Economic Consulting Associates

Insights

Is there a data centre sized hole in flexibility forecasts?

23 September 2025

The AI boom and associated computing demands is fuelling unprecedented levels of investment in data centres. Alphabet, Amazon, Microsoft, and Meta are projecting US$350bn of data centre capex in 2025, rising to over US$500bn by 2030, an approximately fourfold increase from 2021-23.[1]

Such investment is set to substantially increase electricity load. In Great Britain, the National Energy System Operator's (NESO) latest Future Energy Scenarios (FES), published in July 2025, foresee data centre electricity consumption rising from 7.6 TWh today (currently driven by London-based data centres serving finance and trading needs) to 33 TWh by 2035, with scenarios in the range of 30-71 TWh by 2050.[2] This would be equivalent to adding 11-26% of present day electricity consumption. In the US, data centre electricity consumption could rise from 176 TWh in 2023 (4.4% of total US electricity consumption) to 325-580 TWh (6.7%-12.0%) by 2028.[3]

Such a major step-change in electricity demand is raising concerns (and BBC headlines[4]) about the impacts on grid and supply chain capacity, sustainability, and consumer bills. While these are valid concerns, standard assumptions about data centre electricity load also require further scrutiny, and recognition of an opportunity.

Is data centre electricity demand really that demanding?

As highlighted by researchers such as Tyler Norris, a common misconception (including among system planners and regulators) is that data centres operate 24/7/365 at near-maximum demand.[5] However, there is much heterogeneity within and across data centre types. Data centres may operate over 80% of the time when training AI models, but this drops to ~40% when utilised for AI inference. For hyperscale data centres providing mass scale internet services, this figure may be 50%, while multi-tenant, strategically co-located data centres may have operation rates of only 20-35%.[6] Average power draws are increasing, particularly for the most advanced 8 GPU configurations, but this trend is far less pronounced for less-advanced / traditional data servers.[7] UK Power Networks were recently 'surprised' that their existing data centre connections were only using 30-40% of their connected capacity.[8]

Data centres may insist on 24/7 availability for their customer needs, but this does not necessarily translate to 24/7 *actual utilisation*. A 2024 Lawrence Berkeley National Laboratory report lamented the general lack of transparency from data centres as to their actual expected workloads (often citing

---

[1] T. Kinder, 14 August 2025, 'Absolutely immense': the companies on the hook for the $3tn AI building boom, Financial Times.

[2] NESO, July 2025, Future Energy Scenarios: Pathways to Net Zero, V.4, Pages 110-111.

[3] A. Shehabi, et al, 2024, 2024 United States Data Center Energy Usage Report, Lawrence Berkeley National Laboratory, Report #: LBNL-2001637, Figure 5.5.

[4] Z. Kleinman and K. Shveda, 15 August 2025, Data centres to be expanded across UK as concerns mount, BBC News.

[5] T. H. Norris, 8 August 2025, The Puzzle of Low Data Center Utilization Rates, Substack: Power & Policy.

[6] A. Shehabi, et al, 2024, 2024 United States Data Center Energy Usage Report, Lawrence Berkeley National Laboratory, Report #: LBNL-2001637, Figure 3.6.

[7] Ibid, Figure 3.7.

[8] D. Blackman, 11 July 2025, UKPN director 'surprised' by level of data centre demand, UtilityWeek.

confidentiality concerns),[9] which is exacerbating grid planning uncertainty.

Traditional large loads, such as major factories, have a long history of closely collaborating with grid planners and regulators as to timing their peak loads, with flexible tariff designs and interruptible connection arrangements to match. Can such tools be adapted for the data centre era?

Harnessing data centre flexibility

Data centres' idle power draw may only be ~30% of their maximum or rated power.[10] Their load can vary by 10s or 100s of MWs in a matter of seconds. This pairing suggests a rapid demand response opportunity. To date, data centre demand response has been limited due to a reluctance to jeopardise performance benchmarks outlined in their service-level agreements and insufficient incentives from traditional demand response programs.[11] However, with data centres facing long connection queues and grid bottlenecks, this has renewed investigations of options for flexing data centre load and integrating this into system planning. Options include on-site generators (or batteries), shifting compute workloads to other locations, or temporarily reducing workloads, such as delaying non-urgent tasks.

In the US, Google recently announced contracts with two utilities that commit their data centres to providing flexibility as part of resource adequacy planning in exchange for quicker connections.[12] Google has also piloted shifting non-urgent compute tasks to other times or locations to limit grid stress.[13]

In Ireland, where data centre load already has a history of straining the grid, the regulator, CRU, and grid operator, EirGrid, effectively placed a moratorium on new data centres in 2021. The CRU is grappling with the issue in a 2025 proposed decision paper, including a proposal to require data centres to source dispatchable on-site or proximate generation and / or storage capacity, which must also participate in wholesale electricity markets in exchange for not being subject to mandatory demand curtailment provisions.[14] Irish system operators are also tasked with publicly documenting locational grid constraints to provide added clarity to potential connection applicants.

The CRU highlights mandatory participation in wholesale markets as a revenue opportunity. There could be parallels to bitcoin miners in Texas, who are reportedly earning more revenue from participating in the electricity balancing market than from actual mining.[15] Similar to how BESS technologies have revolutionised rapid response ancillary services, could massive computing load bring a new generation of rapid demand response offerings?

Data centres' high electricity consumption does not have to translate to higher consumer bills. If they can flex their load to reduce system peaks (and reduce network reinforcement needs), offer rapid (and lower cost) grid services, and help spread fixed costs over a higher volume of consumption, data centres could *lower* consumer bills. If grids fail to realise these opportunities, data centres, frustrated by long connection queues, appear ready to resort to off-grid solutions.[16]

Adding data centres to NESO's FES flexibility forecasts

In Great Britain, the National Energy System Operator (NESO)'s July 2025 Future Energy Scenarios (FESs) explicitly do not include data centre demand flexibility (nor locational optimisation) in its latest modelling due to present uncertainty.[17] Likewise, Ireland's 2023 Tomorrow's Energy Scenarios (TESs) simply modelled data centre demand as flat (while highlighting data centre

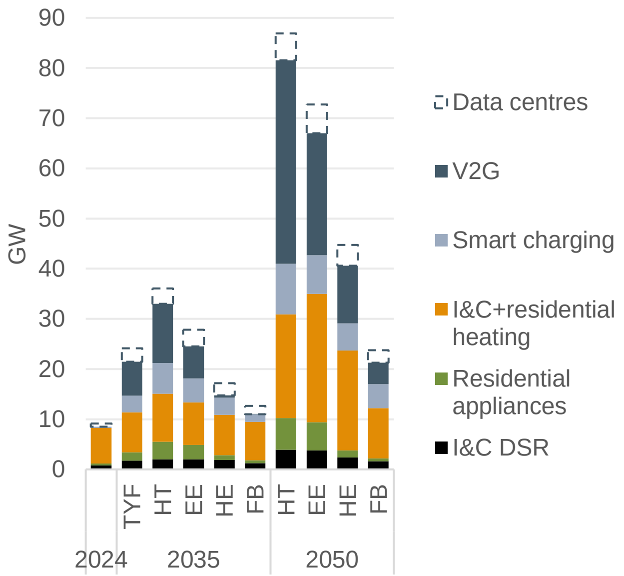
[9] A. Shehabi, et al, 2024, 2024 United States Data Center Energy Usage Report, Lawrence Berkeley National Laboratory, Report #: LBNL-2001637, Page 68.
[10] Ibid, Figure 3.5.
[11] T. H. Norris, et al, 2025, Rethinking Load Growth: Assessing the Potential for Integration of Large Flexible Loads in US Power Systems, NI R-25-01, Durham NC: Nicholas Institute for Energy, Environment & Sustainability, Duke University.
[12] T. H. Norris, 8 August 2025, Google's Big Flex: Can AI Data Centers Become Grid Stabilizers?, Substack: Power & Policy.
[13] Google, 4 October 2023, Supporting power grids with demand response at Google data centers.
[14] Commission for Regulation of Utilities, 18 February 2025, Large Energy Users connection policy, Proposed Decision Paper, CRU/202504.
[15] The Economist, 27 August 2024, Why Texas Republicans are souring on crypto.
[16] See reports of data centres exploring building on-site gas turbines, which could also jeopardise decarbonisation ambitions.
[17] NESO, July 2025, Future Energy Scenarios: Pathways to Net Zero, V.4, Page 111.

This ECA Insight was prepared by **Scott Edmonds**. For more information please email scott.edmonds@eca-uk.com

flexibility as a future objective).[18] As an illustrative exercise, we consider how NESO's forecasts could evolve if data centre flexibility emerges.

The FES's Ten Year Forecast suggests data centre peak load could reach 5.4 GW by 2035.[19] Its scenarios have a range of 4.9-11.5 GW by 2050. Assuming that 50% of this data centre peak load can flex (Figure 1), this could increase available flex capacity by 1.6-3.3 GW (or 9%-16%) by 2035 and 2.5-5.8 GW (or 7%-12%) by 2050.

Figure 1 Demand flexibility at peak, Great Britain Future Energy Scenarios



Source: NESO Future Energy Scenarios Workbook and ECA calculations. Assuming 50% of data centre peak load can flex. I&C = Industrial & Commercial. V2G = Vehicle-to-Grid. TYF = Ten Year Forecast. HT = Holistic Transition. EE = Electric Engagement. FB = Falling Behind.

While this back-of-the-envelope data centre load flex calculation is limited relative to the FES's other flex categories, it should be considered in context:

- The 2035 values would already translate to 20-40% of present flex capacity (81% of which reflects 'old school' storage heaters).
- They could *immediately* offer large scale load flexibility as soon as they are installed.
- In contrast, the FES scenarios foresee significant flex potential from smart charging and vehicle-to-grid but this is highly uncertain and subject to multiple potential obstacles: well-designed smart software and tariffs, aggregating distributed loads, potential impacts on vehicle battery life, consumer consent, distribution network bottlenecks, etc.

Some data centres will be unable (or unwilling) to offer flexible load, and will likely protest the imposition of mandatory demand flex for large loads. However, given cost socialisation risks, regulators and system planners need to develop their toolbox for scrutinising, and challenging, the demands of data centre connections. This can turn into an opportunity for a more resilient and cheaper grid through higher volumes, more granular network planning, flexible tariffs and connection charges, and more dynamic flexibility and ancillary services markets.

[18] EirGrid and SONI, 2023, Tomorrow's Energy Scenarios 2023, Final Report, Page 51.

[19] The FES assumes a fleetwide load factor of 70%, split between 80-85% load factors for data centres primarily dealing with AI and a lower load factor for telecoms and finance data centres: NESO, FES: Pathways Assumptions 2025 (Excel workbook download link).

This ECA Insight was prepared by **Scott Edmonds**. For more information please email scott.edmonds@eca-uk.com