November 2025

Data Centre Impact Study

An analysis of the growth and impact of data centres in National Grid Electricity Distribution's licence areas.

nationalgrid DSO

Contents

Purpose of this report

This report, undertaken by Regen and National Grid DSO, summarises research and modelling of the potential load growth and wider network impact of data centres in National Grid Electricity Distribution's (NGED) licence areas.

It summarises the potential future uptake of data centres, the geographical factors driving AI growth in particular areas, and commentary on the wider network and societal impacts of data centre development in the UK.

Prepared by: Tamsyn Lonsdale-Smith, Senior Analyst, Regen & Malachi Moses-Gair, DSO Engineer, National Grid DSO

Approved by: Ray Arrell, Associate Director, Regen & Oliver Spink, Head of System Planning, National Grid DSO.

1. Introduction

Background and policy context

Data centres have become an increasingly prevalent topic for the energy sector, policymakers and wider society. The development of data centres as commercial-scale premises has been increasing since the 1990s, in conjunction with the development of the internet and online services. In recent years, the UK has also seen the growth of larger colocation and hyperscale sites, serving a wider variety of services.

The scale and energy impact of individual data centres has increased significantly. Very large individual projects have started to appear in the past five to ten years, and with the increasing use of streaming services and the development of AI, the prospect of further growth and increases in project scales in the UK and worldwide is likely.

Data centre sites require a significant amount of electricity to power not only the IT equipment, but also the cooling and air handling equipment that goes alongside.[1] These energy requirements have caused data centre project developments to become a disruptive source of future electricity demand and a development sector that National Grid and other electricity network companies must monitor and explicitly include in their network planning process.

The government has recognised the importance of managing energy demand in the tech sector. In July 2025, the Technology and Energy Secretaries of the UK government met at the second AI Energy Council meeting to discuss the importance of AI energy demand.[2] In this meeting, it was agreed to forecast future trends in AI energy demand within the wider context of improving the grid connections process.

The planning system will also need to adapt. In September 2024, the Labour government reclassified data centre developments as Nationally Significant Infrastructure Projects (NSIP).[3] This reclassification recognises the development of these sites as a specific project class, aiming to streamline the planning process for larger projects. In addition, in December 2024, reforms to the National Planning Policy Framework included direct references to data centre development to facilitate the development of a modern economy and industry.[4]

More recent direct action by the government has sent clear signals to the industry about its support for data centre development. In July 2025, the Secretary of State overruled a decision on a 90 MW data centre development in Buckinghamshire that had been rejected previously.[5] This came just two months after a similar ruling by the Secretary of State to overturn the rejection of a 96 MW site in Hertfordshire. Both sites were initially rejected over green belt land impacts before the new rulings approved them for development.[6]

AI Growth Zones

In June 2025, the UK Government launched a scheme to accelerate AI-enabled data centres through a support mechanism for designated 'AI Growth Zones'.[7] The zones were initially announced in January 2025 as part of the UK government's AI Opportunities Action Plan, which sets out a national strategy for harnessing the benefits and growth of AI.

AI Growth Zones are open to applications from regional and local authorities in GB and industry representatives. Successful zones will benefit from streamlined spatial planning processes, enhanced

[1] CIBSE Journal, October 2019. Cost model: data centre cooling.
[2] UK Government, July 2025. Technology and Energy Secretaries chair second AI Energy Council meeting.
[3] UK Government, September 2024. Data centres to be given massive boost and protections from cyber criminals and IT blackouts
[4] UK Government, December 2024. National Planning Policy Framework
[5] Data Centre Dynamics, 2025a. Angela Rayner greenlights twice-rejected Buckinghamshire data center
[6] Data Centre Dynamics, 2025b. Angela Rayner overturns decision to block data center planned for Hertfordshire, UK
[7] UK Government, June 2025. AI Growth Zones open for applications

access to significant power supply connections (e.g. 500 MW or greater) and support for additional enabling infrastructure.

The UK and self-sufficiency

The UK leaving the European Union has also introduced limitations on cross-border data exchange, leading to security, technology and self-sufficiency concerns. Although the EU has granted the UK ‘data adequacy’ status, which allows cross-border data flows, this adequacy decision is subject to periodic review. It could be revoked if UK data protection laws diverge from EU standards. As a result, there has been increased momentum toward developing domestic cloud and AI data centres to strengthen self-sufficiency and ensure that critical data can be stored and processed within UK jurisdiction, reducing exposure to potential regulatory or political disruptions.

On the other hand, some industry representatives argue that complete self-sufficiency is neither realistic nor cost-effective, given the globalised nature of cloud infrastructure and AI development, and that maintaining close regulatory alignment with the EU may be a more pragmatic strategy to guarantee secure, long-term data flows.

Sector context – other research on the impact of data centres on electricity networks

At the time of writing this report, other projects and research around data centres and electricity networks are also being undertaken.

NESO is conducting a study concerning the impact of transmission-connected data centres on the high voltage network. As with the present study, NESO is primarily concerned with the types of data centres and load profiles that could affect system operability. In addition, there is interest in how certain business models, such as AI training, could be leveraged for demand-side flexibility.

2. Network impact – literature review

Regen has conducted a literature review to understand 1) wider data centre market trends, 2) the impact of data centres on energy networks, and 3) geographical considerations for future data centre development.

GB data centres primarily serve UK-based customers, but also provide international services. Major service providers such as Microsoft and Amazon Web Services (AWS) operate UK data centres to serve local users and ensure ‘data residency’. This also contributes to the digital infrastructure and connectivity transformation in the UK while paving the way for global operations and international clients.

With new technologies, new data centre archetypes and the increased demand for AI services, energy consumption trends have changed and may continue to evolve. Conventionally, data centres have had relatively stable loads on electricity networks because their services, such as cloud storage, were relatively constant. By contrast, servers with AI capabilities/services typically have many peaks and troughs, corresponding to levels of computation. As a result, the load impact on networks due to this variation will need to be understood and managed

Geographical considerations will determine where data centre deployment may be seen across NGED’s licence areas. Apart from seeking grid connection capacity and suitable land, data centre developers consider spatial factors such as low latency (proximity to the end user) and the presence of fibre optic and data network infrastructure from existing data centre providers.

UK data centres market

Types of data services

At a high level, data centres are grouped networked computer servers and associated IT equipment, providing various data services. Crucially, the type, or business model, of a data centre does not necessarily dictate the types of services used and delivered by its servers’ occupants. Some examples of data services are outlined below:

Cloud storage refers to the on-demand delivery of IT resources over the internet through providers such as AWS or Microsoft, collectively known as ‘the cloud’. It is a well-established method for digital data storage and has experienced exponential growth over the past decade.[8] Data storage operation typically has a flat electricity demand profile.

Cloud computing refers to computation via remote cloud-hosted services. Examples include platform as a service (PaaS) or serverless computing. This can have more variable demand use throughout the day.

AI inference encompasses technologies capable of performing tasks such as processing large volumes of data to identify patterns or make decisions. UK private capital investment in AI rose from £0.9bn in 2018 to £3bn in 2022,[9] and the UK government estimates that the sector attracted around £200 million of private investment a day in 2025.[10] As AI becomes more integrated into everyday life, the infrastructure supporting it will increasingly need to be located closer to the general population.
The nature of AI ‘prompting’ and the wider range of users, use cases and requirements of AI services across industry and society will likely cause data centre energy usage to be more varied and ‘peakier’ than other services that data centres have historically provided.

[8] Edge Delta, May 2024. 6 Vital Data Storage Statistics You Must Know
[9] City of London Corporation and EY, Nov 2023, AI: Accelerating Innovation
[10] UK Government, Jan 2025, Press release: UK AI sector investment

Types of data centres

Several asset types exist in the development of data centre projects, each with distinct demand and business model characteristics. Data centres can be categorised in different ways, and some may overlap. However, some common archetypes exist across the sector:

Hyperscale data centres, often owned by big technology companies, might host in the order of 5,000 servers and around 10,000 square feet of floor space and require demand capacity connections into the 100s of MW. These could host anything from cloud services to storage, and increasingly AI servers. The proportion of servers dedicated to AI in this type of business model is increasing and, along with it, the variability of peak power load.[11]

Of the hyperscale business model, there are two broad subcategories. The first is the AI factory, which represents most of the largest sites entering the market today. These are 100% AI-powered and could be used for training AI models or for AI inference.[12] The second are more traditional cloud hyperscale sites. These typically see a more equal split of 50% cloud and 50% AI server racks.

When a hyperscale data centre is wholesale, its entire facility is dedicated to one customer. In contrast, some hyperscale data centres lease out servers to multiple clients; these can be referred to as hyperscale colocation sites.

Colocation data centres are third parties that lease servers to tenants. They provide all components required for companies that don't have the space or IT capabilities to host themselves. Importantly, colocation models can be hyperscale or small buildings. Some active colocation data centre providers in the UK are Equinix, Digital Reality and NLighten.

According to an analysis of 850 data centre survey respondents by Uptime Institute, colocation business models are set to grow more than any other business model type out to 2026.

Edge data centres[13] are also emerging as a key business model archetype. They are smaller than hyperscale sites and designed for AI inference services with low latency enabled by proximity to end users. As services such as autonomous vehicles, high-demand streaming and Internet of Things (IoT) become more in demand, the demand for edge data centres could increase

Enterprise data centres are smaller in scale and tend to be owned by businesses themselves. They may either be located on site or at a location nearby to the business, so as to maintain optimal latency and security, and typically have no or very limited AI-enabled server capacity.

Server types

The server type, as a core technology underpinning a data centre site's operations, is a key consideration for the power demand, energy consumption over time and wider network impact of a data centre.
Each server type has unique energy load requirements across a given operational day, with some being more predictable than others. AI servers have the highest energy density (50-120 kW/rack), followed by cloud computing and IT services (30-85 kW/rack) (see Table 4).[21]

Crucially, servers used for AI services are less likely than traditional cloud and IT servers to have predictable load profiles, due to the varied computational requirements as the volume and intensity of end user queries vary. Assumptions can be drawn on the average proportion of

---

[11] Data Centre Frontier 2022, Understanding the Differences Between 5 Common Types of Data Centres
[12] AI inference refers to the utilisation of an already trained AI model by users.
[13] Note: Edge data centres have been modelled under the same modelling archetype in NGED DFES 2025 as Collocation data centres, since many Edge centres follow a similar business model on multiple occupancy.

AI-enabled server capacity for major data centre archetypes, although individual data centres will deviate from the average.

Hyperscale data centres are set to have increasing proportions of AI servers, from 40% of the data centre's server capacity demand in 2023
to 65% in 2030.[14] Colocation data centres traditionally have lower proportions of AI servers than hyperscale data centres (c. 20%).

Still, they could have higher AI rack density rates if tenants have high computing or AI-service needs, for example in edge business models. Enterprise data centres had a smaller proportion of AI servers (10%) in 2023, but this could increase by 2030 as AI use for everyday business operations becomes more widespread. A summary of data centre types and typical characteristics is shown in Table 1.

Table 1: Characteristics of data centre archetypes

Type	Description	Typical demand profile	Typical % of demand used by AI servers	Typical size (MW)
Hyperscale–cloud, storage and AI	Large, more than half operated by Google, Amazon or Microsoft	Consistent heavy baseline, with increasing variability depending on % of AI servers	40% in 2023, 65% by 2030[14]	>20
Enterprise	Medium, owned by businesses that need large IT or data services	Variable with peaks during work hours	10% in 2023, 30% by 2028	<20
Colocation	Leased spaces with multiple tenants	Fluctuating and lagging cooling response	20% in 2023, 40% by 2030	Variable, sometimes hyperscale, sometimes 1-2 MW
Edge	Small, modular and micro-scale data centres located close to end users	Variable and associated with end user needs	Likely high, 50%	<20, potentially smaller

Other, more bespoke data centre archetypes may also exist. While the above archetypes are relatively well known in the sector, alternative approaches to categorising data centres could evolve as the market grows and end user segmentation widens.[15] Business models and project developments are evolving and growing. As has happened in the battery storage sector, asset classes, business models and archetypes for data centre projects may continue to change further.

[14] McKinsey 224, AI power: Expanding data center capacity to meet growing demand
[15] Cisco Systems, 2025, What is a data center?

Cooling equipment

A key driver of power demand is the type of cooling system used. With the increased efficiency of AI racks, the demand for cooling also increases. Different rack densities and cooling systems are further discussed in the Energy efficiency section of this report.

Active data centre developers in the UK

Regen has conducted high-level market research for this assessment to map some of the active UK data centre developers. The research is not comprehensive, but it provides an overview of some active organisations and the frequency of acquisitions. A summary of the main UK data centre market actors is detailed in Table 2.

Table 2: Data centre developers active in the UK market

Developer name	Data centre business models
Developer name	Data centre business models
Ark Data Centres	Hyperscale (wholesale or colocation)
STACK Infrastructure	Hyperscale/AI-ready
NTT – Global Data Centers UK. (Former Gyron Internet Ltd)	Hyperscale/colocation
Kao Data	Hyperscale/HPC[16], independent
Virtus Data Centres	Hyperscale/colocation/AI/HPC, independent
Yondr Group	Hyperscale
Global Switch	Hyperscale/colocation
Vantage Data Centres (acquired Next Generation Data)	Hyperscale/colocation
CyrusOne (acquired Zenium in 2018)	Hyperscale (wholesale or colocation)
Equinix (Acquired Telecity)	Hyperscale/colocation, carrier-neutral
Digital Reality (Acquired Interxion in 2020)	Hyperscale/colocation/edge
Telehouse (KDDI)	Colocation, carrier-neutral
nLighten (UK) (Formerly Edge Data Centres, acquired Proximity in 2023 and EXA in 2024)	Edge colocation
Space Data Centres	Regional colocation, independent

[16] HPC: High-performance Computing. Includes AI & Machine learning, scientific simulations, engineering analysis, financial modelling, rendering, etc.

Pulsant (Acquired Lumison, ScoLocate, Onyx, LayerV)	Edge colocation
UKFast / ANS (Merged 2022 under Inflexion)	Enterprise colocation
DC Vantage (Acquired Next Generation Data in 2020)	Hyperscale conversion

Market and business model trends

The evolution of the data centre development pipeline from 2017 to 2023/24 shows that a larger share of data centres in the future will consist of hyperscale business models.

In 2017, small on-premises enterprise data centres comprised 60% of the market, and hyperscale data centres only accounted for 20%.32 By 2023, hyperscale data centres represented 40% of the market, and could reach 60% of the market by 2027.32

There has been a drive for cloud service providers to develop more facilities for businesses that wish to outsource their IT infrastructure. This drive is partly due to Brexit, as the UK sector seeks to become more self-sustaining.[17]

Future data requirements and network connectivity

There is an increase in the adoption of 'smart' household appliances in the UK with the ability to connect to the cloud. The Internet of Things (IoT) model also encompasses connected devices that generate and exchange vast volumes of data, from household appliances like kettles to more complex systems such as autonomous vehicles. In the future, there is the potential for self-driving vehicles to upload around 3 TB of data per day, requiring rapid processing through 5G-enabled edge hubs before transmission to regional or hyperscale data centres.[17]

Low-latency requirements mean these facilities need to be located close to end users. Fibre optic cables will be supported by high-capacity fibre routes within motorway corridors in the UK. Fibre network highways extend from London and Slough through to Bristol and Cardiff via the Great Western Rail Line and Grand Union Canal.

Rising demand for high-resolution content, including the shift from 4K to 8K video, will also potentially overwhelm existing data network infrastructure, driving the need for more localised processing and data storage capacity. This will lead to significant investment in regional data centre infrastructure, particularly in areas with robust fibre connectivity. Both the growth of IoT and the increasing scale and quality of digital content are expected to place substantial new demands on UK data centres and, by extension, on the electricity systems that power them.

Electricity network connection considerations

Another factor for DNOs to consider is the likelihood that data centres may target distribution connections instead of transmission connections. Several factors, including import capacity requirements, proximity to transmission network substations, and the data centre archetype and business model, could influence this.

Hyperscale models will likely aim to connect to both the distribution and transmission networks, but site size will ultimately determine whether a transmission connection is required. While there are exceptions, sites with 250 MW import capacity or higher have not tended to apply for a distribution connection to date

[17] Knight Frank, 2024. Data Centre Development Report

(relating to distribution high voltage tiers capping at 132kV). This indicates that larger sites may be targeting the transmission network. Sites below 250 MW could connect to either distribution or transmission, but the proximity to infrastructure will likely determine whether a distribution or transmission connection is pursued.

In general, a very large demand connected to the transmission network is a different consideration from generation. Under The Electricity Act, Section 6, the government can grant licences to generation, transmission and distribution assets, but not for large demand loads such as data centres. This poses a barrier to data centres owning and operating their own transmission voltage equipment. Demand that connects to transmission will be required to physically energise to a 132kV connection bay or new substation at a 400kV grid supply point (GSP).

With the scale of data centre prospective development seen by network operators, NGET and the UK government could explore options around these existing licensing restrictions. A data centre could, for example, obtain a licence exemption through The Electricity (Class Exemptions from the Requirement for a Licence) Order 2001. Alternatively, the government could create a specific demand licence as a follow-on to enhanced infrastructure support through the AI Growth Zone process. The nature of licensing and connecting very large sites to the transmission system could be something that Ofgem and DESNZ explore further, including as part of the RIIO-3 business planning process.

Regarding a higher-level estimate for the split between data centre sites targeting the transmission or distribution network, the NESO FES modelling has previously assumed c.40% of data centre capacity will be distribution-connected. As more edge and other modular data centre business models emerge, we may see higher levels of future capacity manifesting as smaller sites with a higher likelihood of being connected at distribution voltage levels. Equally, the volume of much larger hyperscale sites seeking transmission connections could dwarf these figures. Overall, the potential for growth in distribution-connected centres rivalling future hyperscale capacity in the longer term is unclear.

Today, the pipeline of prospective data centres and the network tiers they target across transmission and distribution in GB is unknown. This is due to there being no equivalent project pipeline register(s) for demand as there are for generation (e.g. DNO Embedded Capacity Registers and NESO's Transmission Entry Capacity Register). However, discussions with organisations close to multiple prospective data centre projects suggest that the pipeline of data centres across GB could be into the 10s of GWs.

A very high volume of prospective data centre connections may trigger the need to look at a similar, perhaps more nuanced, queue management intervention to the NESO connections reform process, which has been applied to generation and battery storage. Such a process will have knock-on impacts on data centre developers targeting transmission or distribution network connections.

Planning permission process

Historically, data centres have been considered by local planning authorities. Planning requirements for successful data centre applications can include: pre-development agreement from water companies, power connections, district heating network feasibility studies, local employment obligations, biodiversity net gain and public benefits packages.

Moving forwards, large enough data centres will be considered as Critical National Infrastructure, and will therefore be subject to Development Consent Order (DCO) approval. The threshold of this is still being determined. The previous version of the National Planning Policy Framework (NPPF) did not mention data centres when outlining how planning policies should support local and inward investment. However, the December 2024 update to the NPPF requires planning authorities to identify suitable data centre locations when developing local plans (paragraph 86(c)). Additionally, paragraph 87 emphasises that local planning authorities should actively support proposals for new or improved data centre facilities and related infrastructure needed for economic growth.

For data centre projects falling under the NSIP regime, the Ministry of Housing, Communities and Local Government (MHCLG) is currently consulting on options to streamline infrastructure planning processes and policy.[18] While no explicit proposals for data centre projects have been specifically included in this consultation, the wider potential for the NSIP process to be streamlined may impact the planning assessment timeframes for larger data centres.

This NSIP consultation may also relate to the proposed support and fast-tracking approaches for projects seeking to be developed in the UK government's AI Growth Zones.

The more data centres seek to connect, the more specific guidance and processes will be issued around how they are to be assessed by both local planning authorities and the NSIP regime (and devolved equivalent processes in Wales and Scotland).

Data centre energy demand

Data centre energy demand has traditionally been relatively level. However, this trend is changing, where data centres hold higher proportions of AI-ready and high computational units.

Current and future energy demand

Current demand

Data centre growth in the UK has been exponential in recent years, with the UK reaching the top third of countries in the world for data centre development and ranking second in Europe, behind Germany. As of March 2025 the UK was host to approximately 523 data centres.[19]

According to NESO, in 2024 data centres made up 2.4 GW of capacity from grid-connected facilities and consumed an estimated 7.6 TWh.[20] This is equivalent to 2% of GB electricity demand. Many are concentrated in areas with high-speed internet connectivity in London and Slough.

Future demand in the UK

Several studies have looked at the future of data centre electricity demand in the UK. According to Aurora Energy Research, data centres already make up approximately 4% of GB electricity demand, and are projected to increase their share to 11% of all demand by 2035.[21]

NESO's Future Energy Scenarios (FES) 2025 also model a major increase from 7.6 TWh in 2024 to between 20 and 41 TWh by 2035.
Several organisations are looking at what proportion of energy demand will be specifically AI-driven. The UK government's Compute Roadmap (developed in collaboration with UKRI) aims for 6 GW of AI-enabled capacity by 2030.

Future demand in the EU and globally

[18] UK Government (Ministry of Housing, Communities and Local Government), September 2025. Consultation on streamlining infrastructure planning.
[19] Statista, 2025. Leading countries by number of data centers as of March 2025.v

[20] NESO, July 2025. Future Energy Scenarios: Pathways to net zero
[21] Aurora Energy Research, 2025. Impact of Datacentres on the GB Power System

Other studies have focused on data centre energy demand at the European or global level. The International Energy Agency (IEA)-4E has compared a wide range of estimates published since 2014, showing how uncertain these projections can be.[22] Early work by Andrae & Elder (2015) suggested demand could reach as high as 8,000 TWh. However, a later study by Andrae (2020) placed the highest scenario at 974 TWh by 2030.

Most projections, however, fall below 2,000 TWh by 2030, and none of the 50 modelling studies reviewed by the IEA-4E since 2018 exceed 2,100 TWh. In 2024, Independent Commodity Inteligence Services (ICIS) calculated that 3.1% (96 TWh) of European demand would be from data centres.[23] By 2035, data centre electricity consumption could rise to 5.7% of European demand. The rate of growth was seen to be 10% year on year. The report also found that Ireland's data centre demand represented 18% of national electricity demand, showing that extreme cases are also possible where low corporate tax rates can attract key global players.

The IEA has modelled global data centre energy demand scenarios, with AI making up 3% of energy consumption by 2030 in the Base Case. [24] Under its highest scenario, AI demand makes up 4.4% of global electricity demand. Another IEA study estimates that 35-50% of energy use from data centres by 2030 could be from AI use, compared to 5-15% today.[22]

Claims of AI's net energy efficiency benefits

Several studies have investigated ways in which AI-driven data centres, and digitalisation more generally, could reduce energy demand. The IEA has considered how AI could bring about efficiency and operational gains in the energy sector, for example, by reducing outage durations by 30-50%.[25] Other claimed benefits included unlocking up to 175 GW of transmission capacity without upgrades through the use of remote sensors and AI-based management.

European Economics concluded that, in the case of both AI-translation and e-books, AI and digital solutions can have significantly lower energy usage than physical counterparts (e.g. printing books or using a human translator).[26] Video streaming, however, remained as energy-intensive as lifecycle emissions associated with Blu-ray discs.

Some voices are more sceptical about AI and data centres' potential benefits. A paper by The Minderoo Centre for Technology and Democracy (2025) cautions that data centres could emit more than self-reported emissions figures imply due to a lack of detailed reporting of AI-specific carbon standards.[27] It further warns against claims that AI efficiency gains could bring net benefits for climate solutions if misuse leads to AI-driven profit seeking.

Additionally, an already-constrained grid capacity will struggle to accommodate new large demands while also maintaining housing targets. Similar cases such as these are already being seen in west London.[28]

Load profiles

Traditional cloud or enterprise data centres

Cloud and enterprise data centre demand profiles are relatively flat,
with limited peaks; 95% of end-user demand in London is from cloud customers with interconnectivity

[22] IEA-4E, 2025. Data Centre Energy Use: Critical Review of Models and Results
[23] ICIS, 2024. Data centres: Hungry for power
[24] IEA, n.d.. Energy Demand from AI.
[25] IEA, 2025. Energy and AI.
[26] Europe Economics, 2025. Impact of Growth of Data Centres on Energy Consumption
[27] The Minderoo Centre for Technology and Democracy, 2025. Big Tech's climate performance and policy implications for the UK
[28] BBC, 2022. New homes may be delayed by power grid capacity.

requirements in pre-defined Availability Zones.[17] Since cloud services are the main service to date, energy demand on the networks has remained less variable. As AI demand increases, however, there could be a shift in the typical demand profile and network loading times for future data centre connections, as well as for existing sites experiencing retrofits and IT equipment overhauls.

Traditional data centres require uninterrupted power supply (UPS) to ensure users always have access to their services. The degree of UPS performance can be summarised using the tier system summarised in Table 3.

Table 3: Traditional data centres tier classification system

Tier	Description
I	Basic capacity required for an office or individual business. Requires UPS, dedicated cooling equipment and backup generation.
II	Covers greater redundancy requirements for power and cooling to safeguard against power disruptions.
III	Does not require a system shutdown for maintenance, unlike Tier I and Tier II.
IV	Has isolated, independent systems to prevent disruption events from disrupting multiple systems. All equipment is fitted with a fault-tolerant power design and requires continuous cooling.

Source: Uptime Institute

AI training data centres

Energy demand forecasting is challenging for AI data centres, and inference demand looks different compared to AI training. Inference will have more peaks throughout the working day, while AI training load curves are characteristically flat, with quick ramp-up or ramp-down times requiring fast responses from the networks. According to real demand profile data from NERC, ramp periods can be as quick as 1.9 in 250 milliseconds.[29]

The figure below shows a training model with rapid fluctuations in energy demand during periods of training (c. 0.6 p.u.) as well as during checkpoint saving (c. 0.17 p.u.). Needless to say, ramp-up and ramp-down times between these two states of AI training are very fast.

[29] North American Electric Reliability Corporation (NERC) 2025, Characteristics and Risks of Emerging Large Loads.

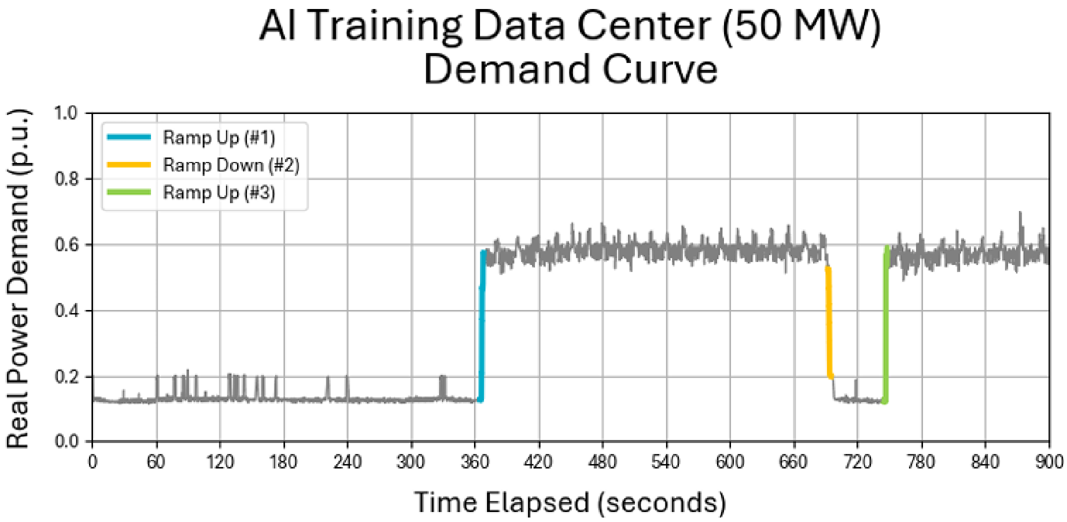

Figure 1: AI Training Centre load profile example

Source: North American Electric Reliability Corporation (NERC) 2025[29]

It is estimated that between 80 and 90% of all energy use from AI computing is for inference, while the remaining 10-20% is for AI model training.[30]

AI inference data centres
AI Inference load profiles lack the characteristic rapid ramp-up and ramp-down of AI training centres. The use of AI inference for day-to-day applications of business or personal customers reveals more variable and peaky profiles, making hourly load profiles more difficult to predict.
It is not clear whether the majority of AI data centres used in the UK will be for training or inference.

[30] MIT Technology Review, 2025. We did the math on AI’s energy footprint. Here’s the story you haven’t heard.

Energy efficiency

Power Usage Effectiveness (PUE)[31] is a key metric used in the industry, defined as the ratio of total facility energy consumption to the energy consumed by the IT equipment alone. The closer a PUE is to 1.0, the more efficiently the facility is operating.

$$PUE = \frac{total\ facility\ power}{IT\ system\ power}$$

Energy demand in traditional data centres (e.g. since 2007) has been steadily declining due to significant efficiency improvements. Power usage effectiveness (PUE) has improved, from an average of 2.5 in 2007 to below 1.6 by 2024.[32] According to Uptime Intelligence's data centre survey of over 850 developers, PUE ratings have remained stagnant over the past five years, with average ratings at 1.56 in 2024, while recent builds achieve closer to 1.3.[33]

Leading hyperscale data centres in the UK and globally can achieve PUEs near 1.1. Google has reported a PUE of 1.09 across its large-scale sites in 2025.[34] This is thanks to strategies like hot/cold aisle containment, efficient cooling systems, and AI-driven load and thermal management. There is likely to be diminishing returns on this as energy efficiency and operational optimisation causes PUE to get closer to 1.

PUE is just one of many different measures of data centre energy efficiency. At least 17 different metrics have been developed, each with its own merits, some preferring simplicity while others are more involved, requiring more robust data and calculations.[35]

Rapid technology improvements

Another aspect of efficiency is the efficiency of AI and computing technology over time. In today's data centre industry, technological advancements are tangible and quantifiable. For instance, hyperscale developers and innovators are reporting annual energy efficiency gains of 1.34 in computational performance per watt across AI supercomputers between 2019 and 2025.[36]

Processor and chip efficiency are improving notably. Different types of data centre racks have different energy densities. NVIDIA chips themselves have been improving in computing capacity over time.

A single high-density AI rack can consume 240 times the power of a small enterprise rack, as illustrated in Table 4. One study projects 8-15% of yearly gains in server power efficiency for AI workloads.[37] However, improvements in microchip speeds, capacity and performance is likely to slow and eventually flatten over time as the limitations of physical spacing and size is reached.

[31] Note: Although it has become a widespread industry standard, PUE has some downsides. In particular, this metroc doesn't consider efficiency gains from the IT equipment itself.
[32] ICIS, 2024. Europe Data Centre Power Demand
[33] Uptime Institute, 2024. Global Data Center Survey 2024.
[34] Google, 2025, Google data center PUE performance
[35] IEA-4E, 2022. Energy Efficiency Metrics for Data Centres
[36] Epoch AI, 2025. Trends in AI Supercomputers
[37] Goldman Sachs, 2024. AI, data centers and the coming US power demand surge.

Table 4: Traditional data centres tier classification system

Component	Typical Power Draw
NVIDIA A100 (1 GPU)	c. 400 W
NVIDIA H100 (1 GPU)	c. 700 W
8× GPU AI server	c. 3-6 kW per server
1 AI rack[21]	50-120 kW
Small enterprise rack[21]	0.5-10 kW

Source: multiple sources

In 2024, rack sizes of 4-6 kW were the most common according to 40% of Uptime Institute survey respondents.[33] The number of respondents responding 7-8 kW rose from 18% in 2023 to 24% in 2024, indicating a trend towards higher-density racks over time. The average rack rating across all respondents was 7.1 kW, excluding outliers.

Stakeholders interviewed by Regen highlight that the installed capacity of edge data centres for 100% AI use could decrease from 28 MVA to 20 MVA within very short timescales for the same computational output. It is feasible that, in order to compete and stay in the market, these efficiencies will need to be adopted by most operational data centres.

Until recently, efficiency gains have resulted in a stabilisation of data centre demand. However, with the recent rise of AI there has been a stark reverse of this trend, with energy usage increasing significantly. This shift is expected to continue alongside increased efficiency gains. This has been considered and reflected in NGED's Distribution Future Energy Scenarios (DFES) 2025 long-term data centre capacity projections out to 2050, limiting the extent of projection growth.

Cooling efficiency

Cooling demand varies widely depending on the type of cooling, IT equipment and processing needs. For example, AI data centres with powerful GPUs require specific temperatures and levels of humidity. Data centres sited in areas with variable temperatures and humidity levels may have higher cooling equipment usage requirements.[29]

Traditional **evaporative cooling** is very water-intensive. **Air cooling** methods use cold channelled air through HVAC units, whereas **free cooling** methods can be leveraged in cooler climates to utilise ambient air temperatures. **Liquid cooling** methods of various design utilise water or other liquids to cool the infrastructure. Finally, innovative cooling technologies such as Evaporative Fibre-Membrane Cooling are also being explored that could drastically reduce energy demand from cooling.

Innovative cooling methods can increase the overall efficiency of a data centre, with varying degrees of efficiency gains for water versus electricity usage. The benefits of air cooling over liquid cooling include better use of water resources. However, more energy usage is required for air cooling, while free cooling drastically reduces energy usage by making use of natural cooling methods. Stakeholders interviewed for this report stated their intention to use air cooling systems to reduce water resource reliance.

Another innovative cooling solution involves **closed-loop cooling**, which significantly improves energy efficiency and limits water consumption. Using this method, plant efficiency can increase by 48-90%.[38]

Data centre decarbonisation options

Current and future data centre energy consumption is a key concern for the electricity networks. There are, however, several solutions that could help reduce energy consumption throughout the year and improve sustainability at data centre sites.

Long Duration Energy Storage (LDES) is capable of storing electricity for several hours a day, and could address data centre UPS requirements. In Scotland, Highview Power is spearheading a groundbreaking LDES project: the Hunterston facility, designed to deliver 2.5 GWh of storage capacity and capable of powering approximately 650,000 homes for 12.5 hours. This project exemplifies how long-duration storage can stabilise clean energy delivery and transform data centre power provisioning, paving the way for more resilient, low-carbon digital infrastructure in the UK.

On-site renewable generation allows data centres to produce renewable energy directly at their facilities, cutting dependence on the grid and reducing transmission losses. By definition, data centre buildings potentially lend themselves to rooftop solar arrays. In the UK, Iomart has installed 560 rooftop solar panels at its Maidenhead data centre, generating around 250 MWh annually and avoiding c. 96 tonnes of CO_2 emissions. With an estimated site demand of c. 15,330 MWh per year (2.5 MW at 70% load factor), the rooftop solar provides c. 1.6% of annual energy consumption. Other examples, such as NLighten's 478kW installation at Milton Keynes, boast meeting total power requirements during certain periods of the day in the summer months.

Private wires involve establishing a direct, behind-the-meter power connection between a generation source and the data centre, bypassing the main electricity grid. This arrangement ensures a reliable (usually lower-carbon) electricity supply, avoids network charges, and can improve resilience against grid outages. With rising energy demands, private wire systems could become prevalent and financially attractive for UK data centre operators. Some consultancies are now providing bespoke services to explore private wire opportunities for data centres.

Long-term Power Purchase Agreements (PPAs) allow data centre operators to lock in stable, competitive prices and guarantee renewable energy supply contracts – often encouraging additional generation capacity to come online. In the UK, these contracts are gaining traction as both cost-containment and reportable carbon reduction actions. For strategic projects, early engagement with energy regulators and network operators is especially critical to ensure alignment with grid capacity, connection timing and cost-effective procurement.

Flexibility services could be leveraged by data centres, providing load shifting, moving energy demand to align with times of peak renewable output, mitigating curtailment and enabling more alignment between renewable output and IT load. Alternatively, data centres could be used for peak shaving. Data centres that host high-performance computing (HPC) or AI training processes may be more amenable to flexibility requests than those that require 24/7 connection from their customers.

Several recommendations to enable participation of data centres in demand side response (DSR) flexibility schemes were proposed in a recent paper by Energy UK. This included making flexibility markets more accepting of, and more attractive to, data centre participants and changes to network connection processes.[38]

Heat network supply is an option for data centres to reduce cooling requirements and generate revenue from the generated waste heat.
Data centres generate substantial waste heat, which can be captured and repurposed for district heating

[38] Energy UK, 2025. Powering the cloud: how data centres can deliver sustainable growth.

schemes, turning a liability into a community benefit. One UK project – funded by the Green Heat Network Fund – aims to use surplus heat from data centres in the Old Park Royal area to supply heating for more than 10,000 homes and 250,000m² of commercial space, delivering around 95 GWh of heat by the mid-2020s.

Small Modular Reactors (SMRs) offer the prospect of low-carbon, baseload power close to data centre campuses. The government-backed Great British Nuclear programme is advancing SMR deployment, positioning it as a potential energy option for future digital infrastructure. While still emerging in terms of commercial readiness, SMRs could provide highly reliable power with near-zero operational emissions, although cost, regulation and public acceptance remain challenges.

Geographical considerations

Regen's research shows that there are some key geographical factors that should be considered when understanding the potential location of data centre development hubs. This research has included identifying explicitly allocated zones and technical feasibility spatial factors.

AI Growth Zones

As discussed previously, AI Growth Zones are designated regions designed to accelerate AI infrastructure development in GB. They offer streamlined planning procedures, dedicated support infrastructure and access to high-capacity power supplies – crucial for AI-ready data centre deployments. Local authorities, regions and industry have been invited to apply and must evidence an array of technical criteria. These include factors such as evidencing at least 500 MW of power capacity by 2030, as well as suitable land, water, connectivity and planning requirements.

The first AI Growth Zone and site is set to be located in Culham, Oxfordshire, announced by Keir Starmer in January 2025.[39] The plan is to roll out a scalable AI-enabled data centre, beginning with 100 MW and expanding to 500 MW, with sustainable energy R&D incorporated. According to the Compute Roadmap published in July 2025,[40] Wales is also set to host at least one AI Growth Zone.

Several proposals have already been presented, indicating where some future AI Growth Zones may be, if successful. These include:

- North Lincolnshire: proposed four AI-enabled data centres, totalling up to 1.5 GW and £15 billion in private investment.
- Cardiff and Swansea: identified for data centre growth after Wales announced plans to host at least one AI Growth Zone.
- Doncaster: positioned as a Centre of Excellence for AI, submitting a bid to host growth zone infrastructure.
- Scotland's Silicon Glen corridor: post-industrial and coastal towns benefit from cool climate, abundant land and green energy capacities.

As of April 2025, 200 expressions of interest have been received for AI growth zones across the UK.[41] The scale of power requirements in these zones could mean transmission network connections are prioritised. However, the specific site design, individual data centre project buildout and network connection requirements within each zone may mean a mix of transmission and distribution network connections could

[39] Cities Today, 2025. UK announces first AI growth zone
[40] UKRI & DSIT, 2025. UK Compute Roadmap
[41] UK Government, April 2025. Press release: Investors and local authorities gear up as AI Growth Zone delivery gathers speed

be required.

Availability Zones

Before the announcement of the UK's AI growth zones, data centre developers had been historically targeting **Availability Zones** in the UK and globally.[42]

> *"An Availability Zone refers to a geographical area where a cloud service provider has decided to create their network by using capacity in a multitude of existing data centres."*
>
> ***– Knight Frank Data Centres***

The idea behind this model is that if one data centre fails, other data centres within the zone will step in to take over operational services. Data centres within each zone must be within a 20km radius of the primary data centre.

While other providers cannot typically build directly inside of an Availability Zone, they will likely be attracted to the areas based on the geographical factors that drew these zones in the first place. This trend is already exemplified by London's Slough Trading Estate and West London corridors, where hyperscalers and independent operators are clustered together. Similarly, in Dublin, many colocation data centres share power and fibre infrastructure within Amazon, Microsoft and Google's Availability Zones. An Availability Zone can be considered an industry-led hub area for data centre development. This is distinctly different from the government-defined AI Growth Zones.

From an interview with a data centre developer to inform this study, Availability Zones are currently more influential in their site-searching process than AI Growth Zones. As the AI Growth Zones policy progresses and zones are defined, there may be a shift in this trend.

Geographical factors

From reviewing further literature,[21] several geographical factors are prioritised for all data centre types, including:

1. Speed of getting a grid connection
2. Proximity to fibre network landing stations
3. Cooling requirement/access to lower ambient temperatures
4. Proximity to end users and low latency
5. Access to a skilled workforce
6. Land availability, especially brownfield sites.

The importance of these factors differs depending on the data centre archetype. Hyperscale data centres see access to green power as a high-priority factor to meet sustainability goals, as well as access to significant amounts of land. Conversely, colocation and enterprise data centres don't consider cooling requirements, land availability or access to green power as high priorities. There is also a consideration that services provided by data centres may be more replicable and reliable if they are located in proximity.[43] This could create efficiencies around shared data network connections and shared power infrastructure.

---

[42] Amazon Web Services, n.d.. Regions and Availability Zones
[43] NetApp Instaclustr 2025, Data center infrastructure: 5 key components and best practices

There is also potential to enable contingency or standby/data service redundancy across multiple sites, reinforcing the potential value for site clustering.

For AI data centres, geographical factors may differ depending on whether the facility is used for AI model training or AI Inference. The former is less dependent on proximity to end users, whereas the latter relies on low latency and can be found in smaller Edge centres.[23]

Fibre optic networks
The proximity to fibre optic networks is an important consideration for data centres. In the UK, the Great Western Rail Line and Grand Union Canal contain main fibre ducts, directly through NGED's licence areas. The potential for siting projects near this network is a significant locational driver.

> *"Sites in proximity to these fibre ducts are extremely attractive and have been a driving factor in the development of data centres around certain areas."*
>
> ***– Knight Frank Data Centres***

Water availability
Data centres require substantial cooling, a function that water can provide. Analysis indicates that many, particularly large-scale, facilities are located close to water sources and use this resource for cooling. However, there is increasing pressure on UK water resources. This, combined with the availability of advanced cooling technologies that require little or no water, such as dry/adiabatic hybrid systems and direct-to-chip liquid cooling, is shifting developer preferences. Heightened risks from water scarcity, tighter licensing requirements, and regulatory scrutiny further encourage the adoption of low-water solutions. Stakeholder feedback confirmed this trend, with operators prioritising cooling systems that minimise water demand.

While water remains a material siting consideration, it is increasingly secondary to other factors. Approximately 43% of data centre providers report on water usage when recording corporate sustainability metrics, compared to 89% that report on power consumption.[33] For new developments, low-water cooling technologies offer a lower-risk, more adaptable approach in the context of evolving environmental and regulatory pressures.

3. Load profile analysis

To understand the impact of future data centres on the distribution network, load profiles were created using measured data from existing data centres connected to NGED's network to model their expected behaviour. As part of DFES, these can then be aggregated with other technologies connecting to the distribution network to account for diversity between technologies, as not all peaks are coincident. To explore this, data centre demand profiles have been created for each half-hour across four representative days. Representative days cover a range of potentially onerous conditions: Winter (Dec, Jan, Feb), Summer (Jun, Jul, Aug), Intermediate Cool (Mar, Apr, Nov), Intermediate Warm (May, Sep, Oct).

Due to their potential size, it is not uncommon for a singular data centre to connect to a substation. Consequently, the absolute peak utilisation rate of any data centre has been used for the final load profiles in order to plan the network for edge-case peaks. The mean peak utilisation rate across all NGED data centres is also explored in this section for discussion.

Method

Meter Point Administration Number (MPAN) readings were collated from 32 data centres connected to NGED's network for every half hour since 01-01-2024. Utilisation rate was calculated as the demand divided by the agreed supply capacity of a site to normalise profiles across different sized sites. Sites were divided into 19 enterprise and 13 colocation sites and readings were divided into four representative days and weekdays/ weekends.
Mean peak utilisation rate is the mean peak across all data centres for a given half hour, whereas absolute peak is the highest recorded measurement at any data centre for a given half hour. Utilisation rate on the y-axis of graphs in the results & discussion section have varying scale to demonstrate profile shape. There were no zero readings in the mean peak to impact the result.

Results & discussion

Mean peak utilisation

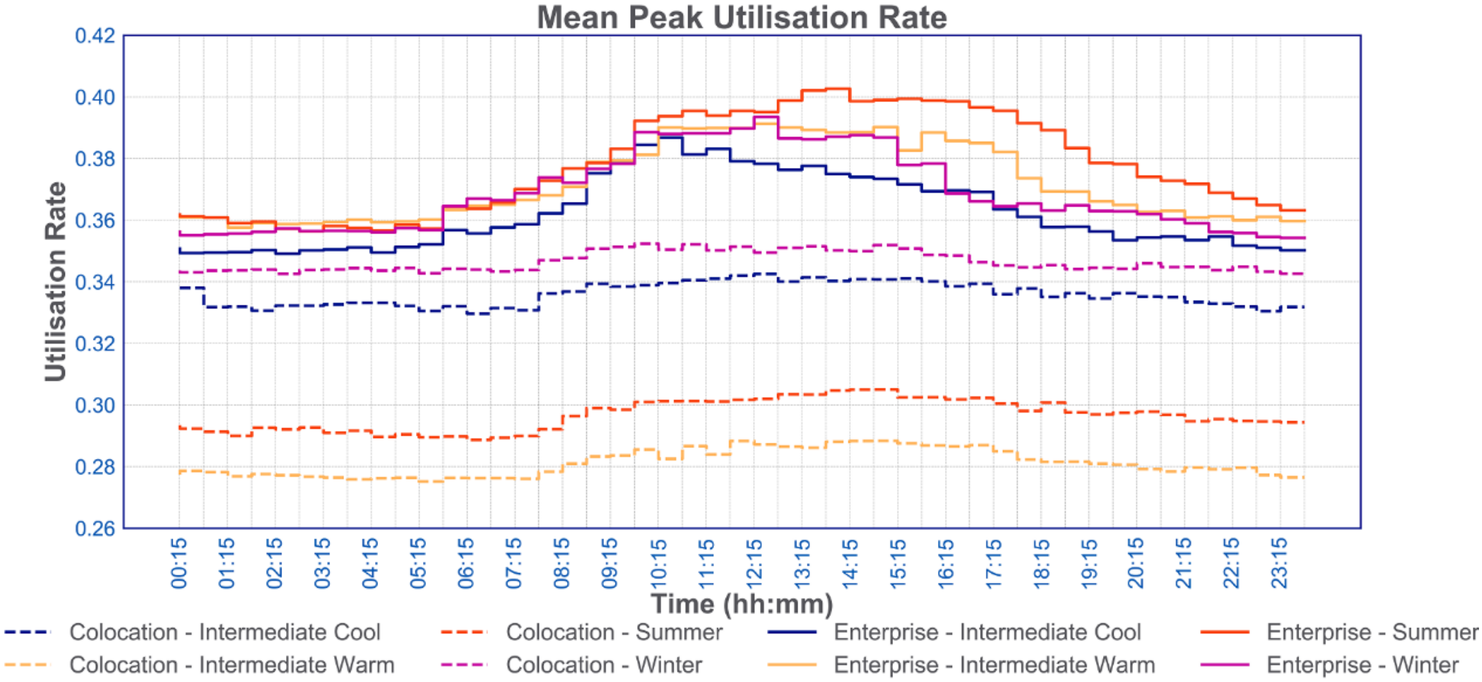

Figure 2: Mean peak utilisation across NGED data centres per type and representative day.

For every half hour and representative day, mean peak utilisation rate was higher in enterprise than colocation data centres. This could be due to enterprise data centres having a greater understanding of their business requirements as the site and associated connection agreement was created with their specific organisation in mind. Colocation sites are more likely to overestimate their demand when agreeing a supply capacity as they will rent out their server space to external users.

Colocation data centres highest mean peak occurs during the winter for every half hour. Enterprise mean peak alternates between summer, intermediate warm, and winter depending on the hour, with summer being the highest for the majority of time from 8am to 1am. A consistent summer mean peak would suggest that cooling load is causing the peaks due to the warmest average temperature increasing the demand. However, the results show that computational load is the likely driver of these peaks as the summer representative day is never the highest mean peak in colocation and not convincingly greater than other representative days in enterprise. Furthermore, winter has the highest mean peak in colocation and only the lowest for one half hour in enterprise, suggesting that colder temperatures do not significantly alleviate total data centre demand. As data centre AI server rack density increases in future and hyperscales connect, it's possible that ambient temperature will play a larger role in the peak utilisation rate as they require greater cooling.

Mean peak utilisation increases during typical working hours in enterprise data centres, likely reflecting the increase in computational load from business use. Additionally, the summer representative day working hour peak extends later in the day than intermediate warm which also extends later than winter. This could suggest that computational load from business use somewhat follows daylight hours, although intermediate cool does not follow this pattern. Although lower, data centre mean peak utilisation stays relatively close to average during non-working hours, including the middle of the night, reflecting the baseload from cloud storage. Colocation data centres display a fairly flat mean peak utilisation profile throughout the day. This may be because of tenants using colocation sites for services that don't necessarily follow typical 9-5 working hours such as website hosting and cloud storage.

Weekday/weekend

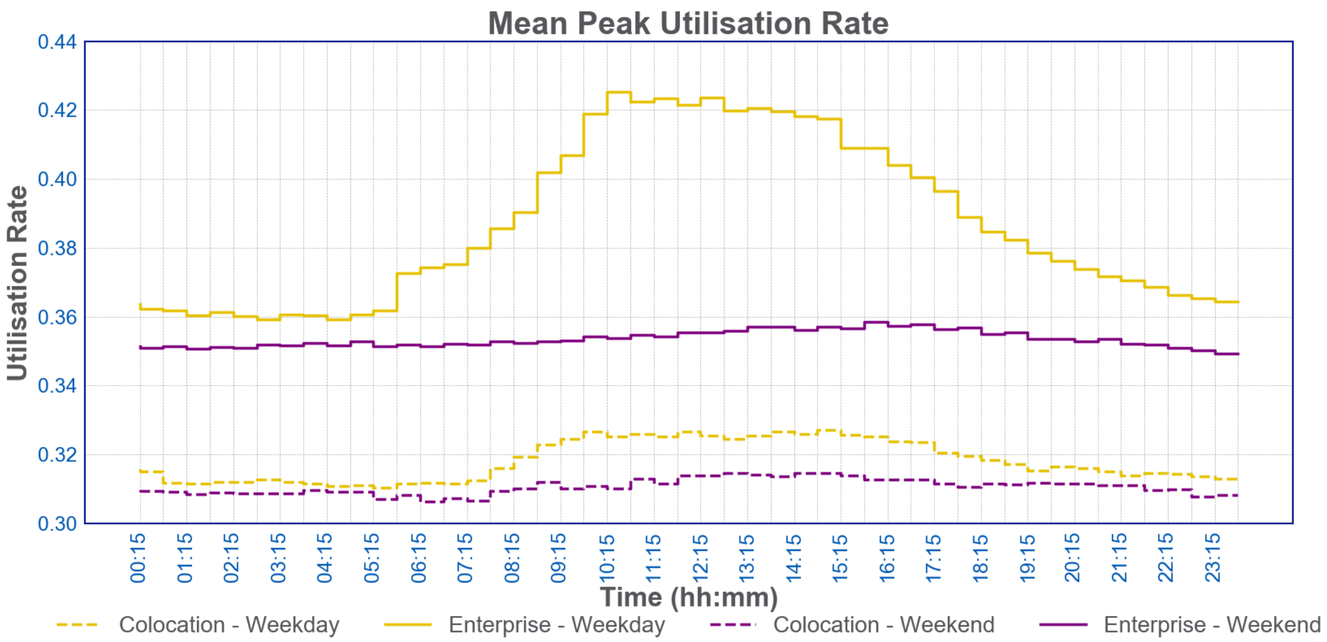

Figure 3: Mean peak utilisation rate across NGED data centres per type and weekday/weekend.

Mean peak data centre utilisation rate is higher during the weekday than the weekend across all half hours in enterprise and colocation data centres, particularly during typical working hours. The increase during

typical working hours during the weekday is far more apparent in enterprise data centres than colocation, increasing by up to 0.07 compared to 0.02. These results reinforce the notion of business use driving increased computational load causing peaks in utilisation in enterprise data centres.
The slightly higher utilisation during weekdays and weekday working hours in colocation sites suggests business use has some influence. As colocation sites often have multiple tenants with a variety of use-cases, they may experience intra-site diversity as different sections of server racks have non-coincident demand peaks. The higher mean peak utilisation in weekdays compared to weekends during non-typical working hours could be attributed to having 2.5 times the number of recorded days, increasing the chance of peaks occurring.

Absolute peak utilisation

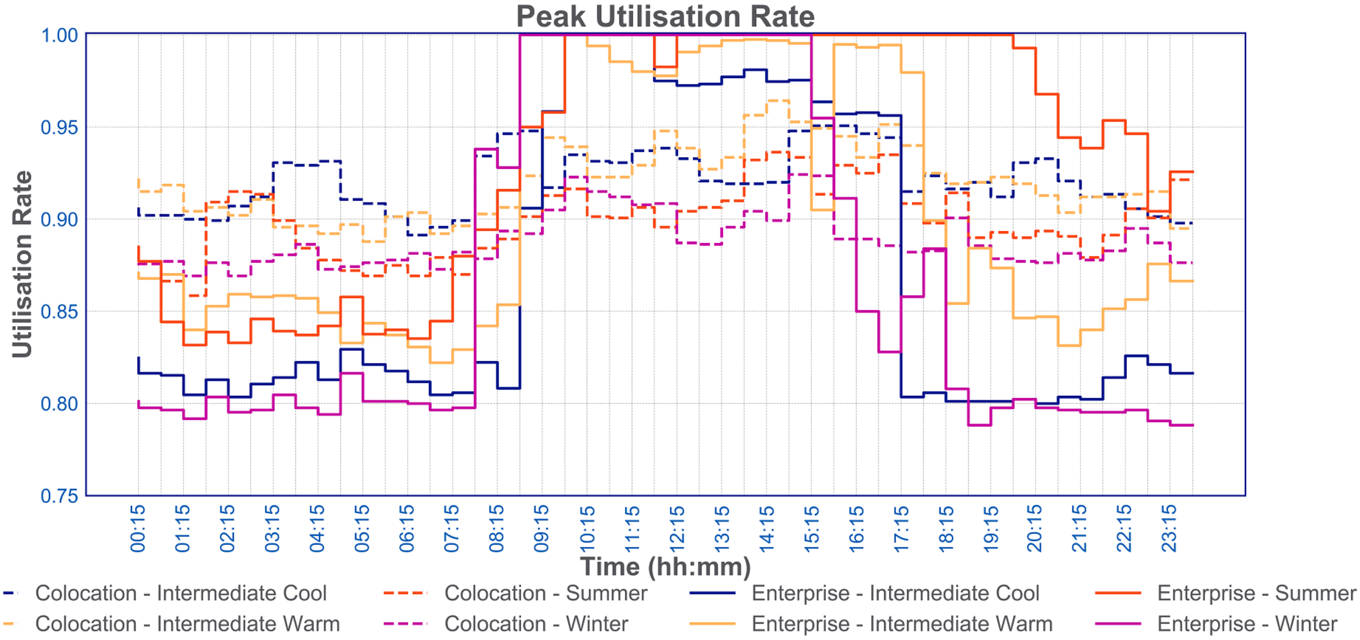

Figure 4: Absolute peak utilisation rate across NGED data centres per type and representative day.

Absolute peak utilisation rate is higher in enterprise than colocation data centres in all representative days from 10am – 4pm and reaches full utilisation for at least part of the day. The summer peak remains at 1 for the longest duration from 10:30am to 8pm.

During non-working hours, absolute peak utilisation was mostly higher in colocation than enterprise data centres during typical non-working hours (7pm-8am), with the exception of during the summer representative day only having a higher peak utilisation rate than enterprise from 1am to 8am. Enterprise absolute peak utilisation show a clear increase during the working day compared to the relatively consistent utilisation of colocation sites. This reinforces the idea of consistent load at colocation sites due to flat load use cases like cloud storage, as well as intra-site diversity from multiple tenants.

Limitations and Future Improvements

As more data centres connect to the NGED network, the sample size can be increased for future iterations of load profile analysis. This will produce more robust results and opens up the potential to split profiles by site size or AI server rack % in addition to enterprise vs colocation if enough data is available.

Currently, the density of AI server racks on existing data centre connections is low but projected to increase significantly on future sites. The highly variable demand peaks from AI server racks are therefore not fully captured in the existing dataset so future iterations of data centre demand profiles are required to refine the results.

The data centre technology in NGED's DFES is broken down into three subtechnologies: enterprise, colocation, and hyperscale. Currently, no measured data is available for hyperscale sites on NGED's network. Once more hyperscale data centres connect to NGED's network, measured data can be used to create load profiles for this category too. This is particularly important as hyperscales typically host the greatest density of AI Server racks.

4. Stakeholder engagement summary

Data centre developers and property developers with existing or planned connections in the NGED licence areas were interviewed to provide industry intelligence on the direction of AI growth and inform modelling assumptions. Key, anonymised insights provided by stakeholders are presented in this section.
Many of the data centre developers applying for grid connections were property developers intending to develop data centres on their assets.

Of the four interviews conducted to inform this assessment, three were data centre developers and another was a property developer interested in siting a data centre on one of their land holdings. Each developer interviewed had different expertise and relevant industry knowledge. The developers are also following different business models.

This engagement provided useful input and validation of assumptions and market trends to inform our scenario projections. Due to commercial sensitivities, specific insight and data were not focused on individual projects. The engagement results have been anonymised and summarised at a high level in this report.

Table 5: Anonymised stakeholder interviewee summary

In-text reference	NGED site capacity (MW)	NGED site DC subtype	UK DC ambition
Data centre developer #1 (DC1)	60	Hyperscale/ large enterprise	Targeting 1 GW across the UK
Data centre developer #2 (DC2)	135	Hyperscale	Targeting hyperscale and a number of AI-powered edge DCs
Data centre developer #3 (DC3)	162 and 8	Hyperscale and colocation	Several edge/ colocation centres in the UK and globally
Property developer #1 (PD1)	2 and 100	Hyperscale and small enterprise	--

Source: Regen conducted stakeholder interviews

Plans for new data centres in NGED licence areas

Stakeholders were asked to comment on their plans for developing data centres within NGED areas. This included any sites already with connection enquiries or those still in ideation phases.

One data centre developer we spoke to is targeting 1 GW of data centre capacity across the UK; each centre will have c. 60 MW of import capacity, targeting hyperscale and/or large enterprise sites.

One developer is focusing on two delivery business models. Across the UK, they are targeting the development of three hyperscale centres with around 50% of servers providing AI services, and a number of edge data centres with near 100% of servers providing AI services. These edge centres were described as hyperlocal, low-latency and located near the end users.

A third developer was a global data centre developer that specialises in colocation edge facilities. They are targeting several sites across the UK, ranging from c. 8-160 MW in NGED's licence areas specifically.

Efficiency improvements over time

DC2 mentioned that they are developing a number of small, AI-powered edge data centres across the UK. Each site is designed to be in c.28 MW scale, with an aim to reduce power requirements in the future, pending new Nvidia chips.

DC3 is targeting colocation data centre development, with their business model being centred around purchasing old data centre sites (that have become available, potentially due to the challenge and costs of keeping up with the rate of change of technology).

Energy consumption trends

Data centre providers were asked to provide information on the energy consumption of different types of data centre business models.

DC2 provided feedback on data centre energy consumption trends relating to AI use. As described previously in this report, this developer supported the idea that AI-enabled data centres have varied loads throughout the day: "AI server racks have an average rating of 80 kW per rack. Baseload AI power demand is likely to stay at 10 kW, while peaks and variations in load will occur throughout the day."

Our discussions with developers also revealed that domestic consumers often do not request energy-intensive AI-powered calculations. In contrast, business use cases for AI are likely to be magnitudes more computationally intense. As a result, the load is heavier during weekdays.

Different data centre models have different densities of AI server racks. Developers highlighted that hyperscale data centres will typically have 50% of racks dedicated towards AI and 50% towards cloud or storage services, whereas edge data centres will contain 100% AI servers. A developer pursuing c.60 MW hyperscale sites said they were looking at between 40-100 kW per rack.

Batteries and Uninterrupted Power Supply (UPS)

Stakeholders highlighted the importance of backup and UPS solutions to ensure that data centres remain connected and maintain data services 24/7. Batteries, especially longer-duration assets, have been mentioned as potential options for a lower-carbon approach to UPS. Developers speculated that the industry may be moving away from lithium-ion battery options due to associated fire hazard risks. They also highlighted that battery lifespans are shortened by AI inference consumption peaks, reducing life expectancy from eight to five years.

If battery usage is focused on emergency backup, however, its use may be minimal, limiting degradation. Charge-discharge cycles could shorten battery life if used as a flexible energy resource.

Geographical considerations

Engagement with developers provided details on geographical considerations for siting projects. Geographical considerations varied depending on the type of data centre archetype and business model being pursued.

Hyperscale data centre providers were more concerned with land size. One developer stated that a minimum of 50 acres is required when considering the location of these types of data centres. Other common land use constraints are considered, such as avoiding green belt land, flood zones and SSSIs. One developer highlighted that the proximity to Availability Zones was more important than AI Growth Zones.

Edge data centres and smaller businesses are more concerned with lower latency than hyperscale centres. Developers highlighted that they prioritise locating projects close to the consumer and are therefore more likely to site projects near population centres. It was highlighted by one developer that population density data from the 2021 census was used as a locational siting factor for its edge data centres. As a result, the most populous cities, like London, are being considered for hosting up to eight of their new data centres.

Colocation models must be close to other properties and have low latency, similar to edge data centres. Availability of brownfield land is also a consideration for some developers.

For all data centre business model archetypes, electricity network headroom was deemed more important than proximity to internet connectivity. DC2 also highlighted that grid reinforcements are more time-consuming and costly than new fibre optic cable routes. However, internet connectivity still influences how data centre developers do their spatial planning.

DC1 spoke about the need for the right social contacts when siting new data centres. London is a hotspot for developers due to the existing network of data centre developers.
Other bespoke considerations, such as supply chain availability and workforce, could also play a role in early plans. One developer highlighted that developing data centres in an area of Wales where gas power stations have been decommissioned presented an opportunity to retrain and redeploy people to work in the data centre industry and sites.

Conversely, feedback from a property developer advised that data centres need fewer employees than other industry sectors, such as the logistics sector. This means that less highway traffic is required, so data centres can be placed in more residential areas, where traffic congestion is more of a concern. It was highlighted that the traffic is five times as frequent in the logistics sector.

There was also a key difference between stakeholders around geographical considerations. The property developer that Regen engaged had not considered a wide array of geographical factors. Rather, the nature of the opportunity was more circumstantial and based on a specific opportunity relating to a property under management. This insight tells us that a certain degree of chance can be introduced where factors fall into place, meaning geographical considerations may not always play a role in the initial site-selection phase. However, this stakeholder stated that the feasibility of sites owned by property developers may be assessed at a later stage against geographical criteria when deciding if the site is optimal for data centre development, once these factors are understood.

Grid constraints and connection costs

Data centre developers were keen to disclose the grid connection costs and connection delays due to associated works. It was highlighted that grid connection costs and dates vary depending on the works required.

One project developer highlighted that their data centre was quoted around £20 million for a 100 MW connection. A 60 MW site being progressed by another developer was quoted in the region of £10 million and given an estimated connection year of 2033, due to the need for transformer upgrades. The same developer mentioned two other sites they are developing outside of the NGED licence areas; one site was quoted c. £17 million, connecting at 132kV. A third, smaller site targeting a 33kV connection was quoted c. £3 million. The developer expressed they would prefer to connect at 132 kV, but went for a lower voltage connection as there were no 132 kV lines at their desired location.

Delivery timeframes

Stakeholders were able to feedback on expected buildout timeframes for various data centre business models. Developers confirmed that hyperscale data centres are almost always built out in phases, sometimes over several years.

DC1 plans the buildout of their 60 MW data centre in phases of 10 MW over three to five years – 20 MW of infrastructure against 40 MW of whitespace, or IT infrastructure. DC2 responded that phased delivery of their 135 MW site would occur in three 40-45 MW phases, with a total delivery timeframe lasting five years. Each phase would take approximately 18 months to build out, which is a standard timeframe within the industry.

Edge data centres, depending on the capacity size, can also have a delayed buildout timeframe. For example, a 15 MW edge data centre might be built in two phases, across a timeframe of two years. Phased buildout data from stakeholders has been used to directly reflect buildout timeframe data centre modelling assumptions.

“Hundreds of millions are being invested. This ramp-up thing depends on that. Some sites we have a year and a half ramp-up, some five years. If we go from 1 to 2 MW, we can do it within a year. If we go from 1 to 50, then it might take five years.”

– DC3

5. DFES 2025 capacity projections

Modelling methodology

Baseline and pipeline research

Baseline research was conducted through engagement with stakeholders and desk research to supplement information and connection data provided by NGED. Several websites that track colocation data centres were used (e.g. Data centre map and Colo-x), press releases, planning documentation and any other information publicly available from online sources. Data collected included rack numbers, location, MW capacity, floor area, data centre type and % of server racks dedicated to AI.

Each baseline and pipeline site identified was categorised based on the following subtypes: small enterprise, colocation or hyperscale. This allowed for some basic assumptions around % of capacity likely to be dedicated to AI servers.

Table 6: Assumptions for AI use by data centre archetype

	Small enterprise	Colocation	Hyperscale
% AI capacity	10%	20%	50%

This initial archetyping was used where no further site information was able to be obtained during site research. If documentation was found that implied a different level of AI usage at a particular site, then this information was used to override the assumptions. For example, some facilities are large colocation sites closer to 50 MW in capacity, which straddle the colocation/hyperscale definition. In these cases, site research might indicate that these sites are closer to the hyperscale business model, and therefore the assumption of AI server usage is increased.

A survey undertaken by Roadnight Taylor indicated that one in five respondents were unable to move forward with their first location.[44]Error! Bookmark not defined. This finding had an impact on how pipeline sites were t reated. Not all sites are modelled to connect, due to the assumption that those still in 'enquiry' stage are speculative and could be considering multiple locations.

Pipeline buildout timeframes

Stakeholder feedback provided around buildout timeframes was applied (where possible) to individual projects, and was also used to inform the buildout of the remaining sites in the pipeline.

Projections to 2050

National long-term data centre scenario pathways, sourced from the FES 2025 data workbook, were used as drivers for the projections in this analysis for data centre demand capacity in NGED's licence areas beyond the pipeline timeframe, out to 2050.

FES numbers for data centre uptake are modelled in annual GWh. A load factor of 70% was applied, as referenced in FES: Pathway Assumptions 2025.

Baseline capacity in NGED's licence areas represents just 5% of GB's estimated FES baseline capacity figures. The size of the project pipeline suggests this could increase significantly over time. Areas with a large current baseline, such as greater London, could reach a limit regarding available land, incentivising developers to focus on other suitable areas across the country.

Geospatial distribution

[44] Roadnight Taylor, 2025. Powering Great Britain's data centre ambitions

Spatial distribution factors have been collected from existing research and stakeholder responses to understand the factors influencing developer decision making. Data compilation and collection followed to assess the availability and quality of existing geospatial datasets.
This process led to a final selection of the top three factors used in spatial distribution:

1. Location of known AI Growth Zones and Availability Zones
2. Presence of existing fibre-optic network infrastructure
3. Population density.

The Strategic Opportunity Areas – spatial distribution factors section of this report provides a further discussion on the specific data sources and reasoning behind the use of each geographical factor.

Baseline

The NGED licence areas have a baseline of 33 data centres with a total installed capacity of 79 MW; 37 MW have been categorised as colocation sites and the remaining 42 MW as enterprise data centres, hosting company-owned data infrastructure or hosted by third-party IT companies. The largest individual site is a 12 MW enterprise data centre in Leicester.[45]

Thirteen of these operational sites (both colocation and enterprise), totalling 13 MW, were assumed to have zero AI server capacity, based on individual site research and business model archetype analysis. Eight larger sites totalling 21 MW were assumed to have 10% AI capacity. Ten colocation sites with a total of 37 MW were assumed to have 20% AI capacity. Finally, two enterprise sites of 4 MW each were assumed to have 30% AI server capacity because they were associated with research institutions.

Table 7: Baseline summary by data centre type

Data centre type	Number of sites	Capacity (MW)	Estimated AI capacity (MW)
Small enterprise	20	42	4.5 (11%)
Colocation	13	37	7.4 (20%)
Hyperscale	--	--	--

South Wales

South Wales only hosts one distribution-connected site, a 2.1 MW colocation data centre in Bridgend. An additional two small enterprise sites operated by IT providers have been identified, totalling 4.3 MW.

West Midlands

The West Midlands has 13 baseline sites totalling 28 MW of capacity. Seven of these are small enterprise, single-user sites (11 MW), including financial institutions and one council-owned asset. The remaining six (17 MW) are colocation sites that host multiple customers' IT requirements.

East Midlands

The East Midlands is home to 12 sites totalling 28 MW across the licence area. Seven of these sites are small enterprise facilities totalling 15 MW, while colocation providers run the other five (13 MW).

South West

The South West hosts five data centres, four of which are small enterprise sites (11 MW). A single 5 MW colocation site is located in Bristol.

---

45 Information age, 2012. Santander to build £100m data centre near Leicester

Pipeline

There are 19 prospective data centre sites with an accepted connection offer in NGED's connections data, with a total capacity of 1.7 GW. All of these sites are modelled to connect up to 2032 at the latest (depending on the scenario). In addition to this, there are 31 data centres, totalling an additional 1.4 GW, that have either submitted a connection enquiry to NGED or are yet to accept their connection offers.

Overall, the pipeline of data centres in NGED's licence areas has much higher levels of hyperscale and large colocation models with likely AI capacity than the current baseline.

Table 8: Pipeline summary by planning status as of August 2025

Status	Number of sites	Capacity (MW)	Estimated AI capacity (MW)
Accepted	19	1,711	823 (48%)
Offered	9	140	49 (35%)
Enquired	22	1,275	516 (40%)
Total	50	3,126	1,388 (44%)

Of this pipeline of projects, a smaller number were modelled to connect. This is due to there being a large pipeline of projects compared to the estimated data centre capacity growth by the FES. To ensure that NGED DFES data centre projections do not significantly surpass a reasonable allocation for each licence area, the following pipeline allocation logic was applied, based on grid connection status and planning status or other online evidence of site progression.

Table 9: Pipeline site modelling logic

Planning status	Accepted	Offered	Enquired
Granted	Connects in all scenarios	Connects in all scenarios except FB	Connects in EE and HT
Submitted	Connects in all scenarios, c. 50% in FB	Connects in all scenarios except FB	Connects in EE scenario only
No information found	100% connect in EE, 50% in HT, 25% in other scenarios	Connects in EE and HT	Does not connect in any scenario

Stakeholder research, NGED pipeline connection notes and Regen site research were then applied to determine anticipated connection dates.

This has resulted in 1.8 GW (58% of the total NGED pipeline) to connect under the most ambitious scenario, **Electric Engagement (EE)**, with 722 MW, or 23% of the total pipeline, modelled to connect under **Falling Behind (FB)**.

Some pipeline sites have been categorised as hyperscale with as low as 25 MW of installed capacity. This is due to individual site research, where it has been determined that these sites are in close proximity to existing hyperscale campuses owned by the same data centre developers. This indicates that these sites are likely extensions of the existing, transmission-connected hyperscale campuses.

Site research findings
Some data centre providers have publicly available information around their data centre developments, which has fed directly into the data centre pipeline assumptions. Vantage data centres is developing several sites in the pipeline, such as one at Bridgend (120m^2).

Licence area-specific pipeline analysis results are outlined below.

South Wales

Table 10: Planning status of data centres in South Wales

Status	Number of sites	Capacity (MW)	Est. AI capacity (MW)
Granted	5	199	97 (49%)
Submitted	2	120	60 (50%)
No information	9	1,164	560 (48%)
Total	**16**	**1,483**	**717 (48%)**

Many data centre sites have entered the pipeline in South Wales, making up 47% of the total NGED pipeline of accepted, offered and enquired capacity. Ten sites (1.1 GW) have accepted their grid offers and are modelled to connect in at least one scenario. Many of the sites in planning are hyperscale models with high potential to host AI servers.

Four of these accepted sites are being developed by Vantage Data Centres, which is also known to have transmission-connected data centres in the area. Vantage Data Centre in Bro Tathan, Vale of Glamorgan, was approved in planning for 120 MW, but only has a 24 MW DNO grid connection offer. With a transmission network connection also located in the area, it is likely that the site may be using a mixture of transmission and distribution network connections. It is possible that small distribution connections are being located near transmission-connected sites under the same operator under the Availability Zone model. A similar story can be seen at a Vantage site in Newport.

West Midlands

Table 11: Planning status of data centres in the West Midlands

Status	Number of sites	Capacity (MW)	Estimated AI capacity (MW)
Granted	2	3	0.5 (17%)
Submitted	3	58	11 (19%)
No information	8	548	253 (46%)
Total	**13**	**609**	**264.5 (43%)**

Of the 13 pipeline sites situated in the West Midlands, two sites (238 MW) hold accepted connection offers. A further three sites, totalling 15 MW, had received a connection offer from NGED, two of which were extensions to existing data centre sites. As a result, five sites with a total capacity of 253 MW were modelled to connect in at least one scenario.

The only site modelled to connect in all scenarios was a 19 MW site in Birmingham developed by SUB1. The facility will be AI-focused, prioritise single-user occupancy and could expand to be 38 MW in the future.

East Midlands

Table 12: Planning status of data centres in the East Midlands

Status	Number of sites	Capacity (MW)	Est. AI capacity (MW)
Granted	2	104	51 (49%)
Submitted	3	270	54 (20%)
No information	10	550	265 (48%)
Total	**15**	**924**	**370 (40%)**

Of the 15 sites in the East Midlands pipeline, five were accepted to connect, totalling 356 MW. These sites were modelled to connect in at least one scenario.

One unique aspect of sites found in the East Midlands was the energy park model, making up four sites with a total import capacity of 470 MW. Three sites in the East Midlands were found to be submitted for planning, all of which will operate at energy parks. The Kettering Energy Park is set to host several technologies across multiple phases of development, including data centres. The site already hosts 23 MW of wind turbines and a future 40 MW solar farm for power consumption on site. According to planning documentation, energy demand at this site will come from data centres as well as other demand, such as manufacturing and R&D.

Therefore, there is uncertainty about how much energy will be generated on-site at energy parks alongside residual imported electricity from the network at different points throughout the year. Full grid connections are likely required to ensure a continuous power supply if on-site generation is unavailable. Only one energy park (40 MW) was modelled to connect in all four scenarios.

South West

Table 13: Planning status of data centres in the South West

Status	Number of sites	Capacity (MW)	Estimated AI capacity (MW)
Granted	--	--	--
Submitted	2	24	5 (20%)
No information	4	85	33 (39%)
Total	**6**	**109**	**38 (35%)**

The business models in the South West lean more towards smaller enterprises and colocation data centres. This has resulted in lower assumptions for total AI-enabled capacity when compared to other licence areas. Three of the six pipeline sites are located in Bristol. Four sites totalling 59 MW had received a connection offer and were modelled to connect in at least two scenarios. One colocation site in Bristol was submitted in planning, with plans to refurbish an old data centre previously owned by Proximity in two separate buildings. Various levels of capacity were modelled to connect due to uncertainty over whether the site is targeting 8, 16 or 24 MW of total capacity across two different grid connection offers.

Projections to 2050

Future scenario projections for data centre capacity were modelled out to 2050, based primarily on FES 2025 uptake rates. Once the known pipeline projects were modelled to build out, it was assumed that the development of data centres would continue, but more gradually, beyond 2032. This trend reflects a surge in data centre developer appetite in the near term, followed by a levelling out of the sector as data centre technology and chips continue to increase in efficiency gains and data centre site prospecting reduces.

In the medium to long term, it is assumed that new hyperscale facility locations will be limited. Additional capacity growth comes from new modular designs and increased capacity due to chip capacity gains at existing facilities. As a result, edge, colocation and enterprise data centre deployment is assumed to represent 100% of post-pipeline projections in all scenarios, assuming AI inference, autonomous vehicles and distributed 5G are developed to enable future uptake.

The results of scenario projections for data centre capacity in NGED's licence areas are summarised in the following pages.

South Wales

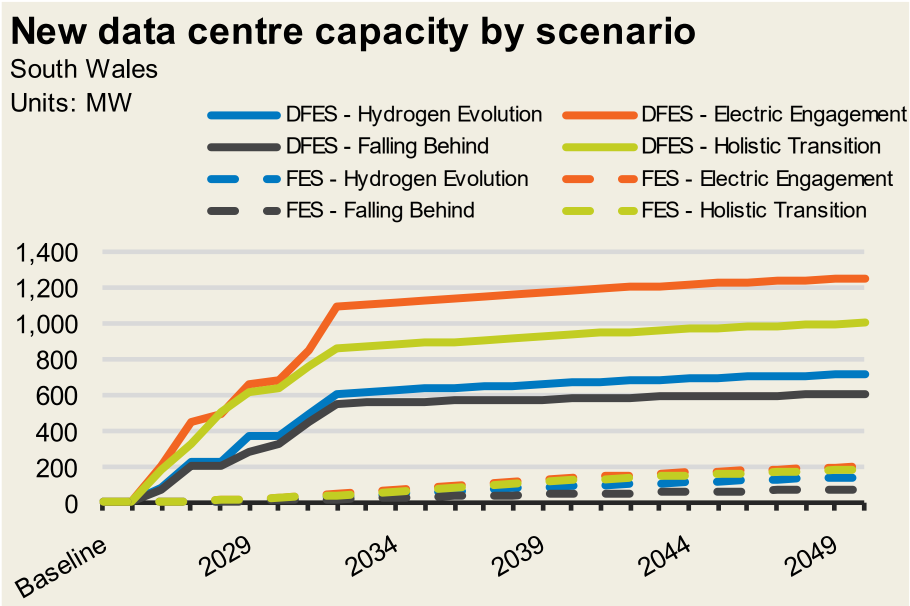

Figure 5: DFES data centre projections for the NGED South Wales licence area.

- Of all NGED licence areas, South Wales is projected to host the largest amount of capacity by 2050: 1.3 GW, equal to 54% of NGED data centre capacity.
- Most of this is due to the 10 accepted and offered pipeline projects totalling 1.1 GW that are modelled to connect.
- South Wales is also set to host a future AI Growth Zone, where data centre development may significantly increase.
- South Wales could host 27% of the estimated FES distributed capacity by 2050 under **Electric Engagement**.
- DFES projections are above the FES regional projections due to known pipeline of sites with accepted connection offers.

West Midlands

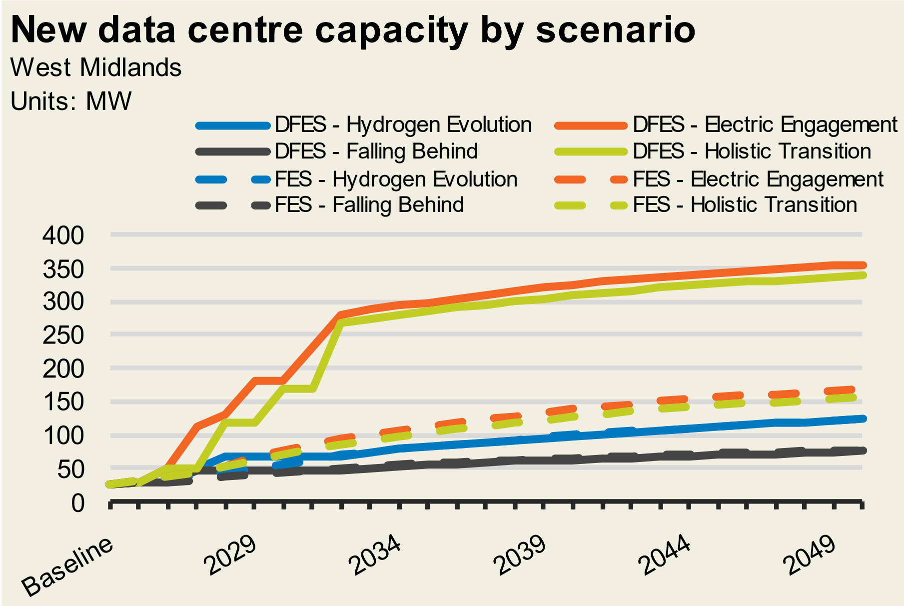

Figure 6: DFES data centre projections for the NGED West Midlands licence area.

- The West Midlands contains high-population urban areas such as Birmingham that are already seeing interest from data centre developers.
- The West Midlands is projected to host around 15% of NGED's data centre capacity by 2050 under **Electric Engagement**.
- The West Midlands hosts a maximum of 356 MW by 2050, representing 8% of the estimated FES distributed capacity by 2050 under **Electric Engagement**.
- The known pipeline of sites with accepted connection offers, modelled to connect in two scenarios, results in the projections being substantially higher than the FES regional projections in two scenarios.

East Midlands

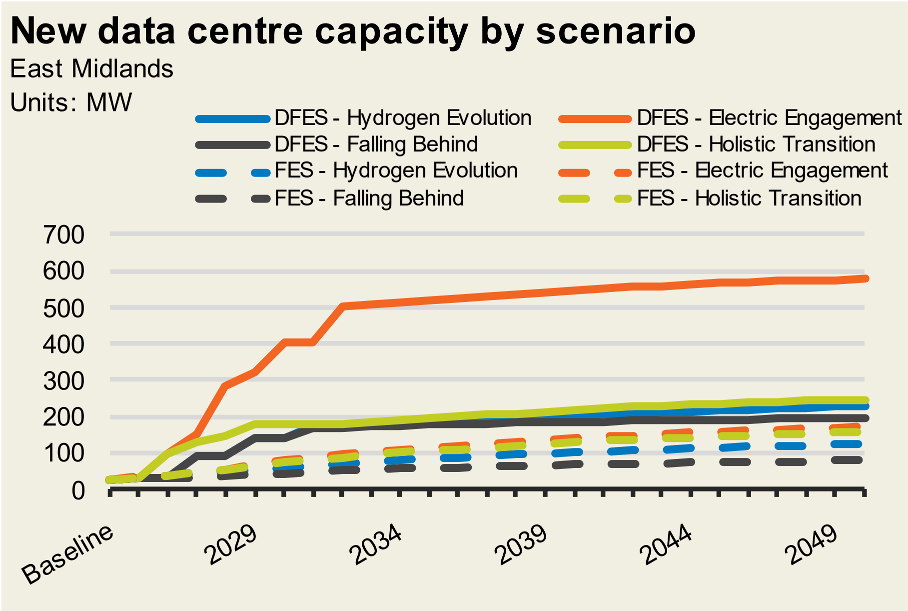

Figure 7: DFES data centre projections for the NGED East Midlands licence area.

- Urban areas in the East Midlands are less densely populated than other parts of NGED's licence areas. Some activity can be seen in Nottingham, where fibre optic network connectivity is present.
- The East Midlands is projected to host around 25% of NGED's data centre capacity by 2050 under **Electric Engagement**.
- The East Midlands hosts a maximum of 578 MW by 2050 in **Electric Engagement**: 13% of FES distributed capacity.
- Known pipeline sites with accepted connection offers are largely aligned with estimated FES distributed projections, except for **Electric Engagement**, where a 200 MW site which has submitted an enquiry for a network connection offer is modelled to connect.

South West

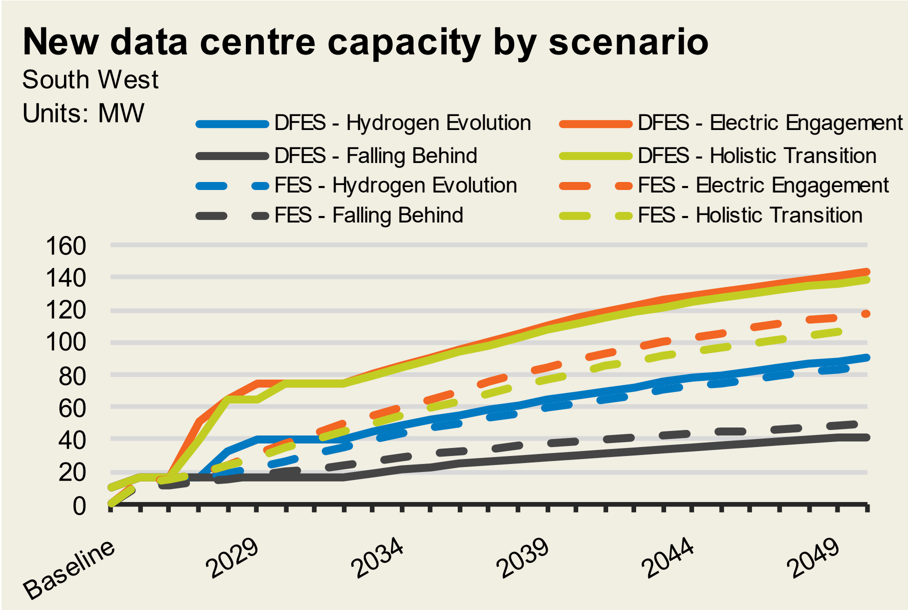

Figure 8: DFES data centre projections for the NGED South West licence area.

- The South West is projected to host 6% of NGED's data centre capacity by 2050 under the **Electric Engagement** scenario. This equates to 143 MW.
- This is due to the region's rural landscape and relatively low appeal to developers in non-urban areas of the region.
- Only one pipeline site (8 MW) has an accepted connection offer in the licence area and is modelled to connect in all three net zero scenarios. A further two sites, totalling 36 MW, have been awarded a connection offer and are modelled to connect in **Holistic Transition** and **Electric Engagement**.
- Urban areas, such as Bristol, will likely host the majority of future data centre capacity in the licence area.

Aligning with the FES 2025

These modelled projections in NGED's licence area have been compared to equivalent NESO FES scenario projections. Distributed totals were not provided by FES 2025, as in previous years, FES 2025 only provided data centre energy consumption data (in GWh) at a national level and does not distinguish between transmission and distribution networks.

Several assumptions were available to enable an estimated FES projection for NGED's region to reconcile with. To determine installed capacity figures from annual GWh consumption, a load factor of 70% was applied. The FES 2024 assumed that 40% of total data centre projects will connect at distribution voltage levels. Industry research and stakeholder feedback have supported this estimate, which was applied to determine an estimated proportion of distributed data centre capacity.

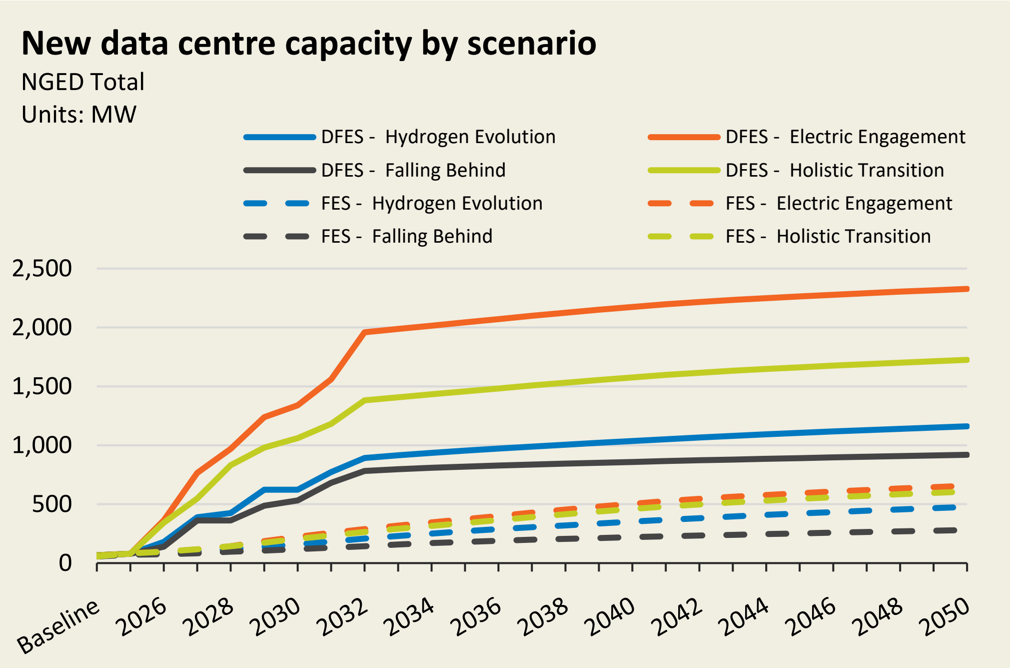

Figure 9: DFES and FES data centre projections.

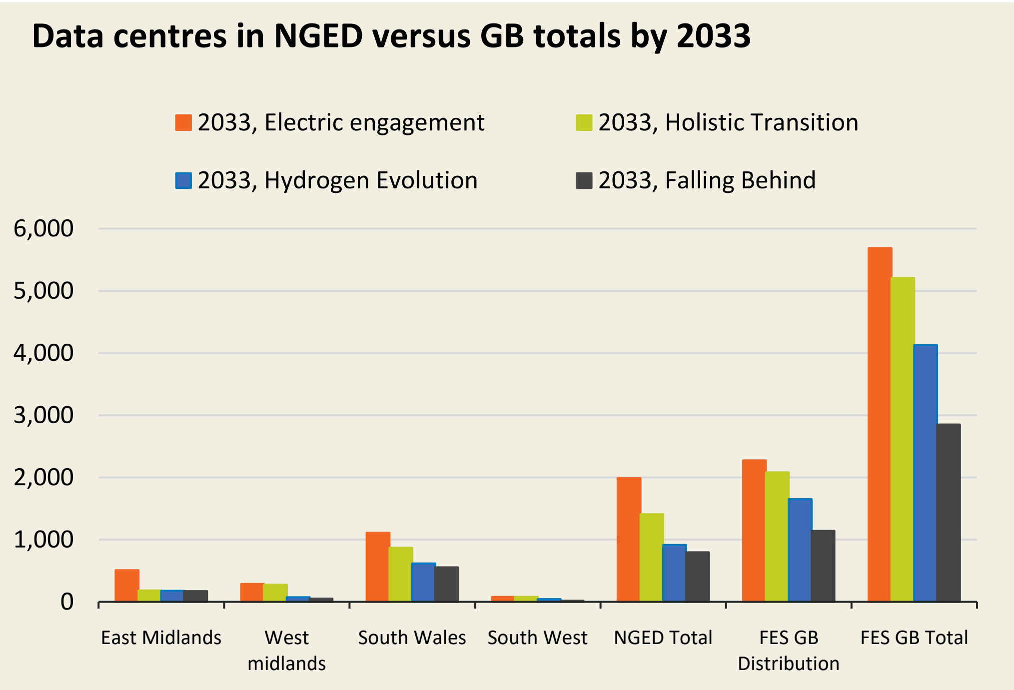

Figure 10: DFES and FES data centre projections by 2033.

Strategic Opportunity Areas – spatial distribution factors

From this report's literature review and stakeholder feedback, it is clear that data centres consider many locational factors when deciding where to site new facilities. The literature review highlighted several important factors for data centre developers, whether they are looking into hyperscale models or modular edge facilities. These factors included: speed to grid connection, low latency and proximity to users, access to low ambient temperatures, proximity to fibre optic networks, access to land and skilled workforce.

Engagement with data centre operators seeking connection to NGED's network yielded insights into data centre siting trends. Hyperscale data centres were most concerned with land availability, network headroom and proximity to Availability Zones. Smaller edge facilities were focused on areas of high population density.

Some spatial factors are more bespoke, such as a property manager or land owner using brownfield land assets with potential for commercial data centre projects. The factors taken forwards in the geospatial distribution of post-pipeline projections are summarised in Table 14.

Table 14: Spatial factors considered to spatially model post-pipeline projections

Spatial factor	Comments
Factors used to inform spatial distribution modelling	
Existing baseline and pipeline sites	The location of existing connected and contracted data centre projects is a key spatial indicator of where future sites could also locate. This is based on NGED connections data.
Known AI Growth Zones and Availability Zones	Information published by UK government. Live example of South Wales AI Growth Zone.
Presence of existing fibre optic network infrastructure	Data available from Neos Networks.
Population density	High-density population areas are a driver for data centres, providing services to businesses and commercial/industrial areas and being located in areas with fibreoptic coverage.
Factors considered, but not used to inform spatial distribution modelling	
Water resources/water scarcity	While a potentially important consideration for data centres' operations, screening out water-scarce areas would discount areas with known baseline and pipeline sites, e.g. London.
Green belt/other geographical constraints	Recent examples of approvals for significant data centre development in green belt land. So considered not an applicable factor to use.
Brownfield sites and land availability	While a consideration, data is difficult to obtain quickly, being spread across multiple local authority platforms.

As a result of the spatial analysis and consideration of geospatial factors, post-pipeline projections have been distributed to different Electricity Supply Areas based on the factors listed in Table 14. The map in Figure 11 shows areas of data centre opportunity based on existing fibre optic network capacity. All modelled projections are strongly weighted to fall in ESAs within these zones or along existing fibre optic lines. Colocation and edge data centres are more distributed towards densely populated areas.

As a result, areas with the highest distributed capacity are centred around Cardiff and Newport in South Wales, Bristol in the South West, Birmingham and Gloucester in the West Midlands, and Nottingham and Warwick-Coventry in the West Midlands.

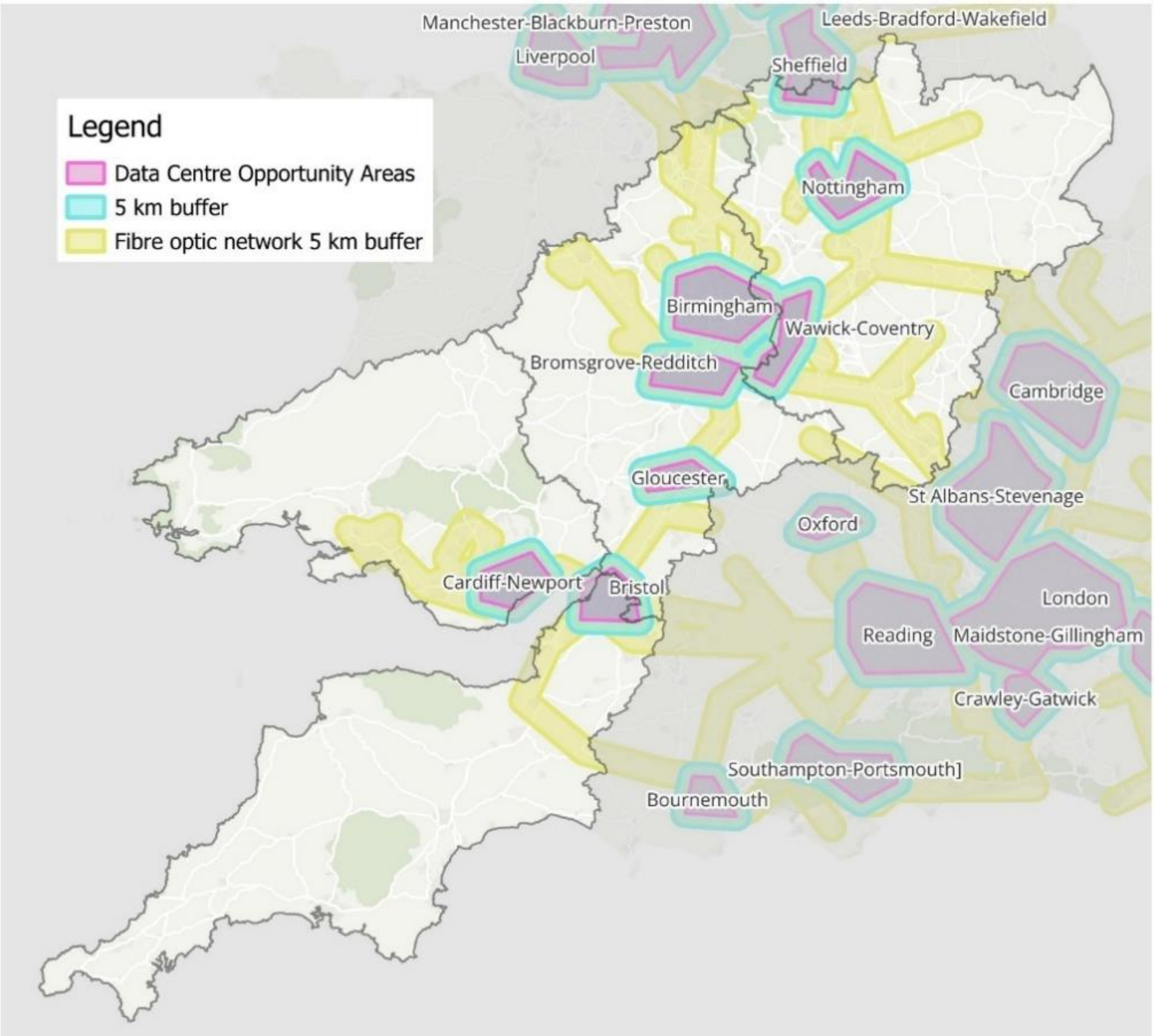

Figure 11: Map of Data Centre Opportunity Areas in NGED

6. Conclusion

This report summarises the first analysis of the potential future capacity of large-scale data centres connecting to NGED's distribution network.

The analysis is based extensively on known pipeline projects, supplemented by stakeholder engagement and validated spatial factors and drivers for future development.

However, the scenario projections provided should be treated as an illustrative view of potential future capacity. The data centre development sector (and the AI use case requirements underpinning them) is a new, constantly evolving space that electricity networks are rapidly trying to understand and adapt their planning processes to.

While NGED may not yet have seen the surge of pipeline capacity that has materialised in other areas (e.g. West London), as a network region, NGED hosts a wide range of geographies, from densely populated urban areas to industrial hubs and rural areas. It could therefore see a number of hyperscale projects connecting in a short period of time, and many businesses and industrial areas could see smaller-scale developments. This is compounded by the AI Growth Zone opportunity identified for South Wales, which could be a hub of deployment and fast-tracked capacity.

A key point to explore further is the potential for developers seeking transmission or distribution network connections. NGED could engage further with NGET, as well as the data centre developer sector, to continue to refine and validate their capacity forecasting, especially for the ED3 period.

REGEN

nationalgrid DSO