

NARUC/NASEO Comprehensive Electricity Planning Initiative for Load Growth Learning Series

Load Forecasting for Large Loads

From the perspective of planning across bulk power and distribution systems

Prepared by Greg Mandelman for Berkeley Lab
June 3, 2025

Today's Topics

Load forecasting for large loads – focusing on data centers and manufacturing – in support of reliability, affordability, and security

Topics	Slides
Key Reference Material	3
Overview of Sector-level Large Load Forecasting	4
Range of Data Center Growth Paths	5-6
Data Centers Concentration & Diversity	7
Size & Diversity of Onshore Manufacturing Loads	8
Examples of ISO/RTO Approaches to Data Center Load Forecasting	9-10
Examples of State-of-the-Industry in Data Center Load Forecasting	11-13
The Gold Standard? Bottom-up Data Center Load Forecasting	14
Questions States Can Ask	15
Actions States Can Take	16

Key Reference Materials for Decision-Makers

These documents provide a balanced, robust, well-evidenced knowledge base on large loads

Background & Context



GridStrategies, *Strategic Industries Surging: Driving US Power Demand*, December 2024

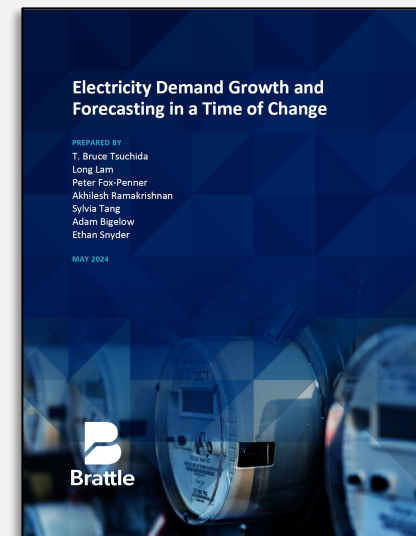
- Outstanding breadth of coverage, with insightful analysis
- If you can only read one document, read this one



The American Clean Power Association, *US National Power Demand Study*, March 2025

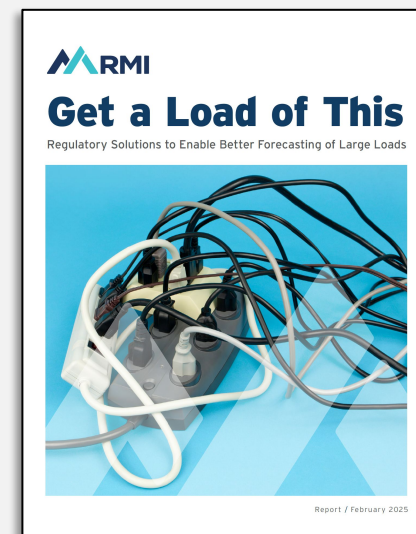
- Completed by S&P Global Commodity Insights, this is an expansive and detailed analysis of generation demand
- Excellent regional coverage at the RTO / ISO level

Load Forecasting Analysis



The Brattle Group, *Electricity Demand Growth and Forecasting in a Time of Change*, May 2024

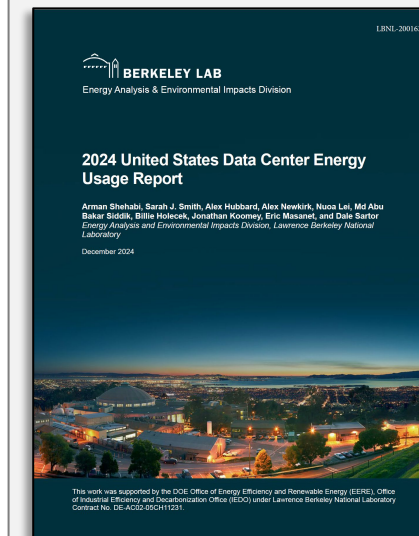
- Excellent historical background on load forecasting and load growth
- Using historic examples, illustrates the risks associated with over- and under forecasting load growth



The Rocky Mountain Institute, *Get a Load of This: Regulatory solutions to enable better forecasting of large loads*, February 2025

- Broad treatment of different types of load forecasting, including data centers and manufacturing
- Offers specific guidance for regulators, including “Discovery questions for load forecasting”

Sector Specific Analysis



Lawrence Berkeley National Lab, *2024 United States Data Center Energy Usage Report*, December 2024

- Comprehensive modeling and analysis of historic data center loads and the potential range of future loads
- Relatively technical, but still accessible
- Want to go deeper? Read these:
 - “[To better understand AI’s growing energy use, analysts need a data revolution](#)”
 - [Datacenter Anatomy Part 1: Electrical Systems](#); [Datacenter Anatomy Part 2 – Cooling Systems](#); [AI Datacenter Energy Dilemma – Race for AI Datacenter Space](#)



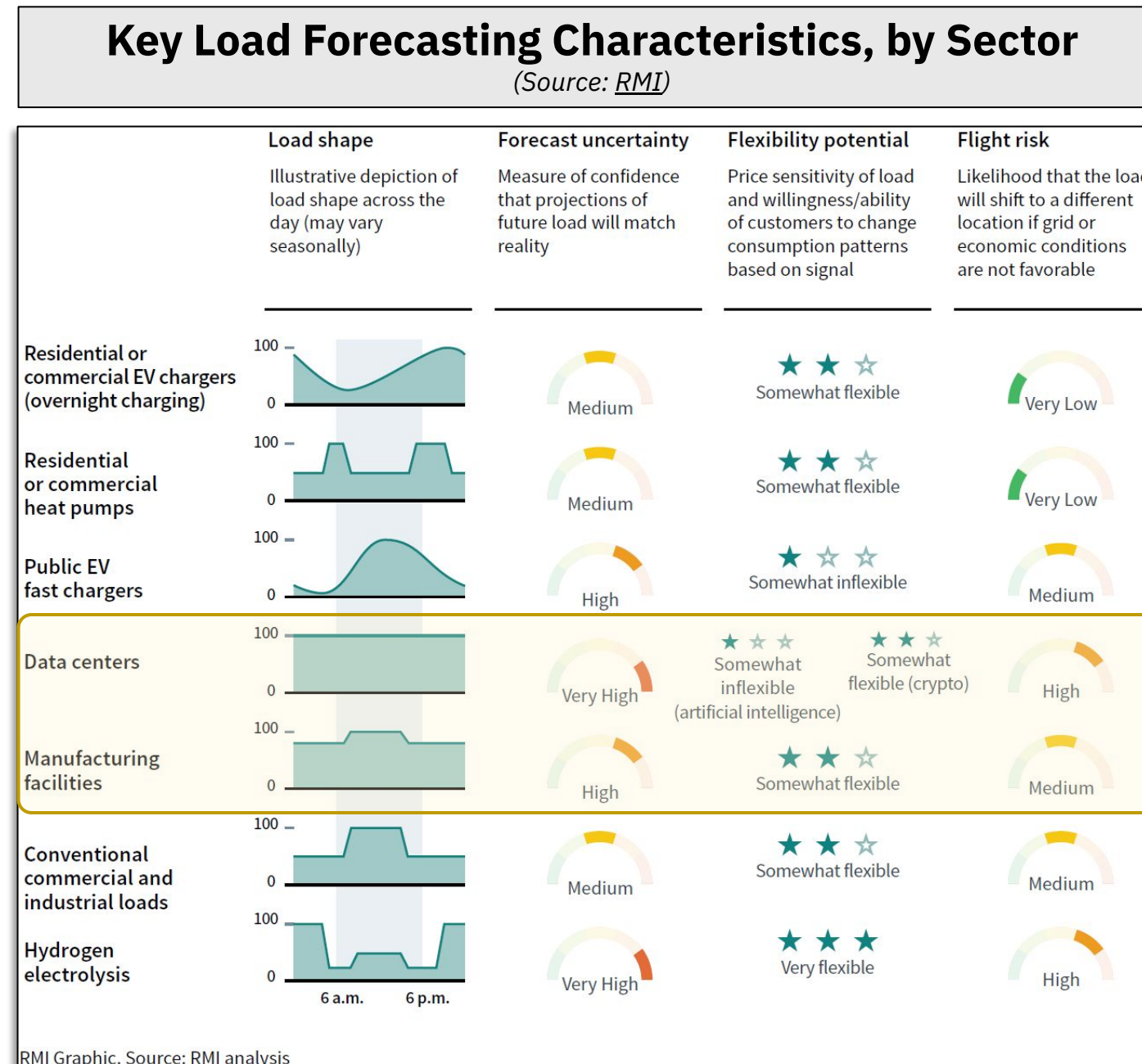
EPRI, *Reindustrialization, Decarbonization, and Prospects for Demand Growth*, July 2023

- Offers useful geographic and subsector-specific information for new US manufacturing electricity demand
- Be aware: EV & battery manufacturing onshoring were projected to be the largest growth sectors, and current policy and potential legislation may negatively impact these plans

Large Loads Comes in Different Sizes & Shapes

Data center (and manufacturing) loads present significant uncertainty in the near- to mid-term

- **Among the sources of large loads, data center loads are significant in terms of:**
 - The overall level of uncertainty associated with their forecasts
 - Their ability to easily move load to different geographies
- **As we will see, there are many different types of data centers, which leads to:**
 - Varying levels of overall demand and ramp rates
 - Differing operating characteristics
- **Unlike other sources of bulk demand, such as manufacturing and hydrogen, data center demand is:**
 - Much larger on an individual customer basis, more dispersed, and greater in overall scale across the industry
 - Relatively less sensitive to federal policy changes



Near-Term Load Drivers

(Source: RMI)

Near-Term Load Drivers	Data Centers	Manufacturing	Electrification
Arizona Public Service	✓		
CAISO	✓		✓
Duke	✓	✓	
ERCOT	✓	✓	
Georgia Power	✓	✓	
ISO-NE			✓
MISO	✓		✓
NYISO	✓	✓	✓
Pacific Northwest	✓	✓	
PJM	✓	✓	✓
SPP	✓		

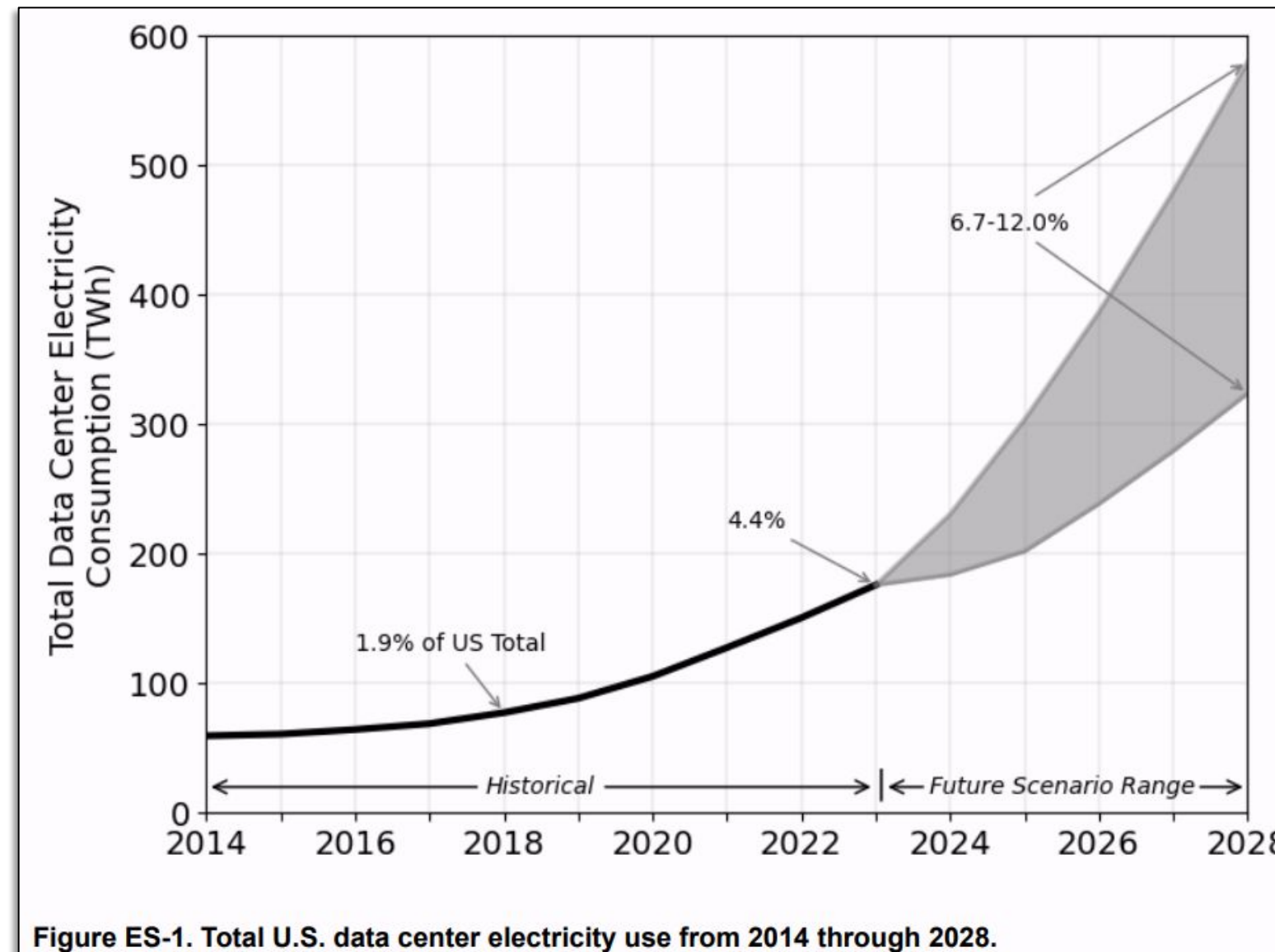
RMI Graphic. Source: Grid Strategies, <https://gridstrategiesllc.com/wp-content/uploads/National-Load-Growth-Report-2024.pdf>

Wide Range of Data Center Growth Trajectories

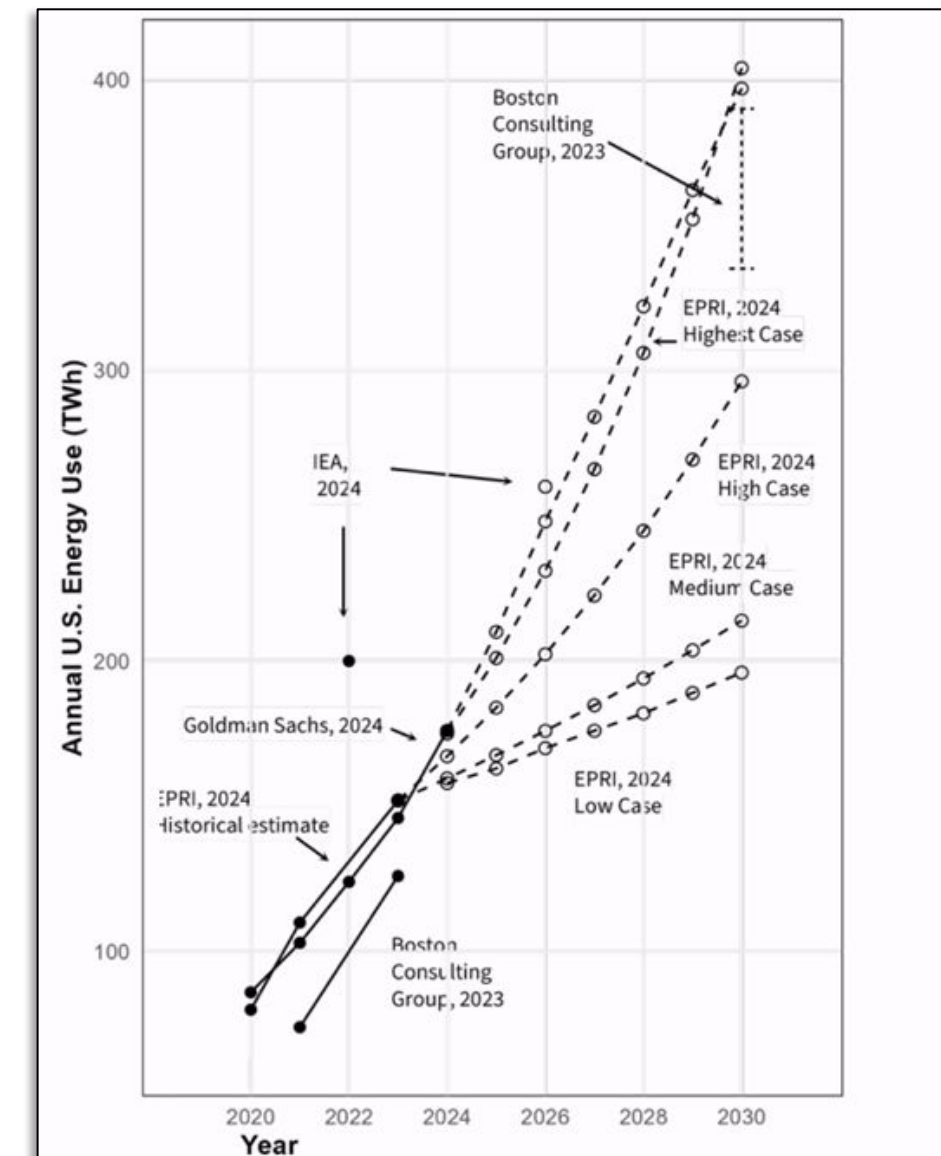
While the uncertainty bands are wide, even “low” data center growth is highly material

- While data centers have been around for decades, several new factors are driving steep load growth, including:
 - Significantly higher power requirements for AI GPUs
 - Large growth in number of AI GPUs
 - Higher utilization of AI GPUs relatively to historic processors
- The spread in the load forecast estimates is driven by many factors with uncertainty, including:
 - Quantities and type of future equipment shipments
 - Operating practices & system utilization
 - Cooling system type & efficiency
- To evaluate data center load growth, consider the following:
 - Is it a top-down or bottom-up analysis?
 - Are the inputs and model transparent?

Historic and Projected Total US Data Center Load
(Source: [LBNL](#))

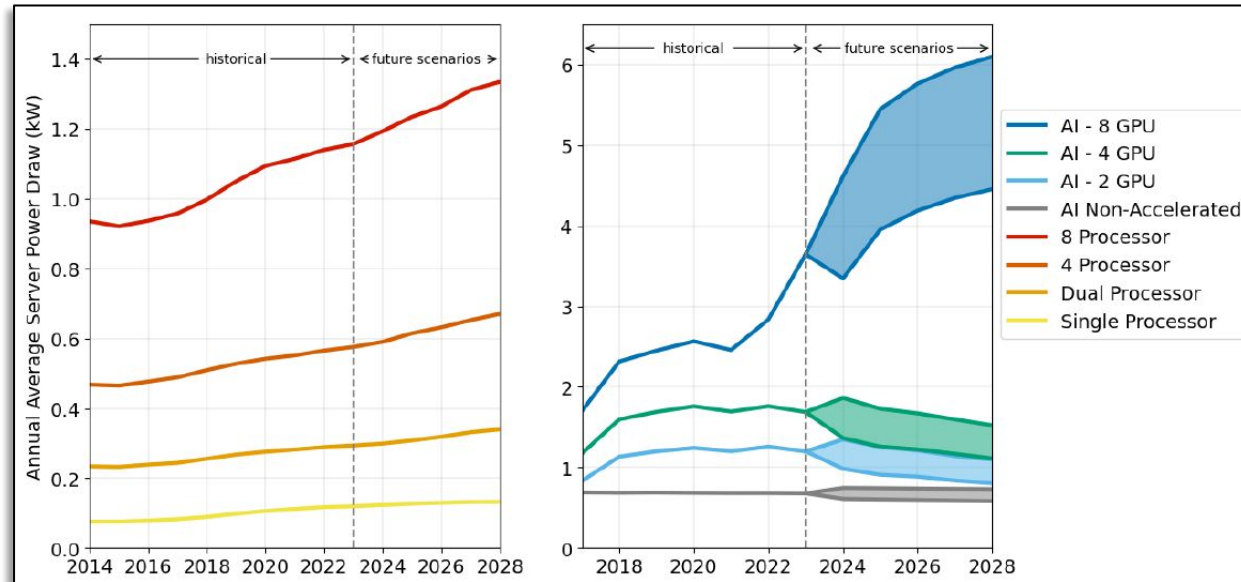


Range of Data Center Forecasts
(Source: [LBNL](#); edited by EPE for clarity)



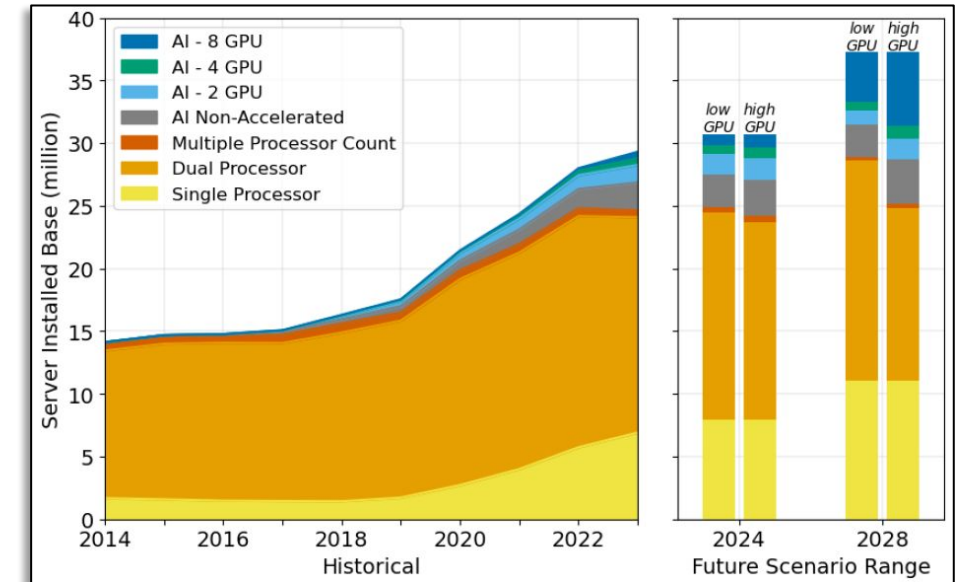
Understanding Data Center Growth in Four Charts

Aggregate average power draw of various server types across each analysis year
(Source: LBNL, Figure 3.7)



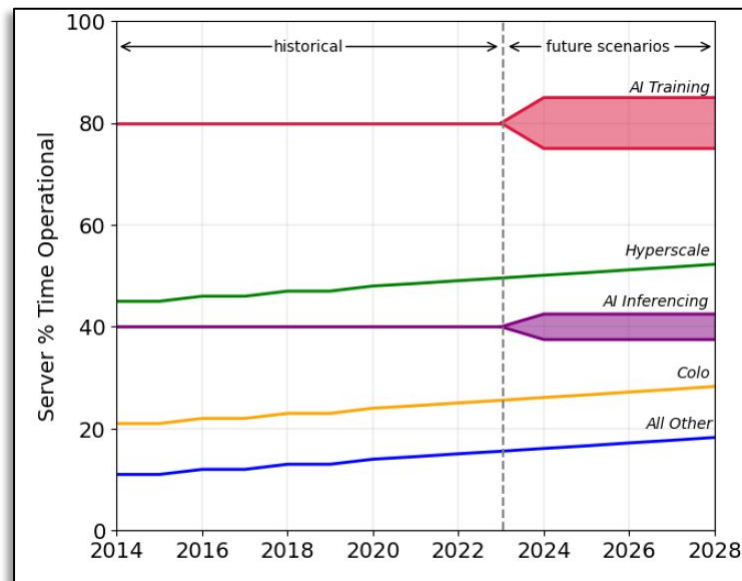
AI GPUs use significantly more energy

Total server installed base for 2014–2028 with higher bound shipments (left). Adjusted installed base with lower bound GPU shipments (right)
(Source: LBNL, Figure 3.9)



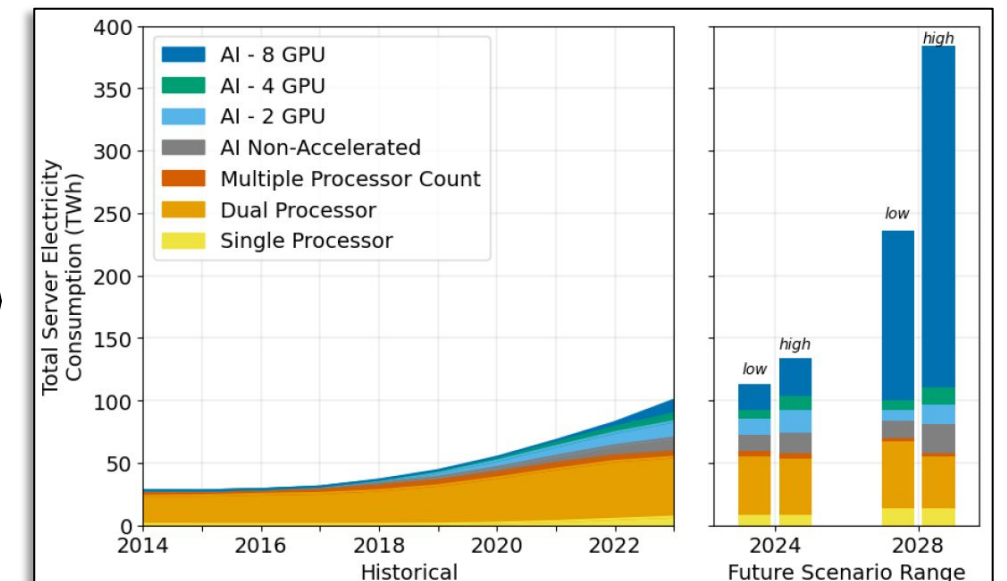
The number of installed AI services is growing significantly

Operational time of servers by server type
(Source: LBNL, Figure 3.6)



Server utilization is increasing with AI

Server annual electricity usage by type
(Source: LBNL, Figure 5.1)

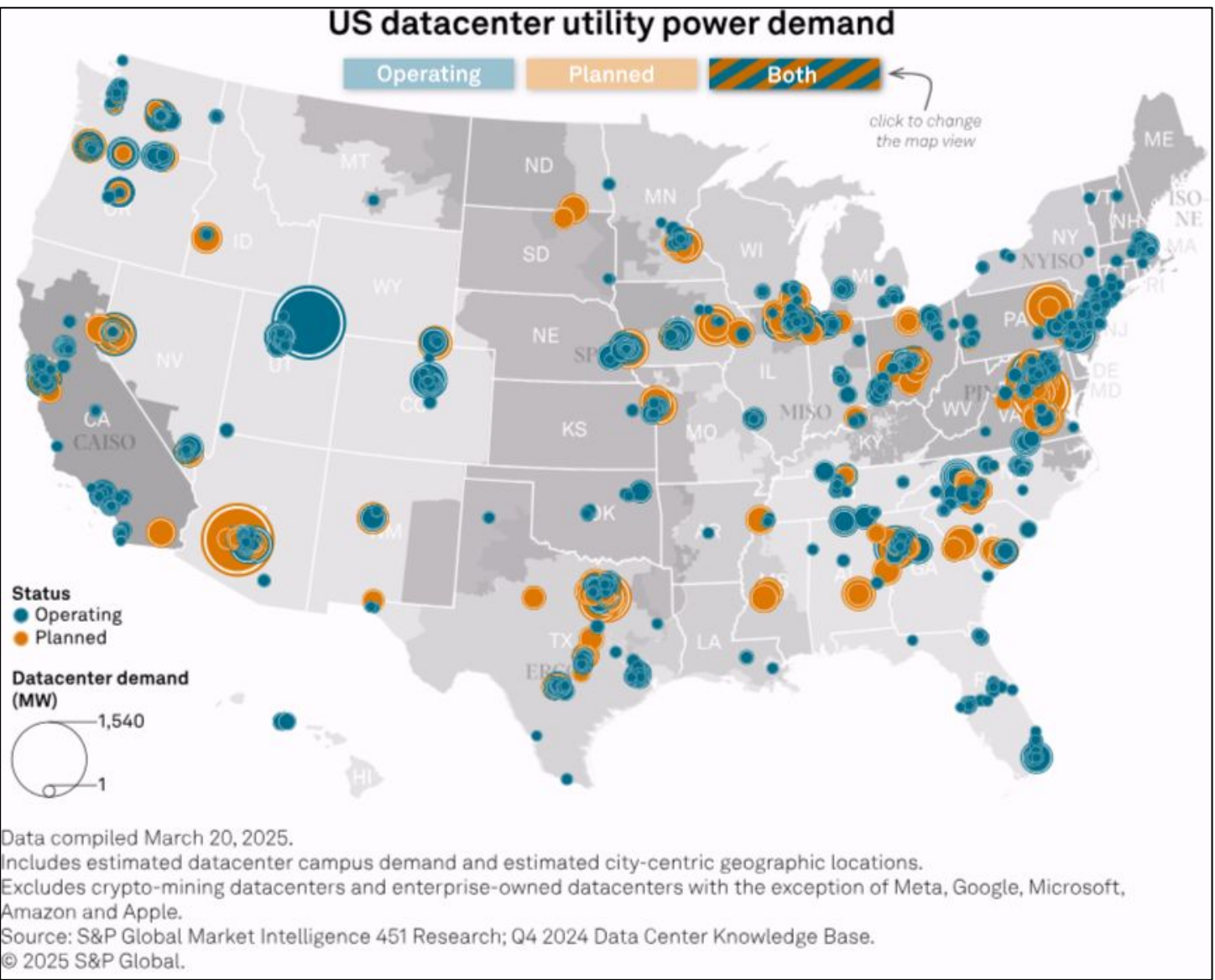


Combined, these three factors lead to dramatic increases in overall data center demand

Data Center Demand Is Concentrated *and* Diverse

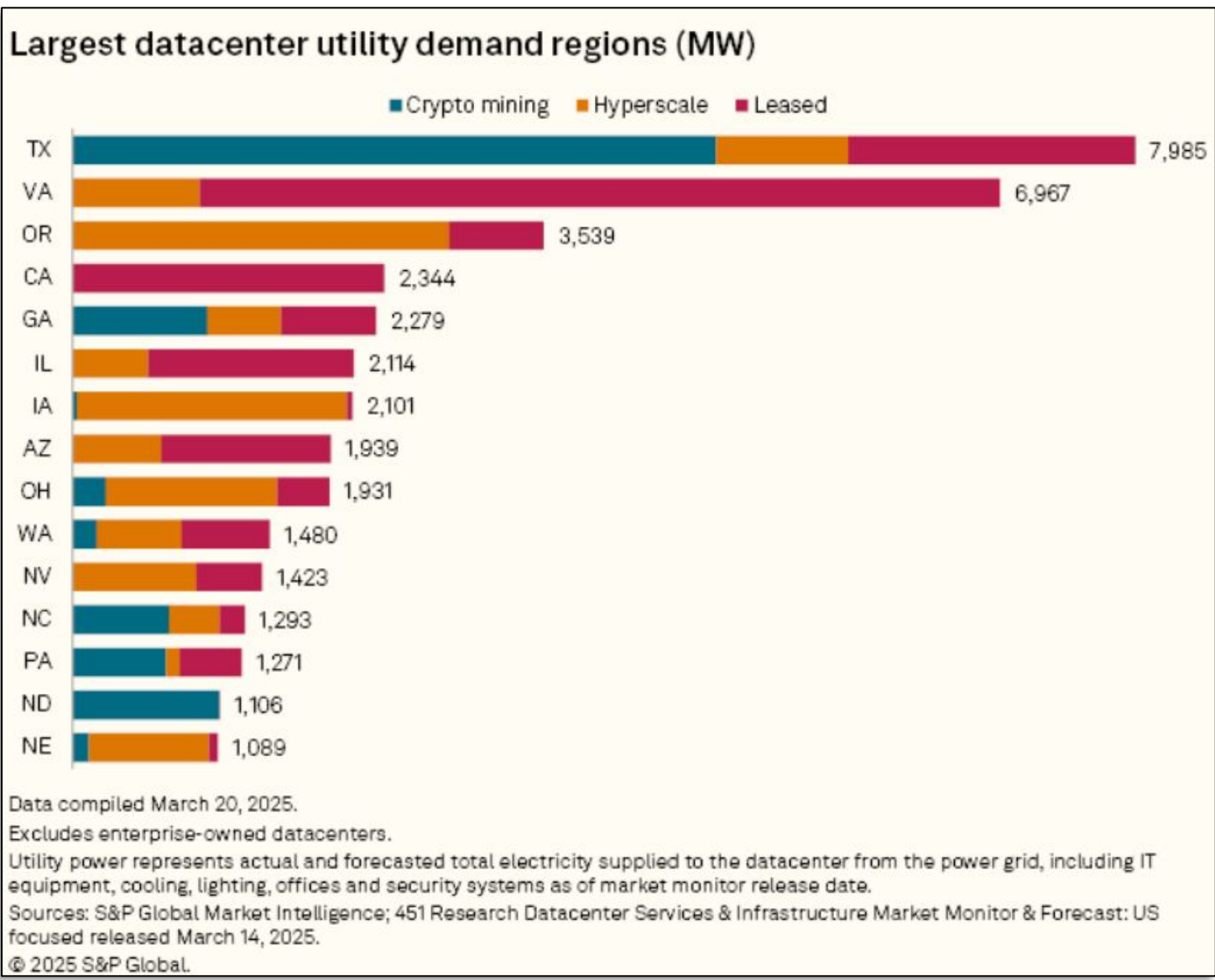
Location, location, location. “Data center” is a broad term, and the details matter.

Location & Size of Operating & Planned US Data Centers
 (Source: S&P Global Market Intelligence 451 Research)



While relatively dispersed, data center loads are also highly concentrated in certain states, like Texas, Virginia, Oregon, California, and Georgia.

Data Center Type for Largest US Data Center Regions
 (Source: S&P Global Market Intelligence 451 Research)



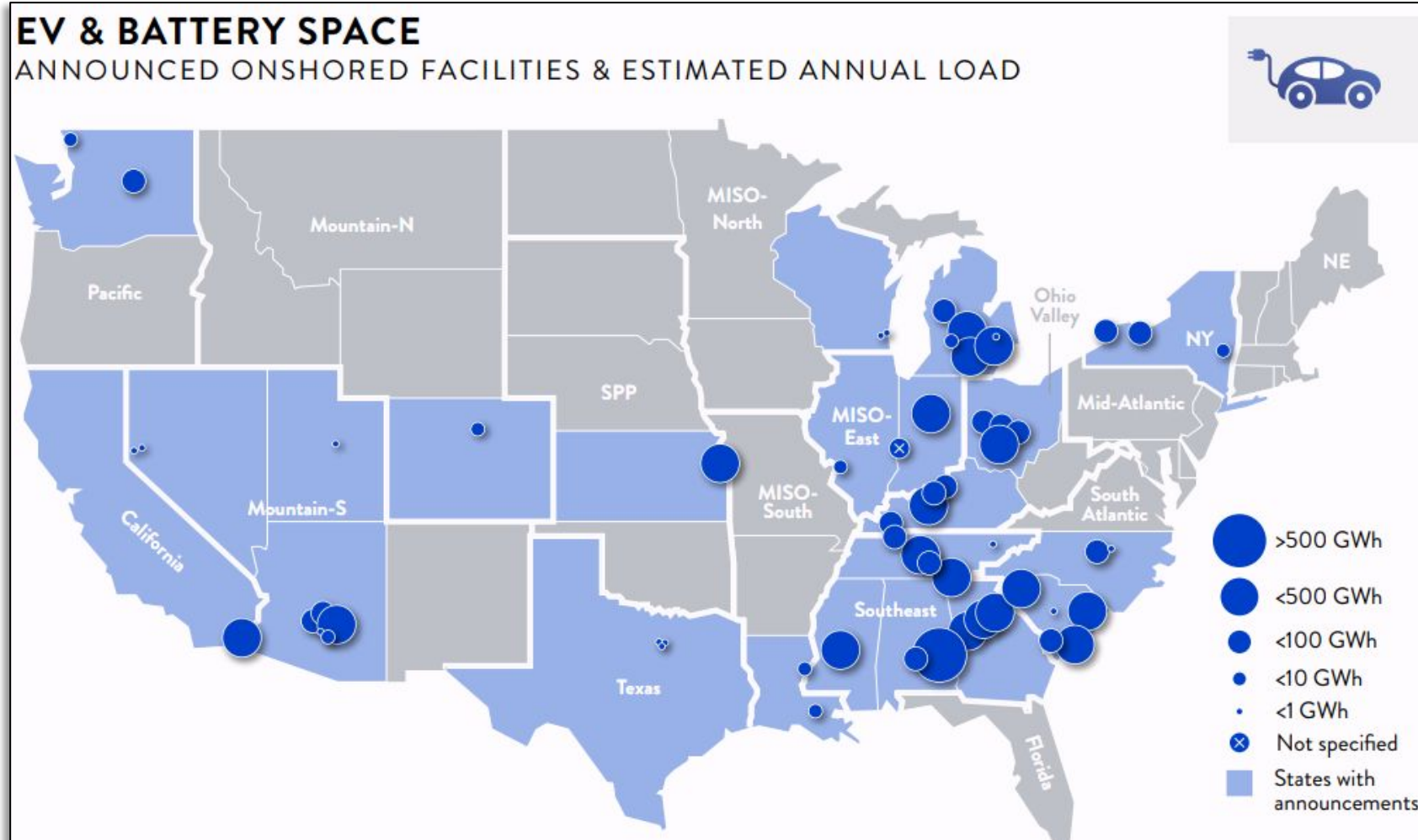
But the composition of those data center loads are diverse, which leads to different policies and forecasting needs.

Onshore Manufacturing Load Growth

EV/battery manufacturing is the largest sector of growth, but it is relatively modest compared to data center growth

Announced EV & Battery Facilities (as of 3/2023)

(Source: EPRI)



Peak Load Impact (MW) from New Onshored Manufacturing

(Source: Brattle, citing EPRI)

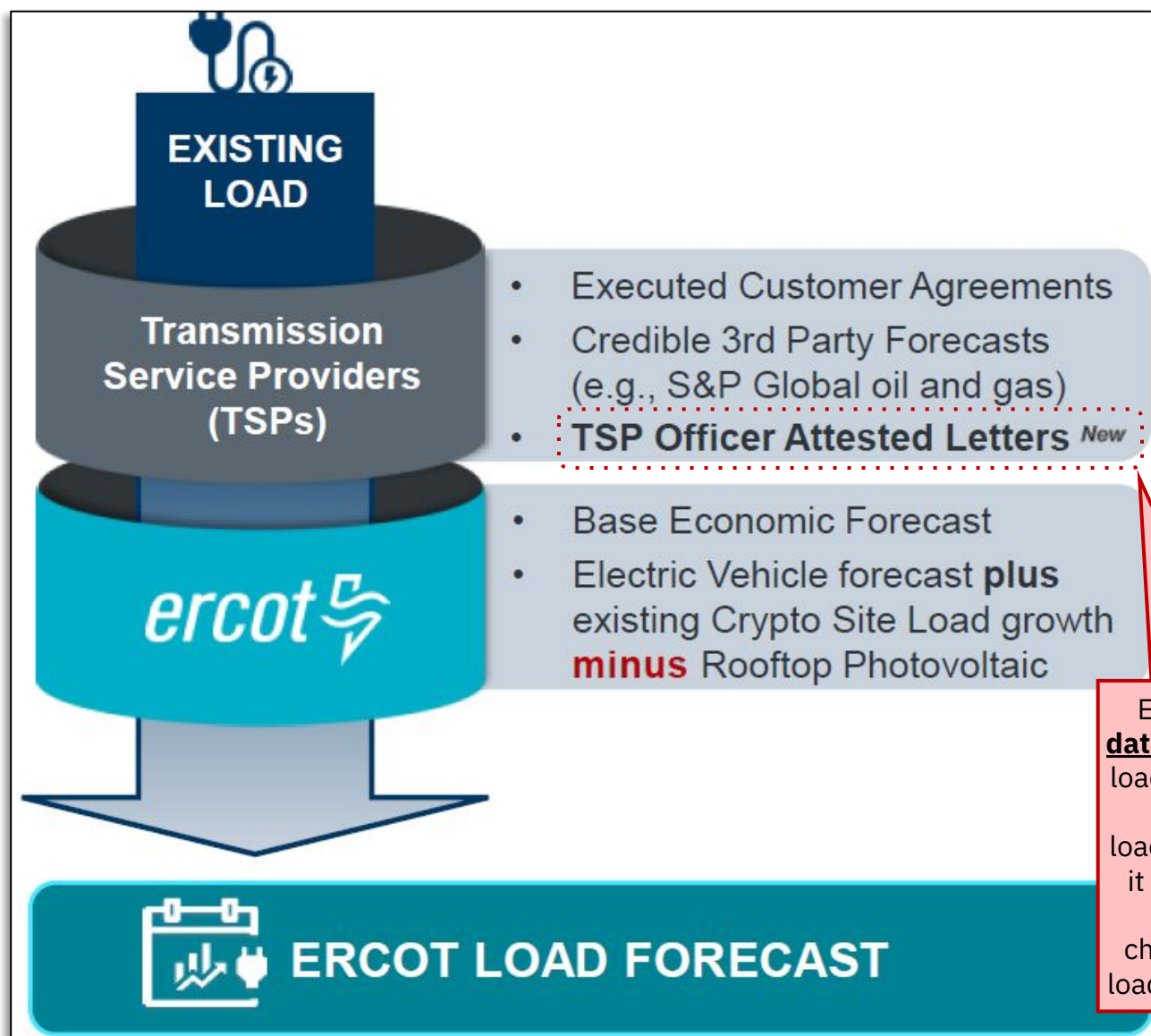
	EV/ Battery	Fuel / Plastic/ Chemical	Metals	Semiconductor/ Electronic	Solar	Wind	Transportation	Other	Total
Southeast	327.3	34.1	8.3	-	25.5	-	-	0.4	395.6
Mountain-South	30.3	2.0	-	170.9	1.6	-	-	-	204.8
Ohio Valley	98.6	4.1	78.7	-	3.5	-	0.0	0.0	185.0
MISO-East	127.5	20.0	-	-	-	-	0.0	-	147.5
South Atlantic	129.8	0.0	1.3	-	1.0	-	-	-	132.1
California	32.3	-	54.8	0.1	9.3	-	-	-	96.5
Mid-Atlantic	-	78.9	0.2	-	0.5	-	6.5	5.5	91.6
Florida	-	-	-	85.0	-	-	-	-	85.0
MISO-South	2.3	52.4	27.4	-	-	-	-	0.0	82.1
New York	12.5	-	-	0.4	0.4	46.8	-	-	60.1
Texas	0.1	52.1	-	-	1.5	-	-	-	53.7
SPP	22.8	-	-	-	-	-	0.1	-	22.9
Pacific	11.6	-	-	-	-	-	-	-	11.6
MISO-North	-	-	0.6	-	0.2	-	-	2.0	2.8
New England	-	-	0.0	-	-	-	-	2.0	2.0
Mountain-North	0.6	-	-	-	-	-	-	-	0.6
Total (MW)	795.7	243.6	171.3	256.4	43.5	46.8	6.6	10.0	1,573.9

ISO/RTO Data Center Load Forecasting Approaches

Conceptually, many ISOs/RTOs follow a similar approach – but the details matter!

ERCOT's HB5066 Load Forecasting Process

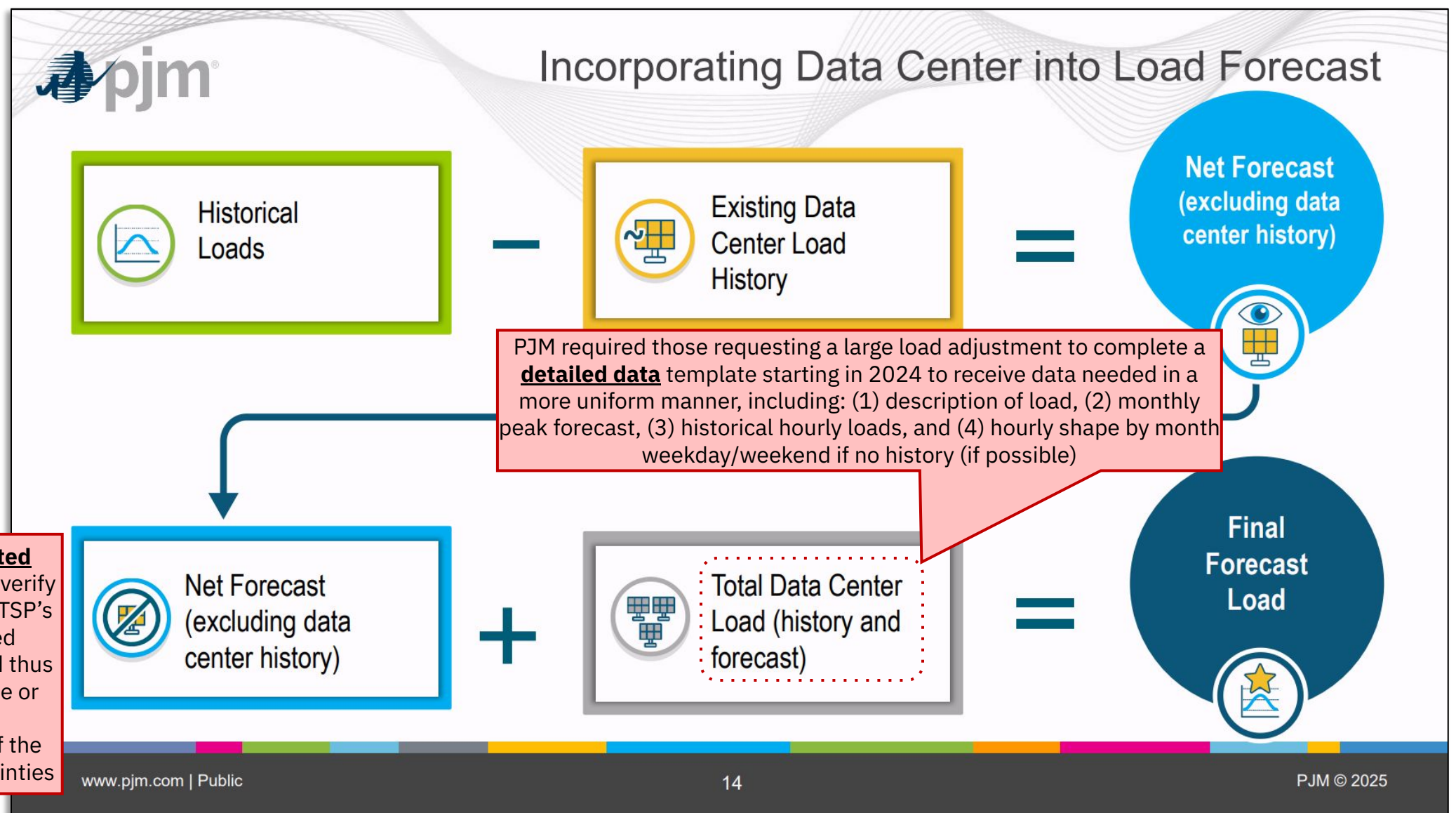
(Source: ESIG Large Loads Task Force presentation)



ERCOT has **limited data** to be able to verify loads provided in TSP's Officer Attested load category, and thus it cannot evaluate or validate the characteristics of the load or its uncertainties

PJM Process for Incorporating Data Center Load into the Load Forecast

(Source: ESIG Large Loads Task Force presentation)



ISO/RTO Data Center Load Forecasting Approaches

Several RTO/ISO approaches leverage some form of derating of data center loads

ERCOT
(Source: ESIG Large Loads Task Force presentation)

Forecast Methodologies

TSP-Provided Load Forecast

- All contracts and Officer Letter Loads based on the in-service dates and MWs that the TSPs provided

ERCOT Adjusted Load Forecast –

- 180 Day delay for all contract and Officer Letter Load
- All new Data Center Load reduced to 49.8% of request
- Then All Officer Letter Load reduced to 55.4% of request

Pre-2024 Load Forecast Method

- Contracts only on the ramp schedule provided by the TSPs

Key Takeaway: ERCOT will begin incorporating an adjusted Load Forecast in analysis that uses historic trends to adjust the TSP-Provided Load Forecast.

ercot

MISO
(Source: ESIG Large Loads Task Force presentation)

Publicly stated data center plans require moderation of forecast assumptions to reflect phased deployment and observed industry realization rates.

Project adjustments for data center forecasts

Provisional Load The total projected capacity of new data center projects as initially announced, either publicly or through interconnection requests.

Deployment Adjustment Reflects the extent to which stated capacities are deployed in practice.

Ramp Rates Projects are expected to phase-in over time, at an expected rate of approximately 20% annually.

Unrealized Load The portion of the adjusted provisional load that has not yet materialized on the grid, accounting for uncertainty in project timelines, execution, or actual usage behavior.

Metered Load The real, measured electricity consumption from operating data centers—used to validate assumptions and refine forecasting models.

MISO

PJM
(Source: ESIG Large Loads Task Force presentation)

PJM Process Summary

- Review the level of commitment:
 - Could include load related to in-service, under construction, with Electric Service Agreement (ESA), or with identified tenant arrangements
- Requests in terms of capacity ask are translated to demand
 - Used 70-90% of max capacity based on info from multiple submitters
- Ramp schedule
 - Data center projects should have ramp into final demand in 3-4 years or more
- Vendor comparison

www.pjm.com | Public PJM © 2025

CEC (CAISO uses the CEC's IEPR forecast, with IOU input)
(Source: CEC)

PG&E and SCE Methodology Updates

For PG&E and SCE, a confidence level was applied by application status:

- SCE T&D planning
- Group 1: completed or to-be completed engineering studies
- Group 2: Active applications prior to initiating engineering studies
- Group 3: Project inquiries

PG&E	Low	Mid	High
Group 1	50%	70%	70%
Group 2	-	-	50%
Group 3	-	-	10%

SCE	Low	Mid	High
T&D planning	100%	100%	100%
Group 1	50%	70%	70%
Group 2	-	50%	50%
Group 3	-	-	10% - 50% per SCE

Source: CEC Staff (with data inputs from PG&E, Palo Alto, SVP, San Jose, SCE) 4

State-of-the-Industry in Data Center Forecasting

Overall, there is “no consensus” among utilities; practices and forecasting methods vary widely

Trends Regarding Utility Data Center Load Forecasting

(Source: [EPRI](#))

General trends in load forecasting with DCs

- **No consensus** among 24 responding utilities* on how DC service requests are included in load forecasts
 - Ten utilities include the full requested capacity (as specified by the DC customer) but 8 of the 10 ramp the capacity over time
 - Another eight utilities include a derated capacity value based on specific weighting criteria (see next page)
 - Six utilities do not presently include DC requests in their load forecasts
- All face challenges around incorporating DCs into load forecasts given the **speculative nature** of some DC service requests

*One utility (out of the 25 total) did not respond to this specific question

© 2024 Electric Power Research Institute, Inc. All rights reserved.

EPRI

State-of-the-Industry in Data Center Forecasting

Overall, there is “no consensus” among utilities; practices and forecasting methods vary widely

How Some Utilities Derate Data Center Loads

(Source: [EPRI](#))

Examples of how five utilities derate requested data center capacity values for use in load forecasting:

- **Utility A:**
 - Derating depends on multiple factors, but DC type has largest impact. Also, generally derate colocation DCs more than enterprise.
 - Derates DCs more than other large loads due to less accurate information from DCs.
- **Utility B:**
 - Considers DCs in the forecast process similar to other large loads. But tends to have less information about DCs, which leads to more uncertainty.
 - For corporate load forecast, consider each DC individually but focus is on forecasting DCs in aggregate.
- **Utility C:**
 - Derates all DC requests by 20% based on prior experience. Will also vary the derating when forecasting multiple scenarios, for example, a 10% and 0% derating.
 - DCs are currently the only sector for which their load forecasts are based on connection requests. However, considering applying a similar derating process for electrolysis projects in the future.
- **Utility D:**
 - Derating based on multiple factors, including historical experience with other similar customers.
 - Currently treats DCs differently from other large loads for load forecasting. Within DCs, derating varies based on type of DC, MW size, etc.
 - Current experience has led to derating crypto mining customers more than hyperscalers given hyperscale DCs tend to more closely adhere to their contract loads.
 - Only considers DC customers in five-year load forecast who have signed an agreement with the utility.
- **Utility E:**
 - Derating depends on multiple factors, but biggest factors are (1) whether the data has submitted a formal connection request and (2) whether there is available infrastructure to support the DC's requested capacity.
 - Derating of DCs for load forecasting is more granular than other large loads given the diversity and uncertainty of DCs as a category.

None of the responding utilities provided an explicit calculation/equation for derating DC for load forecasting.

27

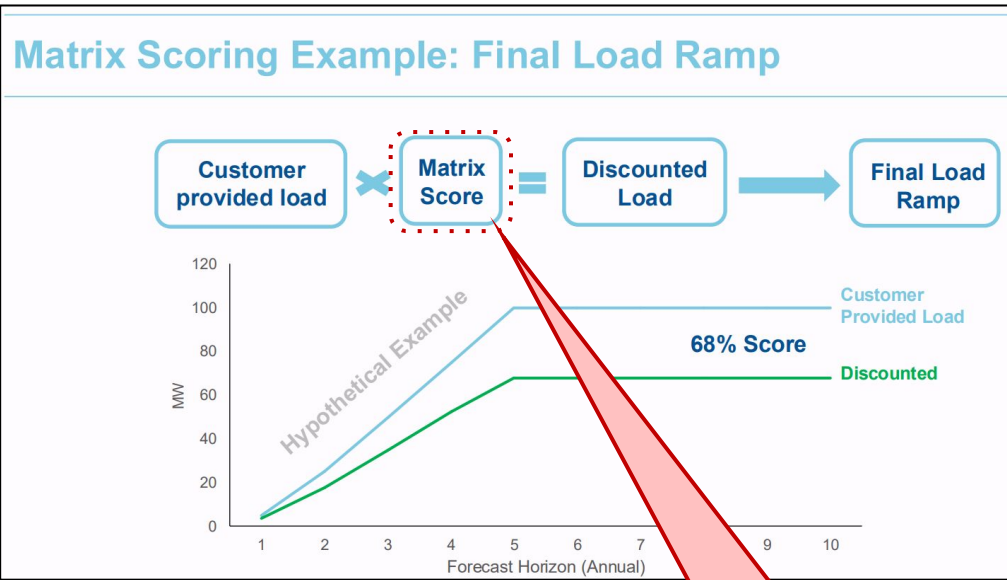
© 2024 Electric Power Research Institute, Inc. All rights reserved.

EPRI

State-of-the-Industry in Data Center Forecasting

Examples of more complex utility data center load modeling approaches designed to better address uncertainty

Southwest Utility Example (Source: ESIG Large Loads Task Force presentation)

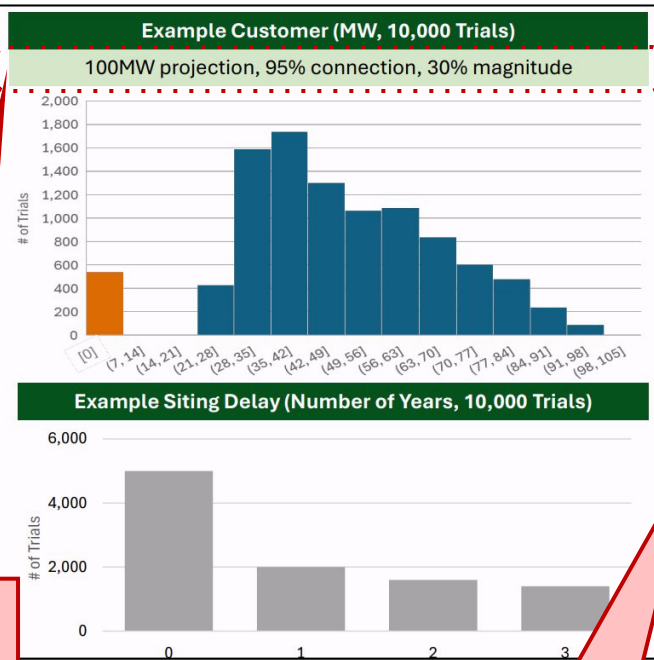


The "Matrix Score" concept is designed to reflect uncertainty across multiple dimensions, such as energy infrastructure, client construction, development permits, land acquisition, and other intelligence.

Southeast Utility Example (Source: ESIG Large Loads Task Force presentation)

- Likelihood of Connection**
 - Binary variable to determine if customer locates on system
 - Uniform distribution (randomized to match probability)
- Magnitude Adjustment**
 - Adjustment to reflect lower usage than projected
 - Triangle distribution (trend to internal projection)
- Siting Delay**
 - Potential delay in customer connection of 0-3 years
 - Left-skewed (no delay)
- 50,000 Trials**

The input parameters that feed the simulation are determined via a combination of data analysis and expert judgement.



Transparent inclusion of historical data compares the actual data against past forecasts.

Dominion Virginia Example (Source: DOM VA 2023 IPR Testimony and 2023 IPR)

Forecasting Methodology
The Company has been tracking data and preparing forecasts for a long period of time and has developed a very robust forecast methodology. Figure 4.1.5.1 compares the Company's forecast to actual data center demand for 2020-2022.

**Figure 4.1.5.1 – Data Center Industry Peak Billed Demand in MW
Company Service Territory**

Forecast Year	Forecast and Results		Variance	% of Variance
	Forecast	Actual	Over/(Under)	To Actual
2020	1,559	1,808	249	14%
2021	2,179	2,302	123	5%
2022*	2,848	2,767	(81)	-3%

* 2022 was the year of the transmission capacity constraint.

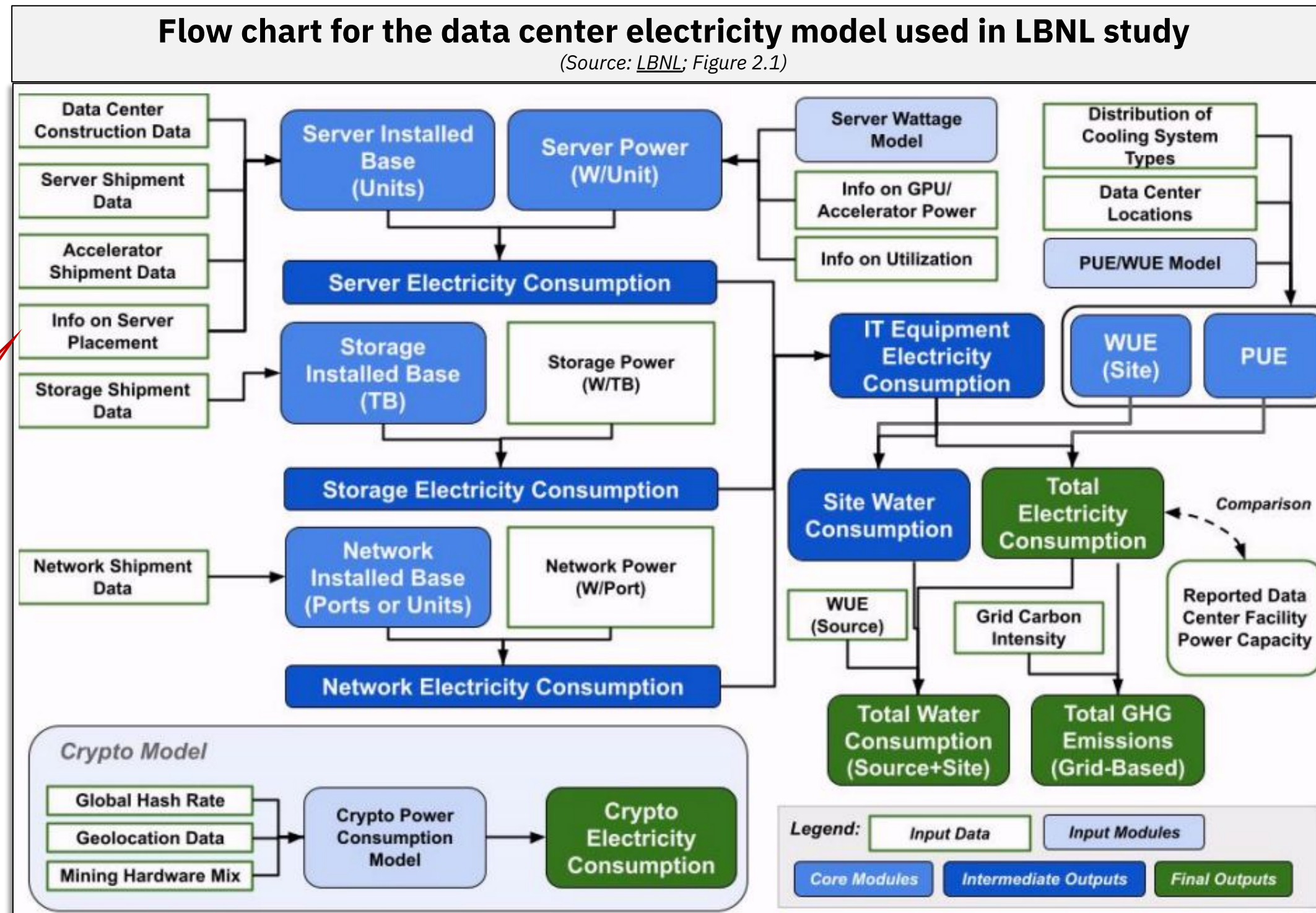
The Company models industry demand growth using the following method:

- Segments the modeling using the eight largest or fastest growing customers and a ninth model consisting of all remaining customers combined into one segment – nine models in total
- Statistically models sales in MWh including lost retail choice sales
- Statistically models demand (MW) using three different approaches
 - Approach 1: linear regression of demand
 - Approach 2: polynomial regression of demand
 - Approach 3: linear regression of sales to demand
- One of these three approaches is selected for each of the nine customer segments based on customer provided intelligence
- Estimate future retail choice conversions (lost MWh sales)
- Develop high, medium, and low demand scenarios
- In total, there are 27 models used to develop the forecast

Inclusion of confidential customer intelligence provides unique insights that can help calibrate a forecast.

Bottom-up Data Center Modeling: The Gold Standard?

Berkeley Lab's *Data Center Energy Usage Report* represents a robust *publicly* available data center load forecasting methodology. Is it a feasible model for others to emulate?



This model requires significant resources to develop and operate, including:

- Third party data purchases
- Multiple sub-models and simulation engines
- Access to industry experts

Questions States Can Ask

- What **data sources** did you use to develop your large load forecast?
 - How did you **collect the data**, and was data collection standardized?
 - How did your **data sources change** from previous years?
 - Are you incorporating **customer intelligence** to develop and adjust your load forecasts?
- What **assumptions** did you make in your modeling?
 - Did you **derate the load capacity** requests? If so, by how much, and what data did you use to inform your derating?
 - How did you incorporate the load's **ramp rate**?
 - If you **conducted simulations**, what assumptions formed the basis for the parameters you used?
- What **modeling methods** did you use?
 - Did you make any **changes to your methods this year**? If so, what impact did those have on your results?
 - How did you **benchmark** your approach against methods?
 - What steps did you take to **stress-test and validate** your results?
 - How do your results **compare against your past forecasts**?

Actions States Can Take

- **Public utility commissions** can:
 - Provide clear guidance to regulated utilities regarding the information and data sources they use to develop their large load forecasts
 - Consider implementing the use of data collection templates to standardize data collection
 - Require reporting on forecasting methods (including data inputs, modeling methods, and data outputs), changes to forecast methods over time, and forecasting error (compared to actuals), an explanation of the likely sources of difference, and what steps can be taken to reduce future error
- **State energy offices, utility consumer advocates and other stakeholders** can:
 - Collaborate with staff and officials in jurisdictions where large load forecasting practices are well established and validated to learn best practices
 - Build trust and understanding by setting up individual meetings with utilities to understand the details of their large load forecasting methodologies, the data sources they use, the improvements they've made over the years, the biggest challenges they face, and how they are planning to address these challenges

Disclaimer

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, or The Regents of the University of California.

Ernest Orlando Lawrence Berkeley National Laboratory is an equal opportunity employer.

Copyright Notice

This manuscript has been authored by an author at Lawrence Berkeley National Laboratory under Contract No. DE-AC02-05CH11231 with the U.S. Department of Energy. The U.S. Government retains, and the publisher, by accepting the article for publication, acknowledges, that the U.S. Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for U.S. Government purposes



Contact

Greg Mandelman: gmandelman@epeconsulting.com

Lisa Schwartz: lschwartz@lbl.gov

For more information

Download publications: <https://emp.lbl.gov/publications>

Sign up for our email list: <https://emp.lbl.gov/mailing-list>

Follow us on Bluesky: [@BerkeleyLabEMP.bsky.social](https://bsky.app/profile/BerkeleyLabEMP.bsky.social)

Acknowledgements

This work was funded by the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, or The Regents of the University of California.

