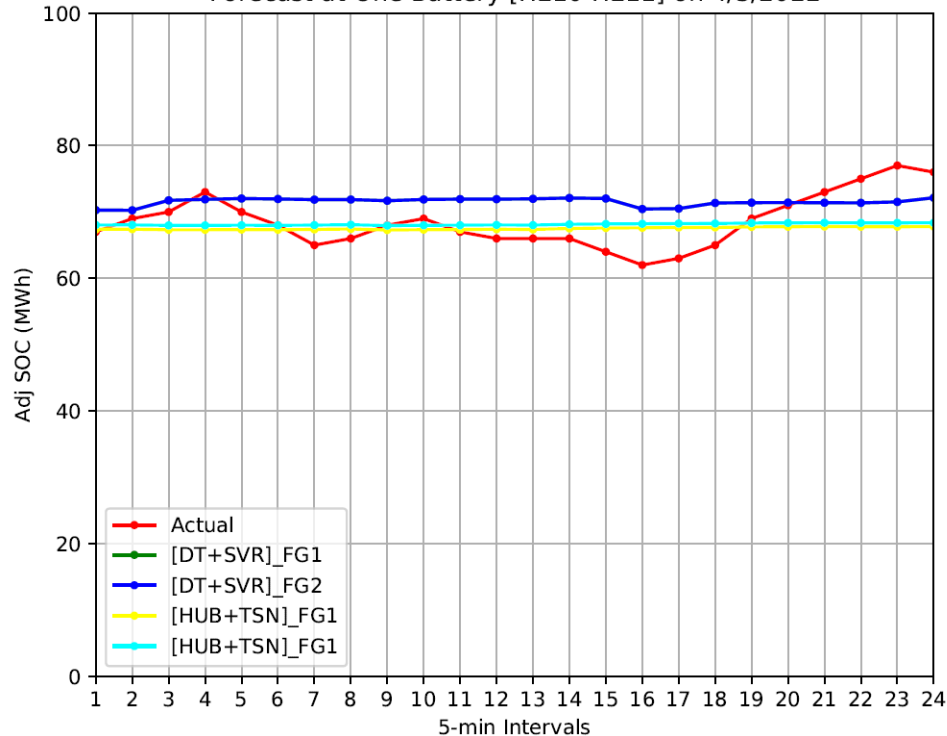
MAPE Value	Level of Accuracy
MAPE ≤ 10%	Very Accurate
10% < MAPE ≤ 20%	Accurate
20% < MAPE ≤ 50%	Medium
50% ≤ MAPE	Less Accurate

MAPE	Equivalent MAE Metrics for ADJ_SOC (MWh) (Based on 95 Percentile)											
	2021-6	2021-7	2021-8	2021-9	2021-10	2021-11	2021-12	2022-1	2022-2	2022-3	2022-4	2022-5
50%	248	257	302	347	452	479	497	485	503	490	598	725
20%	99	103	121	139	181	191	199	194	201	196	239	290
10%	50	51	60	69	90	96	99	97	101	98	120	145

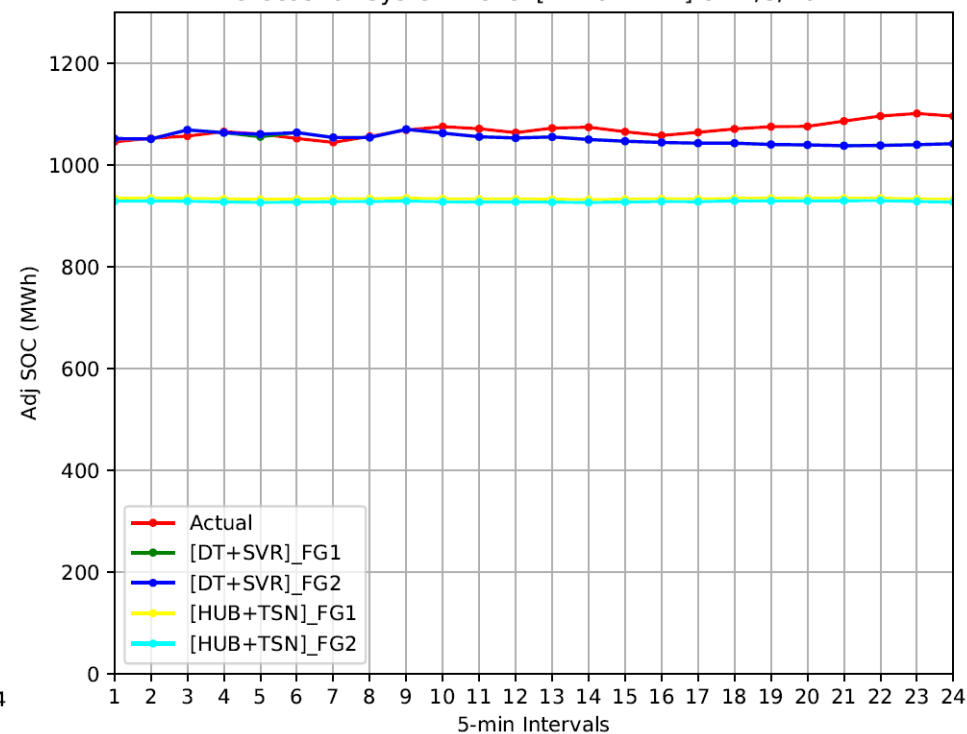
# FORECAST RESULTS

- Forecast models, a.k.a. Predictor 1 (DT+SVR) and Predictor 2 (HUB+TSN), are applied to each of the 70 batteries.
- Forecast result of each battery is summed-up to get system level forecast result.

Forecast at One Battery [HE10-HE11] on 4/3/2022



Forecast at System Level [HE10-HE11] on 4/3/2022



# HEATMAP OF MAPE

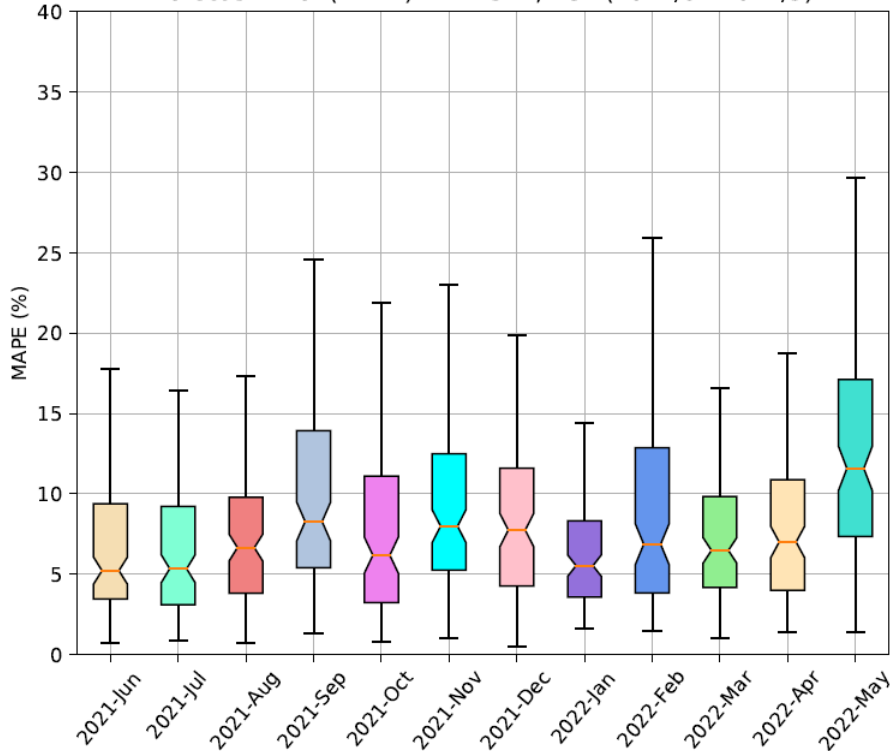
- Heatmap MAPE of 5-min intervals of the next 2-hour duration.
- Average errors gradually increase from the 1st to the last of 5-min intervals.

5-min Interval	2021-06	2021-07	2021-08	2021-09	2021-10	2021-11	2021-12	2022-01	2022-02	2022-03	2022-04	2022-05
1	5.52	4.15	4.18	6.66	5	7.14	6.58	4.2	4.41	4.68	5.32	11.37
2	5.71	4.89	4.61	7.18	5.41	7.1	7.25	4.65	4.91	5.16	5.74	12.43
3	6.17	5.18	5.13	7.68	6.13	7.95	7.37	5.24	5.48	5.6	6.14	12.78
4	6.09	5.52	5.31	7.76	5.98	7.93	7.79	5.74	5.74	5.98	6.21	13.41
5	6.41	5.67	5.5	8.23	6.25	8.3	8.06	5.43	6.44	6.41	6.73	14.02
6	6.74	5.9	5.82	8.91	6.91	8.51	8.38	5.54	6.84	6.64	7.32	14.94
7	7.01	6.17	6.51	9.59	7.01	9.01	8.43	5.79	6.84	6.92	7.42	14.67
8	7.34	6.36	6.37	9.71	7.08	8.75	8.48	5.95	6.98	7.21	7.72	14.87
9	7.45	6.66	6.46	10	7.28	8.7	8.72	5.97	7.09	7.32	7.76	15.01
10	7.88	6.93	7.24	10.32	7.36	8.76	8.98	6.3	7.55	7.62	8	15.17
11	8.23	7.28	7.03	10.35	7.52	9.28	9.34	6.25	8.05	7.68	8.12	15.4
12	8.35	7.24	7.79	10.67	7.86	9.61	9.57	6.33	8.27	7.74	8.24	15.75
13	8.53	7.33	7.87	11.3	7.8	9.44	9.73	6.47	8.95	8.05	8.51	16.38
14	8.48	7.5	8.28	11.28	7.95	9.73	9.92	6.69	9.11	8.34	8.71	16.67
15	8.81	7.64	8.33	11.2	7.71	10.12	10	7.15	9.33	8.56	8.7	16.96
16	9.01	7.82	8.5	11.25	7.99	10.83	9.85	7.15	9.86	8.96	9.1	16.71
17	9.05	7.85	8.89	11.49	8.3	10.86	10.07	7.18	9.76	8.99	9.22	17.09
18	9.35	8.01	8.95	11.24	8.38	11.36	10.28	7.24	9.85	8.79	9.05	17.14
19	9.42	8.35	9.07	10.94	8.74	11.13	10.58	7.19	9.8	9.06	9.35	17.39
20	9.75	8.53	9.03	10.91	8.87	11.8	10.18	7.31	10.02	9.28	9.72	17.9
21	9.58	8.31	8.89	10.83	8.85	11.98	10.17	7.21	10.1	9.21	9.71	18.34
22	9.59	8.57	9.24	11.32	8.91	12.32	10.45	7.46	10.07	9.28	10.05	19.1
23	9.43	8.5	9.67	10.87	8.85	12.41	10.34	7.44	10.17	9.42	9.91	19.29
24	9.45	8.42	10	11.25	8.93	12.76	10.52	7.21	9.98	9.33	9.42	19.93

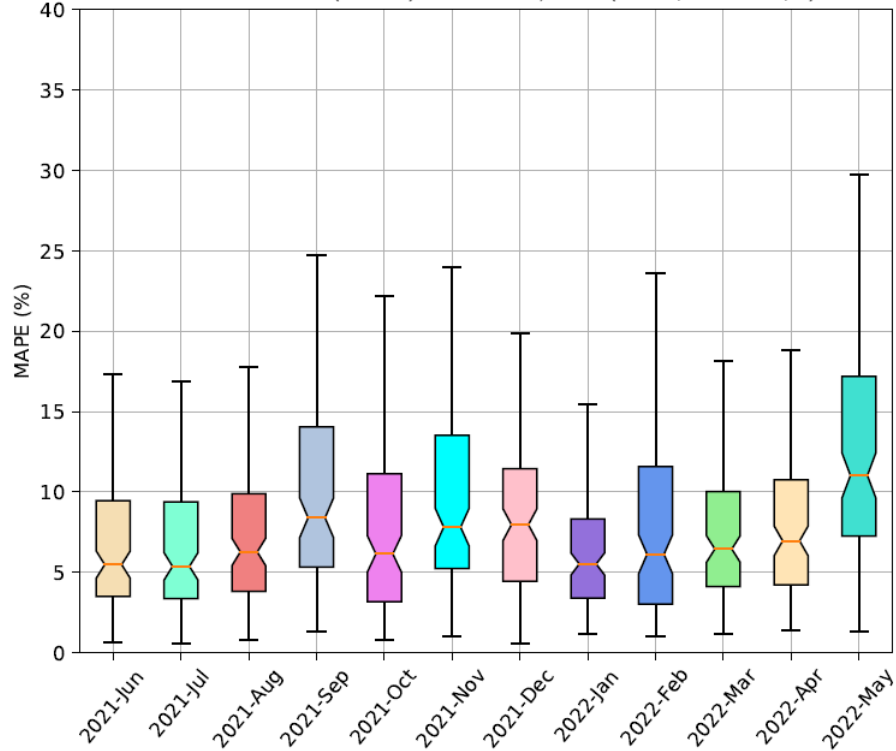
# DISTRIBUTION OF FORECAST ERRORS (MAPE) – PREDICTOR 1

- Increasing number of input features for Predictor 1 (DT+SVR) may not necessarily reduce average forecast errors.

Forecast Error (MAPE) DT + SVR, FG1 (2021/6 - 2022/5)

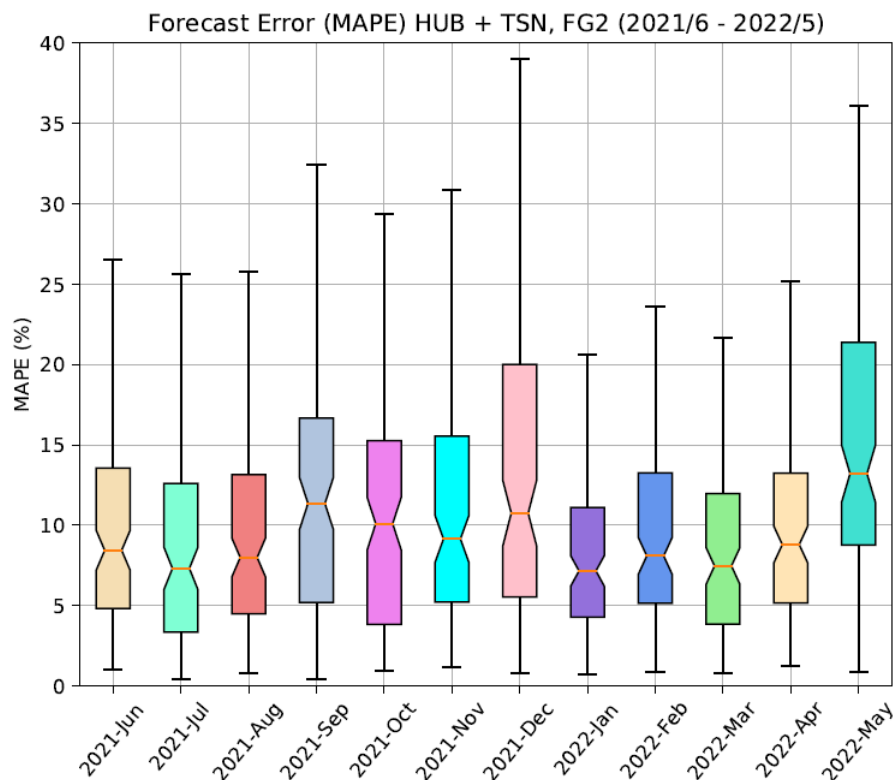
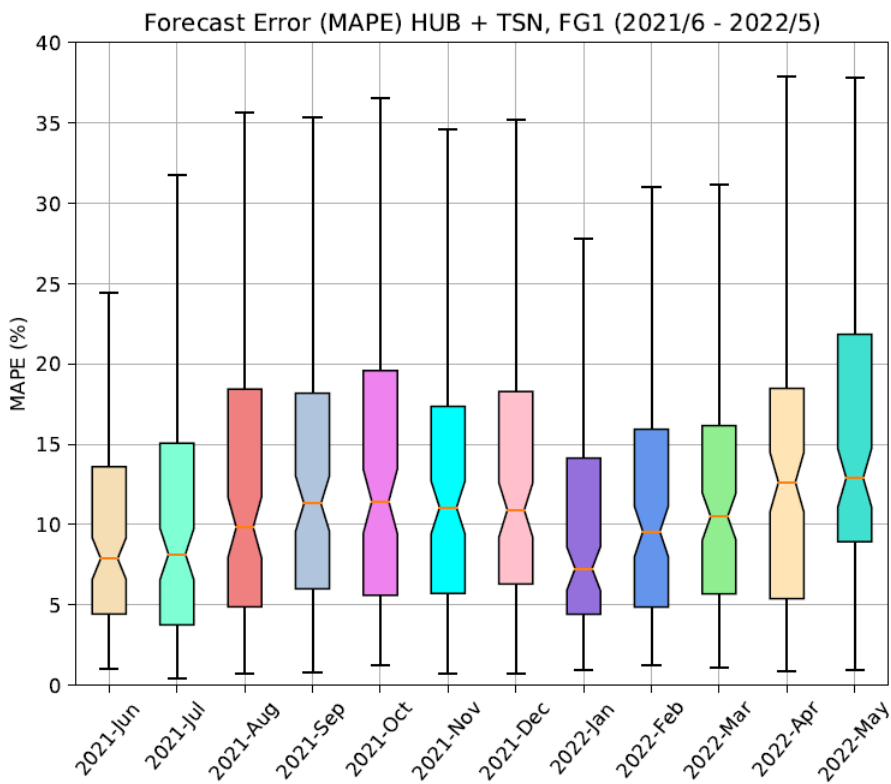


Forecast Error (MAPE) DT + SVR, FG2 (2021/6 - 2022/5)



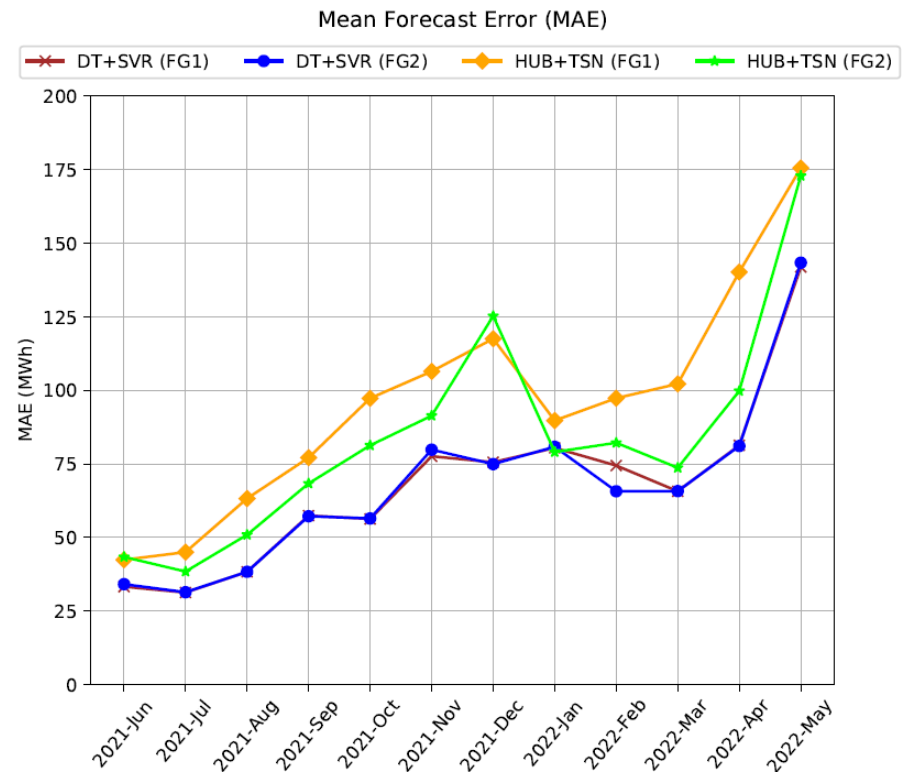
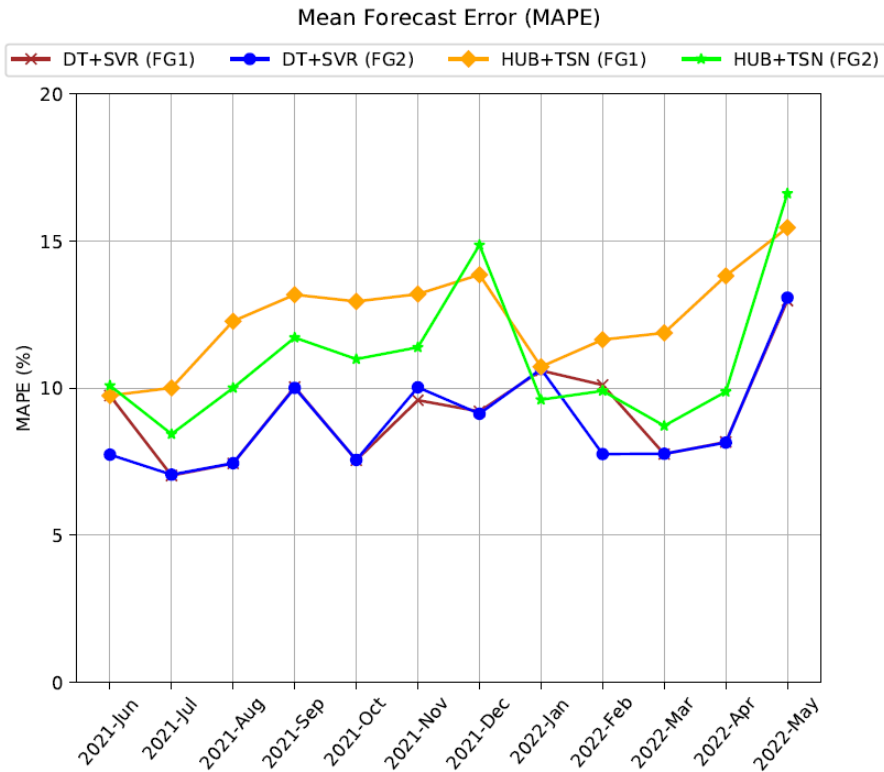
# DISTRIBUTION OF FORECAST ERRORS (MAPE) – PREDICTOR 2

- Increase number of input features for Predictor 2 (HUB+TSN) may slightly reduce average forecast errors.
- Predictor 2 (HUB+TSN) produced higher average forecast errors than Predictor 1 (DT+SVR) on both feature groups.



# COMPARE OF AVERAGE FORECAST ERRORS

- Predictor 1 (DT+SVR) produced more accurate prediction than Predictor 2 (HUB+TSN) for all months.



# DISCUSSION

- Major sources of forecast errors
  - Telemetry data error (due to communication problems, etc.)
  - Wind and solar ramp events
  - Rapid ramp in charging/discharging
  
- Future work
  - Implement parallel version of training/forecasting (in progress)
  - Implement in SAS environment (in progress)
  - Integrate to EMS for real-time operation (in progress)



Thanks  
&  
Questions?