Department for
Energy Security
& Net Zero

Report on the Role of Ancillary Services to Encourage Low Carbon Operability

Carbon Trust Report

DESNZ research paper number: 2023/055

April 2023

Any enquiries regarding this publication should be sent to us at: remamailbox@energysecurity.gov.uk

Contents

Executive Summary

Decarbonisation of Great Britain's (GB) energy system is an essential step in achieving the United Kingdom (UK) Government's 2035 and 2050 targets on climate change. As the GB energy system transforms, it is crucial that power networks continue to operate in a manner that is both reliable and cost-effective. However, the increasing penetration of variable renewable generation technologies is leading to declining electricity system strength due to:

- A rise in asynchronous generation sources, leading to a reduction in the inherent stability of the energy system through declining inertia and short circuit levels.
- More variable sources of generation that are reliant on external input such as wind and sun to generate power, causing a reduction in certainty of supply in periods where there is low sun and wind.
- The highest renewable load factors are typically located at geographical network extremities, exacerbating physical constraints on the energy system as power has to be transported greater distances to demand centres.

These drivers are manifesting as challenges in two key areas:

1. The ability to cost-effectively manage thermal constraints on the energy system, and,
2. The capability of assets to meet system needs in extended periods of low renewable output.

To tackle both of these issues, the appropriate operational signals need to be in place to dispatch assets reliably and cost-effectively. Additionally, suitable investment signals need to be designed to encourage the deployment of the necessary low carbon technologies. One way of creating these signals is through ancillary services markets that National Grid Electricity System Operator (NGESO) design and operate to further facilitate security of supply.

This report considers the scope of ancillary services to address these issues as supplementary or alternative approaches. One such approach is the wholesale market reform under consideration as part of the Department for Energy Security and Net Zero's (DESNZ) Review of Electricity Market Arrangements (REMA). These reforms propose changes to how energy wholesale markets are priced (national vs. zonal vs. nodal pricing) and dispatched (central vs. self-dispatch), which will significantly impact operational and investment signals and reduce the need for some ancillary services (especially in managing network constraints).

However, as the implementation of large-scale market reforms are likely to take several years, some of the following proposals for new ancillary services should be considered for further development and then discontinued if and when market reforms render them redundant.

GB electricity transmission capacity is limited, which leads to higher-than-optimal levels of variable renewable energy curtailment. Ancillary services have a role to play in optimising power flows.

To cost-effectively deliver net zero, significant investment in GB transmission and distribution capacity is needed, irrespective of any reforms to market arrangements or new ancillary services.

Capacity constraints are currently primarily managed by NGESO, curtailing generation behind constraints and switching on generation in unconstrained areas, typically combined cycle gas turbines (CCGTs), using the Balancing Mechanism (BM). The build-out of variable renewable energy is leading to higher Balancing Services Use of System (BSUoS) costs. The dispatch of larger, high-carbon generation assets in the BM occurs for a complex set of reasons, including the ability of these assets to support wider system operability, ease of dispatch, and lack of visibility of smaller assets.

A wide range of ancillary services are currently being used to help create operational price signals needed to effectively manage the GB electricity system. Additional services put forward by NGESO as part of their 5-Point Plan on managing thermal constraints aim to limit costs associated with constraint management and encourage more effective use of existing network capacity. This report finds that further incentives to encouraging effective dispatch of assets and incentivising demand-side response (DSR)/storage are needed.

Some of the wholesale market reforms considered under REMA (e.g. nodal pricing (LMP) or zonal pricing) would provide more efficient operational signals, improving dispatch and reducing the cost of managing thermal constraints. A move to central dispatch would reduce the need for some ancillary services as well as the BM, as the NGESO would be able to optimise asset dispatch unilaterally, though some aspect of real-time balancing would still be needed to resolve forecasting errors or intervene in specific localities where constraints arise.

Even after considering current planned interventions, this report identifies possible gaps in operational and investment signals to cost-effectively manage thermal constraints.

When the transmission assets (cables) are constrained, NGESO needs to keep a small reserve to account for the risk of very short-term (i.e. minutes), variations in wind and demand. This report also assesses a suggestion from a market participant for a ‘shock-absorber’ ancillary service which would allow the transmission assets to be utilised at a higher rate and recommends further investigation into the practical implementation of such a service.

A second suggestion from a market participant is for a ‘strategic cycling service’. Though service would not materially reduce the volume of energy curtailed, it could reduce the cost of thermal constraint management and is recommended for further investigation.

In addition to the above ancillary services, and due partly to the synergy with contribution to capacity adequacy in times of low variable renewable energy output, this report finds that the deployment of long-duration energy storage assets (LDES) located behind constraints could help cost-effectively manage thermal constraints by absorbing curtailed energy during periods of extended constraints.

The future income streams of any LDES are highly uncertain. This report qualitatively assesses investment support for LDES through an ancillary service similar to the Stability Pathfinders but finds that a Cap and Floor arrangement is likely to be more cost effective. Although both approaches combine minimum levels of income to underwrite investment in case the value of stored energy is lower than forecast, a Cap and Floor would limit the upside to the investor in case the value of the stored energy is much higher than forecast, controlling costs to end customers.

Further to the investment signal, this report recommends the further investigation of a Storage Level Signal (SLS). This could be structured as a co-optimised ancillary service for constraint management and a capacity adequacy service to ensure that there is sufficient headroom in LDES to absorb energy during extended periods of constraints, and also sufficient energy to provide capacity adequacy in periods of low wind and sun.

LDES operators could receive availability payments to retain a state of charge at the NGESO's discretion, ensuring that assets are available to respond to system needs at any given time.

Further work is needed to assess the cost effectiveness of LDES for a combined thermal constraint and adequacy role, as well as whether an SLS would be cost-effective.

The need for electricity capacity adequacy in GB is changing. The primary means of bringing forward investment in capacity (Capacity Market (CM)) will need to evolve or be supplemented to ensure cost-effective, low carbon capacity adequacy.

The target of GB net zero carbon power market operation is increasing reliance on variable low carbon generation. The GB electricity system will therefore need to ensure sufficient low carbon capacity in extended periods of low sun/wind (Dunkelflaute periods). This capacity will come from low carbon thermal generation (nuclear, biomass and carbon capture and storage (CCUS)), as well as using interconnectors to import energy from connected countries with available resources. This report finds that LDES is likely to be needed to provide sufficient, low carbon capacity for the GB market.

The CM is currently the primary means of securing electricity generating capacity. However, this report finds that due to the changing nature of capacity adequacy, moving from the need to secure sufficient capacity for winter evening peaks to extended periods of low wind and sun, the CM in its current form is unlikely to cost-effectively encourage the investment needed in LDES and so will need to evolve or be supplemented by other signals.

As outlined above, this report finds that a Cap and Floor arrangement, rather than an ancillary service investment approach along the lines of the Stability Pathfinders, is likely to be more cost-effective for thermal constraint and capacity adequacy management, supplemented and co-optimised by an SLS service.

There is some scope for support schemes like the CM and CfD to be reformed to support cost-effective operability.

As outlined earlier, the CM will need to change or be supplemented to ensure capacity adequacy. This report also considered what other reforms, excluding auction reform, could be implemented to encourage provision of ancillary services. A key finding is that while individual ancillary services markets and the CM are not bureaucratically onerous on their own, lack of coordination and a need to qualify separately for each market likely has the effect of limiting participation, particularly of smaller, less sophisticated service providers. This report therefore recommends further work on streamlining and coordinating asset registration, qualification, and performance assessment to bring additional assets in and, through increased competition, reduce the cost of the ancillary services for end customers.

For generators benefitting from energy support schemes like the Renewable Obligation (RO) and the CfD, participation in any ancillary service that reduces their metered, active power output comes with an opportunity cost of losing the RO or CfD support for the reduced volume. Several options for the reform of CfDs have therefore been considered to reduce or remove this disincentive to participate in ancillary services.

However, as the gross cost of utilising CfDs for flexibility or ancillary services looks high, this cost is largely offset by a reduction on difference payments charged to suppliers. The net cost of CfD flexibility and ancillary services participation is therefore generally low. This report assesses several proposed reforms and finds that while these would reduce gross costs, they would generally have minimal impact on the net cost to customers.

Therefore, this report recommends a combination of enhanced reporting (to ensure that CfD generators do not seek to make excess returns when participating in flexibility and ancillary services) and consideration of BM and ancillary procurement reforms to require NGESO to consider the net cost of CfD service provision rather than the currently used gross cost. This report also recommends considering reforming CfDs to allow for ‘deemed’ CfD payments for reduced generation volumes if and when delivering ‘applicable balancing services’ as currently is the case in the BM.

Key Recommendations

There are ways to improve transmission capacity outside of building new pylons, which warrant further investigation.

While the Government recognises the need for network investment, there are actions that can be taken to speed up this process and ensure good use of pre-existing assets and infrastructure. This could involve retrofitting cables with composite cores or coatings that improve emissivity, which would not require the construction of new cables. Dynamic line ratings are in place on some lines, but an expansion of this technique where feasible could help increase thermal capacity. Improving asset visibility by accelerating energy system digitalisation will also help dispatchers (particularly in the BM) make better use of available assets.

Some ancillary services proposed warrant further investigation to understand their effect on net costs.

Examples such as Zenobē's strategic cycling and short-term (shock absorber) constraint reserve services outline how short-term storage assets can help cost-optimise curtailment actions and allow more power to safely flow through constrained boundaries. However more work is required to understand if these services will have a tangible effect on net constraint and curtailment costs.

The concept of a 'before-day-ahead' constraint price signal should be explored.

Conceptually, changes to wholesale pricing in REMA (nodal and zonal) shift the risk of non-availability from the NGESO to generators. A 'before-day-ahead' constraint price signal could conceivably do this with less complexity than nodal/zonal pricing, encouraging assets to self-curtail. However, more work is needed to understand if this would have any impact on net constraint costs, rather than simply reallocating them.

An investment support scheme should be designed for LDES.

This scheme could potentially be modelled on the revenue cap and floor schemes applied to interconnectors. Such a service would need to de-risk investment in multiple revenue streams, have procurement/contract horizons that match asset lifetimes, and contain a locational signal that ensures LDES is deployed in the correct locations.

Investigate possible service design options for constraint/adequacy ancillary services.

A service that would allow NGESO to control the charge state of storage assets, and that would ensure they are available to respond to system needs (constraints/adequacy), should be explored. Further work is needed to understand the technical service specifications/requirements, along with the inter-seasonal procurement volumes.

Change the cost assessment for CfD dispatch in flexibility and ancillary services from gross to net.

When procuring flexibility and ancillary services, NGESO is tasked by Ofgem to do so at the minimum cost. However, the current procurement approach considers the ‘gross’ cost to NGESO rather than the net cost to consumers in that it does not consider the change in CfD payments that result from dispatching a CfD generator in a flexibility or ancillary service. Enhanced reporting and a change to NGESO procurement rules would more accurately reflect the (lower) net cost of ancillary services to GB end customers.

Streamline and standardise administrative procedures for ancillary services and CM, making the services more accessible to asset operators.

In GB today, there are several different markets for ancillary services, including for local and national ancillary services as well as the CM. As these services often have different qualification, credit and performance assessment criteria, the cumulative administrative and financial burden on service providers (particularly smaller service providers), is likely material. A streamlining and standardisation of participation in these markets could bring forward additional capacity and lower costs through competition.

1 Introduction

The UK Government has set a target to deliver a fully decarbonised power system by 2035, subject to security of supply. This is a crucial step towards reaching the UK's 2050 Net Zero goal. Most of the low carbon electricity that will need to be generated will come from intermittent sources, primarily wind, the natural resources for which are often located far away from demand. With a limited amount of transmission capacity, this has sharpened the focus on operability in general – i.e., how do you make a system dominated by variable generation far from demand function both practically and cost-effectively.

Thermal constraints and capacity adequacy (both at the transmission and distribution level) are the two operability areas of primary concern, with a secondary, but important, focus on frequency management, voltage control, restoration and within-day flexibility.

Thermal constraints are leading to a higher-than-optimal system cost of energy for customers in GB. During periods of high wind, low carbon generators far from demand (mostly Scottish wind) are often curtailed. To balance this reduction, higher marginal cost and higher carbon generators (mostly gas-fired plants) are often dispatched nearer the demand centres in England.

This report focuses primarily on transmission constraints over distribution constraints and will take both a system-level view of the challenge and assess the actions through which the ESO is managing them.

The opposite scenario to high wind and curtailment is that of low carbon adequacy in periods of low wind and low solar output, particularly if lasting for an extended period (one/two days). Meeting this challenge will require a combination of longer-duration storage (including hydrogen). This report will consider to what extent the Capacity Market (CM), the primary tool to ensure adequate capacity, can ensure low carbon capacity adequacy, in addition to the sufficiency of other interventions that are planned to manage this issue. The movement away from ensuring capacity for a few hours, which is clearly within the remit of NGESO, to ensuring capacity over several days, the responsibility for which more appropriately sits directly with the Department of Energy Security and Net Zero (DESNZ), leads to the question of institutional changes that are touched on briefly in the report.

The management of frequency and voltage within acceptable limits is becoming increasingly difficult in a system with falling inertia and shifting generation and demand patterns. NGESO has clear, if sometimes costly, methods of managing these challenges directly e.g., through long-term contracts such as those in Stability Pathfinders, and is progressing the creation of markets to deliver these services. The report will outline these challenges and identify any potential further efforts that could be helpful.

The report also touches briefly on the challenge of developing low-carbon restoration services also considering some of the main relevant areas of the ongoing Reform of Energy Market Arrangements (REMA) process and to what extent some of these reforms can ameliorate the operability challenges identified above.

The report outlines the operability advantages and disadvantages of the main wholesale market reforms, including reforms to: wholesale markets, including nodal and zonal pricing options (forms of Locational Marginal Pricing; LMP)[1]; bringing forward low carbon generation capacity, including Capacity Market (CM) and (ancillary services); reforms to the mass low-carbon power support policies, primarily Contracts for Difference (CfD).

Finally, the report recommends policy changes to address any identified gaps in terms of operability and any further work needed.

1.1 Aims and Objectives

The Carbon Trust was commissioned by DESNZ to expand the evidence base on system operability and its interaction with the changes posed as part of the REMA consultation. Nine questions were assigned from which we have structured our research and analysis.

This report aims to:

1. Present the operability challenges facing the GB energy system and evaluate the existing and/or planned interventions in place to meet them.
2. With reference to the different options presented under REMA, consider how operability challenges and solutions may manifest and behave.
3. Identify possible intervention strategies with a focus on ancillary services.

These objectives were explored to address the core questions that were presented in the initial project Terms of Reference (TOR) and in agreement with the DESNZ project managers, were prioritised as follows:

Core Focus:

- Thermal Constraints (TOR Question 4): What is the scope for addressing thermal constraint management through an ancillary service, as an alternative or as a supplement to the wider structural reforms being considered in the wholesale markets chapter of the REMA consultation which could in theory help mitigate thermal constraints?

[1] 'Locational marginal pricing' (LMP) is used throughout this report to refer collectively to both zonal and nodal pricing. When referring to exclusively one form of LMP we have distinguished accordingly.

- Adequacy (TOR Question 6): What is the role of ancillary services in creating the investment and operational signals for low carbon technologies and long-duration storage assets to meet system needs during extended periods of low wind/sun?

Lesser Focus:

- Support Scheme Reform (CfD/CM; TOR Question 7): What is the scope for reforms to support schemes like the CfD and CM to meet our objectives for operability?

High-level only:

- Long-term contracts vs near-term markets (TOR Question 5): Under what conditions might a long-term contract be appropriate for the provision of ancillary services rather than closer to real-time markets?
- Local ancillary service markets (TOR Question 8): What is the case for local ancillary services markets for the provision of services such as frequency response, reserve, and inertia etc. to both local and national systems?
- Co-optimisation of ancillary services (TOR Question 9): How would co-optimisation of ancillary services with energy dispatch work under: a) a central dispatch model with a single national wholesale price, and (b) a Locational Marginal Price model, also with central dispatch?

Aim (1) has been met through the first phase of the project's activities (Task 1), a summary of which is presented in Chapter 2 and this objective corresponds to Questions 1-3 of the DESNZ project TOR.

Aims (2) and (3) are examined in Chapters 3-6 and cover the Questions 4-9. At the request of DESNZ, Questions 4 and 6 (and to a lesser extent, Question 7) have been prioritised in this report. Questions 1-3 are also a priority but have been covered primarily in the first phase of work.

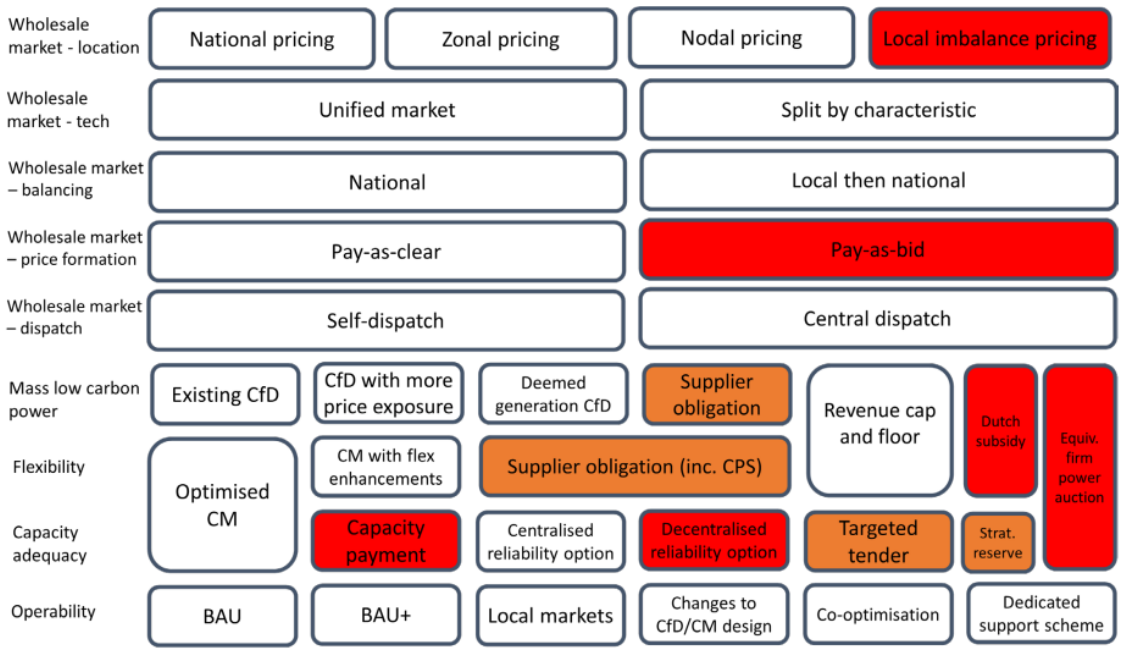
1.2 Methodology

In developing our response to the questions outlined in Chapter 1.1, the Carbon Trust has engaged with over 20 stakeholders including network operators, developers, investors, representative bodies, along with market operators, administrators, and regulators. These engagements, along with a review of desk-based literature, have been analytically assessed for this project. Primary quantitative modelling is out of scope and has not been undertaken for this report. However, where relevant quantitative modelling has been done by third parties (e.g., NGESO), high-level results have been included and highlighted. We have recommended further quantitative modelling be conducted in select areas (e.g., costs and service level) to aide future analysis and guide action.

1.3 REMA Options

In the British Energy Security Strategy[2], the government committed to a detailed evaluation of electricity market design to ensure it remains suitable for maintaining both energy security and affordability for consumers during decarbonisation of the electricity sector. In 2022, DESNZ launched a consultation on REMA, covering all non-retail electricity markets and seeking to identify the possible reforms needed to transition to a decarbonised, cost-effective, and reliable energy system.

Figure 1 - REMA options under consideration.



Source: DESNZ, 2023. Note: Options highlighted red and orange have been discounted by DESNZ based on consultation feedback.

The scope of REMA is extensive, including wholesale market reform, low carbon investment, flexibility, capacity adequacy, and operability (Figure 1). High-level options for potential reforms were put forward in the consultation, which ran until October 2022, with 255 responses from generators, developers, representative bodies, energy infrastructure, academia, suppliers, and private individuals.

The options put forward in REMA will have a significant impact on operability. While REMA explicitly considers changes to how ancillary services are provided and procured, adjacent changes to the wholesale market and support schemes will influence the assets and providers of operability services.

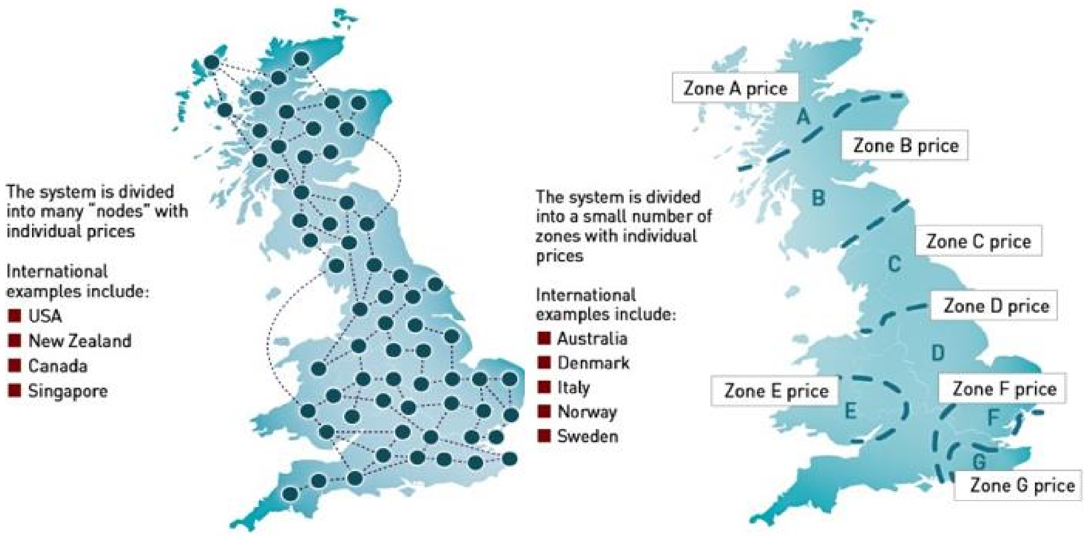
[2] DESNZ (2022) British Energy Security Strategy

While all the components of REMA are important, the changes in the wholesale market would potentially have the largest impact on the way energy is generated, dispatched, and consumed in GB. In the wholesale market, generators offer to sell their electricity to suppliers and other market participants who bid to buy the electricity they need to meet their customer demand. In GB, this market is regulated by the Office of Gas and Electricity Markets (Ofgem) and is administered by Elexon. Market location refers to how wholesale electricity is priced, in which there are three key options under consideration:

- *National Pricing*: GB currently uses national pricing, in which there is one price for electricity across the whole country at any given time. The price is affected by multiple factors including commodity (gas) prices, carbon taxes, demand, and the availability of wind and solar resources.
- *Zonal Pricing*: under zonal pricing, the transmission system is split into multiple zones, (typically in the range of 2-15) and is usually demarcated along congested boundaries. The wholesale price of electricity differs between these zones, but within zones, price remains uniform.
- *Nodal Pricing*: nodal pricing, is the most granular form of pricing setting. The transmission network is divided into, potentially, hundreds of 'nodes', each of which has an individual and separate wholesale electricity price. While complex, nodal pricing more precisely reflects the variability of value of electricity due to location across the country.

In addition to market location, the options of dispatch are also of considerable importance. There are two types of dispatch under consideration as part of REMA – central and self.

Figure 2 - Comparison between nodal (left) and zonal (right) pricing



Source: Watt-Logic, 2022

Dispatch:

- *Self-dispatch*: is a method of dispatching electricity where generators are able to choose when they operate and how much electricity they generate based on their own business interests. In self-dispatch, producers and consumers of electricity bid into the market to sell and buy electricity. The market price (in GB a single, national price) determines how much electricity costs in a given settlement period (currently 1/2 hour). Self-dispatch is the current mechanism used in the GB energy system.
- *Central dispatch*: is a method of dispatching electricity where a central entity, usually the system operator, decides which generators should operate at any given time to meet the demand on the grid. The system operator takes into account a range of factors, such as the availability and cost of different generators, to determine the optimal mix of generation to meet demand. In central dispatch, generators bid in with a price, but the system operator chooses which generator to dispatch and how much electricity they generate.

1.4 Operability

Operability, and its interaction with proposed changes in the wholesale market, is one of the areas where more evidence is required for policy makers to make effective decisions; this report begins to address this need through its primary research questions focusing on the factors concerning system operability.

Operability refers to the action taken by system operators to ensure the safe and efficient movement of power across the network, which is carried out through the provision of ancillary services. NGESO are responsible for ensuring operability and procuring/managing ancillary services and have set the aim of zero carbon operation capability for short periods by 2025, with continuous zero carbon operation by 2035[3]. Ensuring that the GB energy system remains stable and reliable is essential for meeting net zero targets. Understanding how the different options put forward under REMA affect operability is therefore key.

Operability requirements are fulfilled through ancillary services, managed by NGESO, and provided by energy generators and other market participants. To guarantee effective operability, NGESO procures ancillary services from assets either through longer-term contracts (e.g., the Pathfinder projects that source future services)[4] or through short-term markets (e.g., Dynamic Containment markets).

A full evaluation of operability requirements and the interventions in place to ensure the safe and reliable functioning of the electricity system is outlined in Chapter 2.

[3] NGESO (2023) A Net Zero Future
[4] NGESO (2022) Network Option Assessment (NOA) Pathfinders

As part of the REMA consultation, six options are being considered as ways to ensure operability as the electricity system decarbonises:

- *BAU*: a business-as-usual scenario, continues with the status quo and retains current services for ensuring operability.
- *BAU+:* makes incremental changes and improvements to the status quo but does not fundamentally change existing arrangements. This scenario will give the system operator the ability to prioritise low/zero carbon procurement.
- *Local Markets:* developing local markets for ancillary services that may give greater power to distribution network operators (DNOs) in managing operability.
- *Changes to CfD/CM design:* at present there are concerns whether energy support polices (CfD/CM) provide an incentive to deliver ancillary services. Changes to enhance incentives are being considered.
- *Co-optimisation:* The process of scheduling energy, reserve, and other ancillary services (in certain markets) is integrated into a single process to ensure co-optimisation of the two markets (ancillary services and energy supply markets). Through this co-optimisation process, the system operator determines whether an asset should provide energy, ancillary services, or both based on what would deliver the greatest system value. Co-optimisation would be dependent on a move to self-dispatch.

2 Challenges to net zero operability

2.1 Introduction

To answer the specific questions set out in Chapter 1, the first stage of this work focused on identifying the challenges affecting net zero operability and the suitability of current and planned market, policy, and regulatory interventions to meet them. This phase of work answered three key questions:

1. To what extent does REMA's case for change (CFC) identify the operability challenges facing system operators?
2. What policies, regulation, and markets are/will be in place for addressing the challenges facing operability?
3. How far are the policies, regulation, and markets likely to enable net zero carbon operability of the system by 2035?

This chapter will provide an overview of the results of this initial scoping and prioritisation work, highlighting the key operability challenges, the interventions in place to meet them, and whether such policies are likely to lead to a reliable and cost-effective net zero energy system.

2.2 High-level drivers of declining system strength

While there are many sources of declining system strength[5], there are several cross-cutting drivers that highlight where decarbonisation is leading to changes in the electricity system. These issues have been highlighted by system stakeholders and were identified by NGESO in their System Operability Framework (2023). Four key challenges are summarised below:

- *A rise in asynchronous generation sources:* Synchronous generators have large turbines that rotate at rates proportional to grid frequency, providing several inherent by-products such as inertia and stability. As the GB energy system decarbonises, these large synchronous generators are being replaced by inverter-based technologies that currently do not provide the same level of grid-forming capability (e.g., the ability of a plant/asset to contribute services such as voltage, stability, inertia, frequency, short-circuit levels, to the electricity system), leading to a reduction in the inherent stability of the system.
- *A reduction in the amount of easily dispatchable generation:* Traditional thermal assets such as open cycle gas turbines (OCGT) or combined cycle gas turbines (CCGT) provide guaranteed power that can be called upon when needed. Such assets can be

[5] Where 'system strength' refers to the resilience and reliability of the energy system.

quickly dispatched to provide system services and ensure operability, particularly in areas requiring flexibility and fast action such as voltage and restoration. In a decarbonised system, a greater share of generation is from smaller, distributed assets which are less easy to dispatch.

- *More variable forms of generation:* As the electricity system decarbonises, there will be an increase in generators that are reliant on an external input to generate power (e.g., sun or wind). This inputs uncertainty and variability leads to lower control of the output of energy, a problem exacerbated by a fall in the amount of dispatchable generation assets. Having less control over energy output creates challenges for ensuring security and adequacy of supply, for example in periods of high demand with low sun/wind. Additionally, this will result in a reduced ability for NGESO to resolve operability issues as many ancillary services are currently supplied as by-products of energy generation.
- *Geographic distribution of generation sources:* In a low carbon electricity system, some generation (such as offshore wind) will increasingly be located at network extremities. Along with causing localised voltage and current issues, this disparity can have a significant impact on the physical constraints of the transmission system as larger amounts of electricity will have to be transported over greater distances. In contrast, the network was originally designed around large thermal generation assets and is unsuited, with limited capacity, to adapt to the new spatial patterns.

2.3 Operability areas under consideration in this study

NGESO as the Electricity System Operator has statutory obligations to ensure that certain operational parameters and legal standards are met, such as maintaining system frequency between 49.5Hz and 50.5Hz and ensuring that 60% of electricity demand is restored within 24 hours in the case of a blackout. These obligations are set and regulated by Ofgem[6]. NGESO outline seven operability areas in their System Operability Framework (SOF) (2023)[7]:

- *Frequency Response and Reserve*[8]*:* Frequency is the number of times that the grid current changes direction. It is maintained at 50Hz and controlled through Response and Reserve services. Deviations outside of 50Hz (+/-0.5Hz) can lead to equipment damage and safety concerns.
- *Stability:* the inherent ability of the electricity system to resist deviations from normal operating conditions (~50Hz) after an event on the network such as a loss or surge of power. Stability is primarily provided through procurement of inertia and Short Circuit Levels (SCL); inertia refers to energy stored in rotating masses and SCL is a measure of the current that flows on the system during a fault.

[6] NGESO (2023) How we are regulated
[7] NGESO (2023) System Operability Framework
[8] Frequency Response and Frequency Reserve have been combined under one operability 'vector' by NGESO.

- *Reactive Power/Voltage:* Reactive power refers to the out-of-phase power that is generated by alternating current (AC) systems. Reactive power levels affect the voltage level which are maintained through the injection or absorption of reactive power at strategic locations.
- *Restoration:* the post-fault plans and services in place to restore power in the event of a blackout.
- *Thermal Constraints:* the transmission and distribution network have physical thermal limits on their power capacities. Constraint management deals with balancing of generation and demand to maintain operability.
- *Within-day Flexibility:* the ability to manage flexible and inflexible patterns of energy supply/demand throughout the day.
- *Adequacy:* determines whether there are sufficient available resources to meet electricity demand. This operability vector refers to long-term energy security specifically.

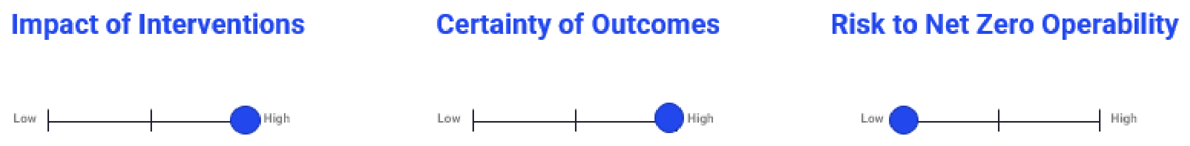
In the following chapter, actions being taken to address the seven operability areas are assessed and evaluated based on three criteria:

1. *Impact of Interventions:* representing the importance of an intervention in addressing the drivers/challenges associated with declining system strength. Projects with niche scope or geographic relevance are scored down; system-wide changes impacting multiple areas are scored up.

2. *Certainty of outcomes:* denoting the likelihood and intervention has the desired effect. Innovation projects and pilots with greater uncertainty are scored down; active markets/policy changes with demonstrated outcomes are scored up.

3. *Risk to net zero operability:* assessing the extent to which the challenges in each area pose a risk to both reaching net zero and ensuring the reliable and safe operation of the electricity system. Challenges with little potential to impact net zero operability are marked as low; challenges with significant potential to cause a detriment to the future energy system are marked as high.

2.3.1 Stability is declining, but solutions are known and therefore low risk



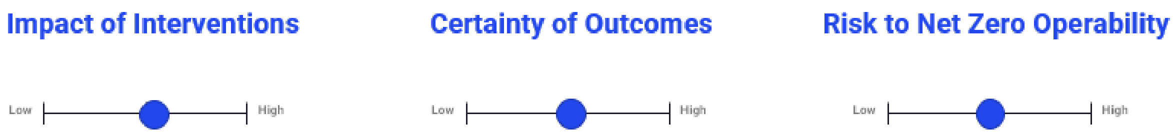
Stability is provided through inertia and SCL. Thermal assets such as coal or gas power plants provide inertia and SCL as an operational by-product, but as these are replaced by non-synchronous generation assets (renewables), the levels of both are falling. Falling inertia levels lead to a faster rate of change in the frequency (RoCoF) of the energy system in the event of a fault, meaning frequency response services and human operators must act with increasing speed to address fluctuations. Declining SCL makes system design more challenging, as electrical equipment such as transformers and switchgears may not be sufficiently sized, resulting in greater potential for failure during a short circuit event. The overall result of these conditions is a decline in the resilience of the system in the event of faults and a heightened vulnerability to faults in the first instance.

NGESO have planned a range of interventions to support existing and new sources of stability. The Stability Pathfinders Phases 1-3[9] were successful in procuring both inertia and SCL. Additionally, Grid Code 0137[10] is expected to provide a significant amount of inertia and SCL by allowing converter-based technologies with grid-forming capabilities to provide stability services. NGESO are also in the process of designing a future stability market to deliver cost-efficient procurement of stability services[11]. With the Stability Pathfinders, sufficient levels of inertia and SCL are secured until 2027. Grid Code 0137 was implemented in February 2022 allowing a low-cost, high impact means of securing stability services. NGESO has calculated future stability needs and is confident that system strength will be met to due to procured inertia levels and forecasts of inertia demand.

The Carbon Trust broadly agrees with NGESO's assessment that stability needs will be met with current or planned interventions, as outlined above, having a high impact. The proven ability to procure stability from new sources, such as battery storage, leads to a high certainty of outcome. The Carbon Trust also finds stability challenges to have a low risk on net zero operability. A risk remains however, due to the potential delays to low carbon synchronous generation, in particular new nuclear, which will be a key source of low-carbon inertia. Such delays may give rise to additional costs, but the Carbon Trust is confident that stability can still be ensured through other means such as synchronous condensers and virtual synchronous machines.

[9] NGESO (2023) NOA Stability Pathfinder
[10] NGESO (2021) GC0137
[11] NGSO (2023) Stability Market Design

2.3.2 Voltage (reactive power) Pathfinders have demonstrated solutions



Voltage control services, otherwise known as reactive power management, are required to maintain voltage levels within a specified range, either through absorbing or injecting reactive power into the electrical system. Challenges around voltage are being driven by increases in asynchronous generation and wider changes in generation and consumption. Asynchronous generation, such as wind and solar, have less capacity to absorb and inject reactive power due to their exclusive production of ‘real power’[12]. Additionally, as more renewable energy sources are integrated into the network, the production of reactive power is shifting from transmission networks to distribution networks as an increasing number of smaller generation assets are connected, which requires new methods for management and consumption. These underlying drivers result in increasing reactive power needs and voltage becoming more difficult and costly to manage. From 2019-2021, voltage management costs increased from ~£3/MVArh to ~£17/MVArh, increasing annual costs for NGESO from ~£70m to ~£190m[13]. During stakeholder engagement, NGESO highlighted a continued rise in reactive power that will need to be addressed in the future.

NGESO have planned several interventions to address the challenges associated with voltage. The Voltage Pathfinder[14] looked to identify the most cost-effective means of addressing high voltage issues in Mersey and the Pennines, areas of high concern. In July 2022 NGESO submitted proposal CM085[15], aiming to modify the Transmission Owner Code to require Offshore Transmission Owners (OFTO) to provide reactive power capability where they are able to do so, in doing so unlocking new sources of managing reactive power. The Voltage Pathfinder successfully completed procurement for reactive power services in locations with voltage issues, and while effective, this project was limited in geographic scope, covering just two areas. While the code modification (CM085) may have a large impact, it has not yet been approved, and thus our assessment rated it as having a reduced certainty of outcome.

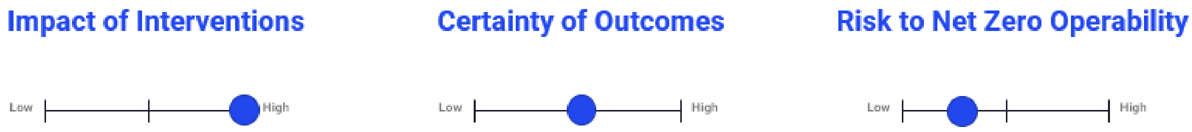
Overall, the Carbon Trust concludes that voltage challenges pose a medium risk to net zero operability as, despite the challenges NGESO face, we have confidence that most voltage regions in the UK could be currently managed with zero carbon solutions. Moreover, the technical solutions for managing reactive power are mature and the deployment of additional

---

[12] Real power is the electrical power that is most useful and used by appliances, lights, equipment etc. In contrast reactive power is non-useful power, created as a by-product and flowing back into the source. Additional equipment can be added to enable provision of stability services, but this comes at a cost.
[13] NGESO (2023) System Operability Framework This increase was significantly impacted by higher power gas, and therefore power, prices.
[14] NGESO (2023) NOA Voltage Pathfinder
[15] NGESO (2022) CM085

shunt reactors, static VAR compensators, synchronous condensers, and static synchronous compensators (STATCOMs) could be scaled up to meet system needs[16].

2.3.3 Restoration is a challenge, but NGESO are confident in this area



Restoration services are being impacted by the decrease in dispatchable generation and the increase in more variable sources of generation. Restoration services were previously supplied by fossil fuel powered thermal assets that can quicky provide power in the event of a disruption or fault. Such assets are often grid forming, meaning they can energise the electricity grid. In contrast, renewable assets such as wind and solar are generally not able to provide power on demand (without storage) due to their dependence on weather conditions.

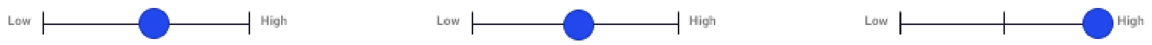
NGESO have a legal requirement to ensure restoration and so have introduced several interventions to provide restoration services. New policies such as grid code modification GC0156[17] and the New Restoration Standard[18] place more stringent requirements on NGESO in the event of a blackout. Additionally, interventions such as the Distributed Restart Project, are looking at how distributed energy resources (DERs) can provide black start capabilities[19]. Also relevant, is the Black Start Wind Tender[20], set to commence in 2023/24 which seeks to prove the feasibility of wind in providing restoration services.

This report finds that the challenges affecting Restoration are a relatively low risk to net zero operability due to the proven ability of DERs to provide restoration and black start services. The increasing ability of DERs, in particular battery storage, to provide restoration services was also highlighted during stakeholder engagements. Our assessment concludes that the current and planned interventions have a high impact with a good degree of certainty of outcome, in large part due to the obligations placed upon NGESO that impel strong action in this area, but also due to the increasing ability of some (grid forming) renewable assets and DERs to provide restoration services.

[16] REGlobal (2023) Reactive Power Management for Renewable Energy Integration
[17] NGESO (2022) GC0156
[18] BEIS/DESNZ (2021) Introducing a new 'Electricity System Restoration Standard': policy statement
[19] NGESO (2023) Distributed ReStart
[20] National Grid ESO (2023) Restoration Services / Wind Tender

2.3.4 Thermal constraints present large challenges

Impact of Interventions | **Certainty of Outcomes** | **Risk to Net Zero Operability**



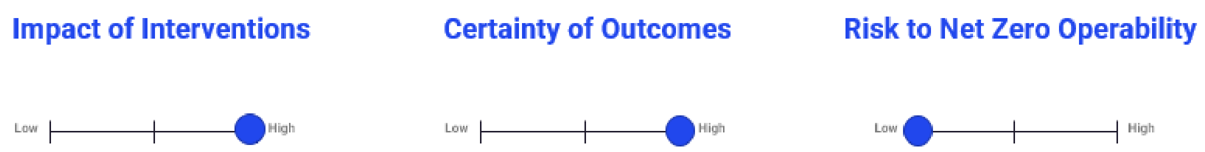
There are several underlying drivers exacerbating thermal constraint: the transmission lines and substations that make up the GB grid have a limited physical (thermal) capacity to transmit electricity; electricity generation is increasingly moving to areas relatively far from large demand centres; and demand-side assets such as heating, and transportation are becoming increasingly electrified. To resolve thermal constraints when network capacity is not sufficient to transport energy, balancing of the electricity network is undertaken by NGESO through the BM. This typically involves the curtailment of renewable assets and the 'turn-up' of high carbon assets (gas or coal) across constraint borders. Balancing requires payments to both sets of assets and results in a greater share of electricity generation coming from high carbon assets. The cost thermal constraints have risen due in part to higher gas and therefore power prices and was £2.4bn in 2022[21].

Thermal constraint management is a key operability area and a priority for NGESO and other actors, where multiple interventions have been introduced. These interventions and potential additional services will be explored in Chapter 3. Key interventions include the NGESO's 5-Point Plan for Thermal Constraint Management, the Local Constraint Market Design, and the NOA Constraint Management Pathfinder. These seek either to alleviate thermal constraints and/or to reduce the cost through more effective utilisation of existing network infrastructure and the procurement of new sources of flexibility.

These interventions are likely to have a high impact, in particular the 5-Point Plan for Thermal Constraint Management and the Local Constraint Market which set out targeted actions for reducing constraints in several ways. Although thermal constraint management is a priority area for NGESO, as these projects are yet to conclude and given the complexity of the challenges, the certainty of their impact is given a medium score. We conclude that thermal constraints pose a high risk to net zero operability due to both the financial and carbon costs associated with the BM. Thermal constraint is one of the operability areas where our we see room for improvement in the provision of ancillary services to ensure continued operation (explored in Chapter 3). Irrespective of constraint management, network reinforcement will be needed to for cost-effective net zero operability. This has been highlighted during stakeholder engagements with participants noting that while such interventions go some way to resolving the thermal constraints, what is ultimately required is a significant investment in network reinforcements and transmission capacity to meet future demand.

[21] Regen (2022) Seven solutions to the rising cost of transmission network constraint management

2.3.5 Frequency is an area with high confidence and adequate interventions

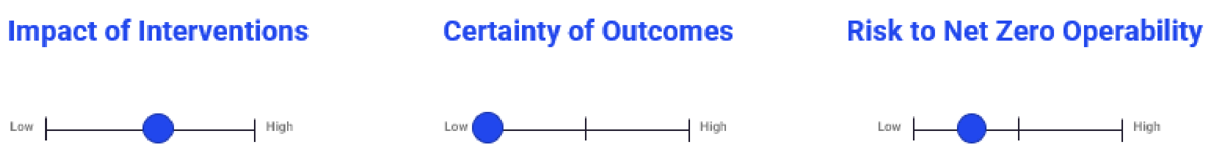


Two main factors are causing operability challenges in the area of frequency: falling inertia and increasing size of the 'largest infeed loss'. Inertia protects against significant frequency fluctuations due the overall, and aggregated, stored kinetic energy in the system. However, as discussed earlier, DERs such as wind and solar provide significantly less inertia compared to gas and coal generation assets. 'Largest infeed loss' refers to conditions whereby the size of new connections such as those for windfarms and interconnectors are increasing such that in the event of a fault, the resulting loss of input is greater. As inertia falls and largest infeed loss rises, the RoCoF increases leading to more challenging system maintenance, in turn needing faster and higher cost frequency services.

To address the challenges associated with Frequency, NGESO have implemented a suite of fast-acting frequency services: Dynamic Response, Reserve, and Containment[22]. These services were successfully implemented from 2020-2022 and despite large price fluctuations, we assess them as having a high impact with a proven outcome. The Accelerated Loss of Mains Change Programme placed new requirements on generation assets, and the Quick and Slow Reserve services[23] address challenges in frequency reserve. This overhaul of reserve services will have a large impact on how reserve is procured through new markets and refined procedures, but their effect will not become clear until the services are operational.

Overall, NGESO are confident in their ability to provide frequency services and the Carbon Trust agrees with their assessment, concluding that frequency challenges pose a low risk to net zero operability with the current and planned interventions having high impact and certainty of outcome.

2.3.6 Within-day Flexibility is 'new' but presents operability solutions



[22] NGESO (2023) New Dynamic Services
[23] NGESO (2023) Reserve Services

Within-day Flexibility has emerged as a recent area of focus associated with the short-term (12-24-hour) patterns of electricity demand. This is a cross-cutting issue that intersects with other operability areas, in particular Adequacy (discussed below). During stakeholder engagements, NGESO highlighted that Within-day Flexibility is as much a means of achieving outcomes as it is an operability challenge. This is also a new area of operability where limited work has been done at the time of reporting.

Electricity demand peaks at foreseeable (and largely inflexible) times during the day that reflect consumer patterns such as domestic evening activity. The electrification of heat and transportation is likely to cause these peaks of energy demand to increase. Managing these demand peaks is important for limiting costs, which tend to be highest during peak times, and reducing the pressures in other operability areas. Lowering peak demand may reduce the pressures associated with ensuring adequacy of supply and reduce the need for thermal constraint management. The increase in more variable and less dispatchable energy sources is also relevant, however Within-day Flexibility as a concept, and the interventions in place, are primarily concerned with demand.

Additionally, as energy efficiency investments and the roll out of solar PV progresses, minimum demand on the system (in particular, during summer afternoons) will likely fall, leading to a decrease in system voltage and inertia in those periods. Within-day Flexibility will help improve system strength in these areas through shifting demand into those periods and reducing system pressures.

Although Within-day Flexibility has emerged recently as a listed operability vector/solution within the System Operability Framework (2023), NGESO have been quick to implement services to ensure its provision. The primary intervention has been the Demand Flexibility Service[24], which seeks to incentivise consumers to reduce their consumption at key times, enabling NGESO to access additional flexibility in peak times such as winter evenings. Further interventions such as the Market Wide Half-Hourly Settlement Scheme[25] and the Flexibility Innovation Programme[26] seek to unlock additional sources of flexibility and if realised, will have a high impact on addressing operability challenges through relieving network pressures and accessing extra capacity through which to resolve challenges. Assigning a ranking to Within-day Flexibility must be caveated with understanding that it is not itself, an operability challenge, but rather a means to support operability. In other words, it is the lack of Within-day Flexibility that poses an operability risk. With this in mind, we conclude that the challenges associated with Within-day Flexibility pose a medium risk to net zero operability with current and planned interventions having a medium impact and certainty of outcome. Due to the characteristics of this operability area, the direct ability of NGESO to affect change is less certain.

[24] NGESO (2023) Demand Flexibility Service
[25] Ofgem (2023) Electricity Settlement Reform
[26] DESNZ (2021) Flexibility Innovation Programme

2.3.7 Adequacy poses the greatest risk to net zero operability

Impact of Interventions | **Certainty of Outcomes** | **Risk to Net Zero Operability**

Electricity supply is becoming increasingly reliant on weather-dependent generation, and this creates challenges owing to potentially sustained periods where output from variable generation is low and/or insufficient to meet demand. The challenge is therefore to ensure that supply can meet demand in a low carbon system at all times. Adequacy is currently primarily provided by capacity from weather-independent generation, mainly fossil fuels such as gas. However, these high carbon generators are being retired in order to meet net zero commitments. Adequacy poses a high risk to net zero operability if not addressed as thermal assets such as coal and gas will need to be retained to provide supply during periods of system ‘tightness’[27]. NGESO have undertaken several research studies exploring the problem and potential solutions to adequacy, though we conclude that these will have a low – medium impact on adequacy.

Similar to Within-day Flexibility, this is a cross-cutting issue that intersects with other operability areas and is exacerbated as responsibility for ensuring adequacy sits with multiple actors, including NGESO, DESNZ, and others. This is an area in which NGESO interventions are not sufficient, with NGESO action being limited to exploratory studies and scoping projects. Significant further work is required to address the challenge of adequacy. It is for this reason that adequacy and potential ancillary services to ensuring its provision is explored in more detail in Chapter 4 of this report.

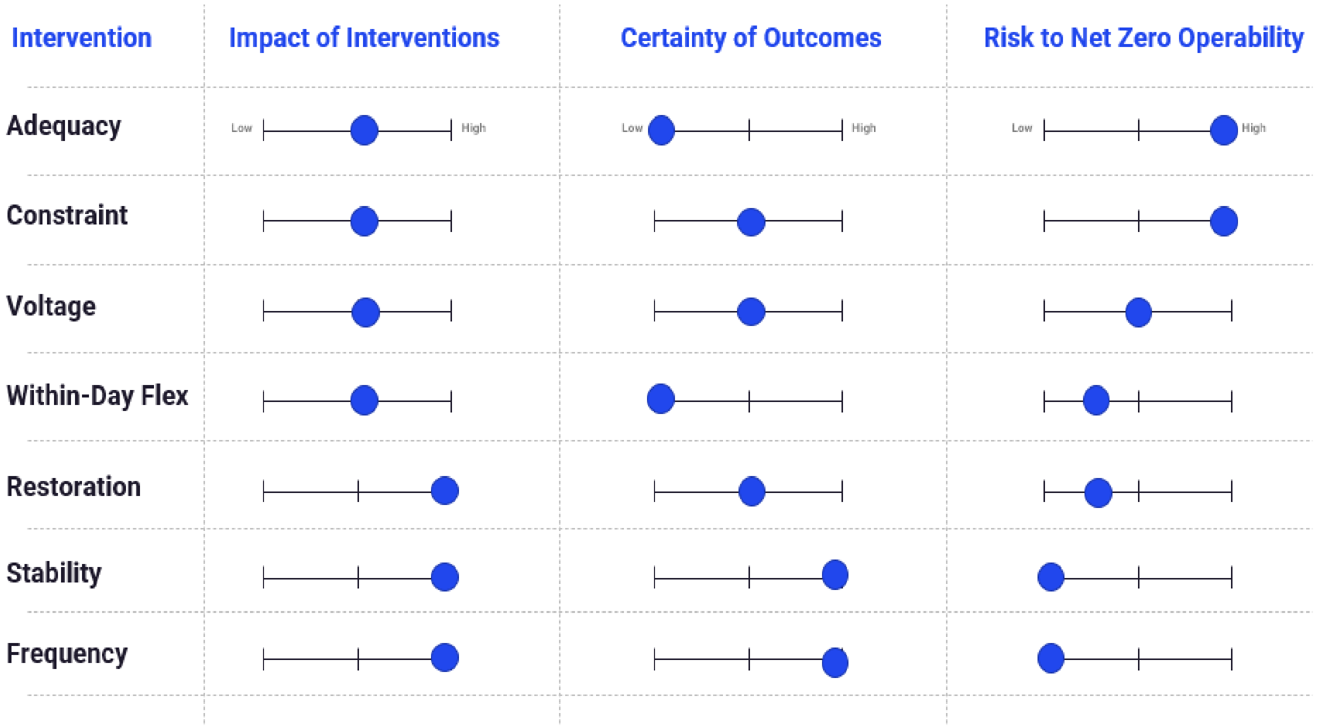
2.4 A ranking of perceived operability risks

A summary of our ranking of operability areas and the interventions to address challenges is shown in Figure 3. We conclude that for a majority of operability areas, the current or planned interventions by NGESO are sufficient to address the challenges, in particular in the areas of frequency, stability, and restoration. While we identify potential issues concerning long-term horizons, notably, potential delays to new, low carbon synchronous generation (e.g. new nuclear and CCUS), we conclude that these do not pose a significant challenge to the delivery net zero operability. Rather, if these delays manifest, the need for alternative solutions along the lines of Stability Pathfinders could be costly. NGESO noted during engagements that 2028 is likely to be a key year for stability as it is when Hinkley Point C Nuclear Plant is planned to start generating, but a delay to this plant would be visible years in advance, giving NGESO

[27] System tightness refers to the balance between the electricity demand and supply on the power grid at a given point in time.

time to put in place alternative arrangements which would secure inertia albeit with potentially higher costs.

Figure 3 - Summary ranking of operability interventions



Source: Carbon Trust, 2023. Note: The ranking of the risk posed to net zero operability indicates an overall level of risk and a rating of 'Low' does not mean that there are no challenges. A rating of 'Low' is based on our analysis where interventions and work by NGESO present low risk to net zero operability.

While this report ranks the interventions for Voltage and Within-day Flexibility as having a medium-level of impact, we believe these are still sufficient for ensuring net zero operability and do not pose a significant challenge to the safe and reliable functioning of a decarbonised electricity system. In particular for Voltage, while no enduring solution is proposed, the Pathfinder model could be extended to meet voltage challenges in these areas. While it has been noted that costs are expected to rise across almost all areas as services become more challenging to provide and procure, we rank Thermal Constraints and Adequacy as the operability areas in which there is the most concern. During engagements with both NGESO and a variety of other stakeholders, the Carbon Trust's ranking was reaffirmed[28], with stakeholders noting that in both of these operability areas, the issues driving the problem are expected to increase and pose a significant challenge to net zero operability if left unresolved. As such, we explore in more detail the challenges and potential solutions to these two challenges in this report (Chapters 3 and 4 respectively).

[28] Our preliminary assessment of operability areas and interventions was added to with input gathered from a workshop held with NGESO in January 2023. Here, NGESO reiterated that in a majority of areas, the key question is around cost-effectiveness rather than ensuring operability per se.

3 The Role of Ancillary Services in Constraint Management

This chapter addresses the question: "What is the scope for addressing thermal constraint management through an ancillary service, as an alternative or as a supplement to the wider structural reforms being considered in the wholesale markets chapter of the REMA consultation, which could in theory help mitigate thermal constraints?"

As highlighted in Chapter 2, the gross[29] costs associated with management of thermal constraints may rise to ~£3bn a year by the middle of this decade[30], driven by a rise in connected renewable generation assets, often located at geographic network extremes. To tackle this problem, a 'whole systems' approach is needed that considers the role of multiple energy vectors, including collaboration between a wide range of system stakeholders (transmission and distribution network operators, generators, storage, and flexibility providers) and the ESO to deliver more efficient and coordinated outcomes for consumers in line with net zero requirements. Amongst other considerations, a whole systems approach in this report also includes the cost to the end consumer from balancing actions, taking into account the gross cost of turn-up and turn-down services as well as the consequential reduction in policy costs from these actions.

At a conceptual level, two key categories of interventions are used to manage thermal constraints:

1. *Increasing Network Capacity:* the process of increasing the capacity and reliability of the electricity network by building new infrastructure or upgrading pre-existing assets, thereby reducing congestion at transmission boundaries. Significant network reinforcement is needed for the UK to meet its net zero goals, a fact accepted by the government, Ofgem, and NGESO. Increasing network capacity can additionally refer to increasing the level of interconnection with other power systems, which can also assist with management of thermal constraints.

2. *Constraint Management:* thermal constraint management looks to effectively manage supply and demand to make best use of transmission and distribution infrastructure. This can be achieved by a number of activities such as curtailment of generation or activation of demand side response (DSR) assets. In GB, the majority of operational signals for constraint management are sent through the balancing mechanism operated by NGESO.

[29] The gross cost of constraint management is calculated as the sum of the cost of turn-down services and related turn-up services. When CfD generators are turned down and if CfD generators do not take advantage of the constraint to profit maximise, the net system turn-down cost for these actions is minimal.
[30] NGESO (2022) Modelled Constraint Costs

From a whole systems perspective, a combination of these approaches will be necessary in all NGESO Future Energy Scenarios (FES). While significant investment in new network capacity is vital, it would be sub-optimal to remove all constraints as, from a whole system cost perspective, this would require investment in costly transmission assets which would then be underutilised. Therefore, even with significant increase in network capacity, there will still be periods of constraint which need to be managed cost-effectively.

This chapter identifies the characteristics required across constraint management services to manage thermal constraints in a low carbon system. Additionally, the following chapter will assess the extent to which these characteristics are delivered through current and planned measures, before detailing the additional ancillary services that should be investigated to manage thermal constraint costs. However, this chapter does not quantify the system optimal combination of increasing network capacity, managing constraints through flexibility and storage, and curtailment as this would require detailed network modelling, which is beyond the scope of this project. However, NGESO do carry out this sort of analysis has part of their Network Options Assessment (NOA) work, which is referenced at various points throughout this chapter[31].

Specifically, this chapter is structured as follows:

- Chapter 3.1 provides a brief overview of options to increase network capacity. This is included for completeness and is not the main focus of analysis.
- Chapter 3.2 introduces the criteria used to assess current and proposed means of managing thermal constraints.
- Chapter 3.3 outlines the current means of constraint management, assessing them against the criteria to identifying gaps in current provisions. This builds on the analysis in chapter 2 (project TOR questions 1-3).
- Chapter 3.4 assesses the scope of proposed ancillary services to fill these gaps against the same criteria, highlighting where gaps still remain.
- Chapter 3.5 assesses the requirement for ancillary services in the context of the different REMA wholesale market scenarios, as outlined in Chapter 3.2.
- Finally, Chapter 3.6 summarises key findings and conclusions for this chapter.

[31] National Grid ESO (2023) Network Options Assessment (NOA)

3.1 Increasing the speed of network reinforcement reduces the need for ancillary services for constraint management

Network capacity can be increased through the construction of new assets, retrofitting existing assets using new materials to improve performance and/or monitoring assets in real-time to maximise use of existing capacity. The need for network reinforcement to reduce thermal constraints was a recurring theme in our stakeholder engagement and this chapter provides a brief overview of these options.

3.1.1 The scale of planned network reinforcement activities

Plans for investment have been published in NGESO's first Holistic Network Design (HND)[32] alongside a refreshed Network Options Assessment (NOA) (July 2022)[33]. Together, 94 asset investments (amounting to some £22bn) are recommend to deliver a network that can accommodate the Government's ambition of 50GW offshore wind by 2030. Longer-term, Ofgem are minded-to introduce a Centralised Strategic Network Plan (CSNP), which will be led by the Future System Operator (FSO)[34]. The CSNP will proactively identify, design and progress investments in the transmission network, considering onshore, offshore and cross-vector aspects.

Historically, significant new transmission assets have often taken a long time to progress through the planning process for several reasons, including public opposition and local environmental considerations. Whilst this is an issue the Government are looking to address, options to retrofit existing infrastructure is one of the potential means of improving transmission capacity more rapidly.

3.1.2 Improving power line emissivity could serve as a cost-effective means of reducing the thermal stress on transmission cables

When under thermal stress, power lines can sag leading to reduced efficiency, increased line losses, and potential leading to line fatigue and failure. One possible solution to this problem is to increase cable emissivity, which refers to a material's effectiveness in emitting energy as thermal radiation. By improving the emissivity of the cable, the cable is better able to radiate heat away from itself, which reduces the operating temperature of the cable and helps prevent sagging.

Cable emissivity can be improved by using materials that have high emissivity, such as aluminium or steel. Additionally, coatings or surface treatments can be applied to the cable to increase its emissivity. Prysmian Group, an organisation that specialises in power distribution

32 NGESO (2022) Holistic Network Design
33 NGESO (2022) Network Options Assessment (NOA)
34 NGESO (2022) Decision on the initial findings of our Electricity Transmission Network Planning Review

technology, claims that emissive cable coatings can result in an up to 30% reduction of line operating temperature and a 25% reduction in line losses[35].

3.1.3 Retrofitting composite cables could increase network capacity without needing to construct new pylons

Traditional power line cables are typically made of steel or aluminium conductors.

Instead of using a conventional steel core, composite cables use a core made of a strong and lightweight material, such as carbon fibre or fibreglass, that has high tensile strength and low thermal expansion. Composite cables have several advantages, including:

- *Higher ampacity:* Composite cables can carry more current than conventional cables without overheating or sagging, because the composite core has lower electrical resistance and higher thermal conductivity than steel. This can improve the efficiency and reliability of the power system and reduce the need for additional lines, pylons, or substations.
- *Lower losses:* Due to lower electrical resistance, composite cables can also reduce electrical losses, reducing overall generation requirements to meet a given demand.
- *Lower weight:* Composite cables are lighter than conventional cables because the composite core has lower density than steel. This can reduce the stress on towers and poles and enable longer spans and higher voltages. This can also lower the installation and maintenance costs and environmental impacts of the power system.
- *Higher durability:* Composite cables are more resistant to corrosion, fatigue, creep, and abrasion than conventional cables because the composite core is protected by a polymer matrix and aluminium wires. This can extend the service life and performance of the cable and reduce the risk of failures and outages.

Retrofitting composite cables with emissive coatings to existing pylons could serve as a relatively low-cost way of improving transmission capacity while avoiding some of the planning constraints of network reinforcement outlined above. Some case studies of where composite cables have been used with success include:

- In China, a 1,100 kV DC transmission line from Changji to Guquan used composite cables (Aluminium Conductor Composite Core) conductors to deliver 12 GW of power over 3,293 km.
- In India, a 400 kV AC transmission line from Gwalior to Agra used composite conductors to increase the power transfer capacity by 1.5 times without changing the towers or insulators.

[35] Prysmian Group (2023) E3X Technology

- In Brazil, a 230 kV AC transmission line from Bom Jardim to Campos used composite cable conductors to reduce electrical losses by 27% and increase the power transfer capacity by 60%.
- In Canada, a 138 kV AC transmission line from Fort McMurray to Fort McKay used composite cables to increase the power transfer capacity by 100% and reduce line losses by 32%.
- In the US, a 115 kV AC transmission line from San Diego Gas & Electric's (SDG&E) Sycamore Canyon substation to its Penasquitos substation used composite conductors to increase the power transfer capacity by 2.3 times and reduce line losses by 40%.

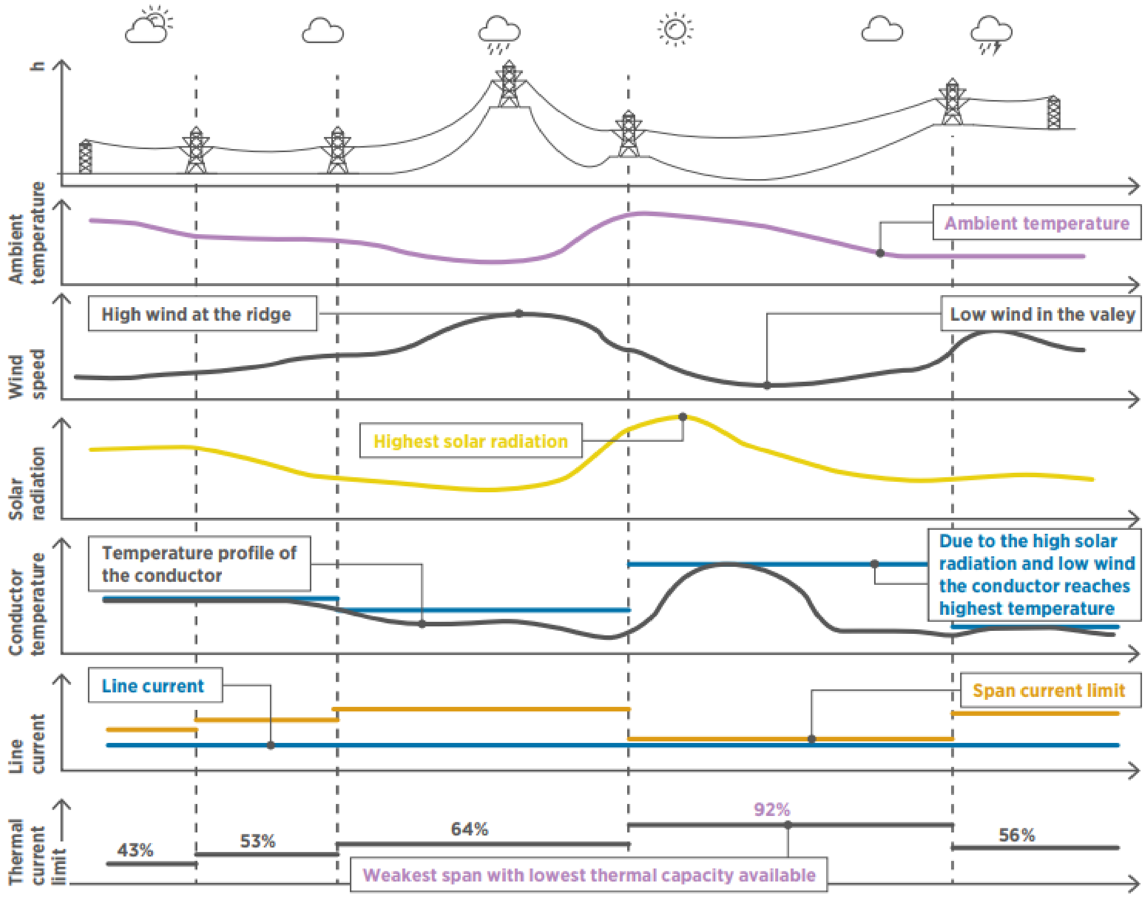
3.1.4 Dynamic line ratings could allow more power to be safely transported through existing network infrastructure

In the past, system operators have relied on static thermal ratings to determine the ampacity of transmission and distribution conductors. These ratings were based on theoretical extreme worst-case scenario weather conditions. However, the ampacity of a conductor is constantly changing and depends on various factors such as line current, insulation, wind speed/direction, solar radiation, and ambient temperature (as shown in Figure 4)[36].

With Dynamic Line Ratings (DLR), sensors are installed on the transmission lines to collect data on temperature, wind speed, solar radiation, and other environmental factors that can affect the performance of the line. These data are then used to calculate the real-time ampacity of the line, based on current conditions.

[36] IRENA (2020) Dynamic Line Rating: Innovation Landscape Brief

Figure 4 - The variability of influencing factors on DLR

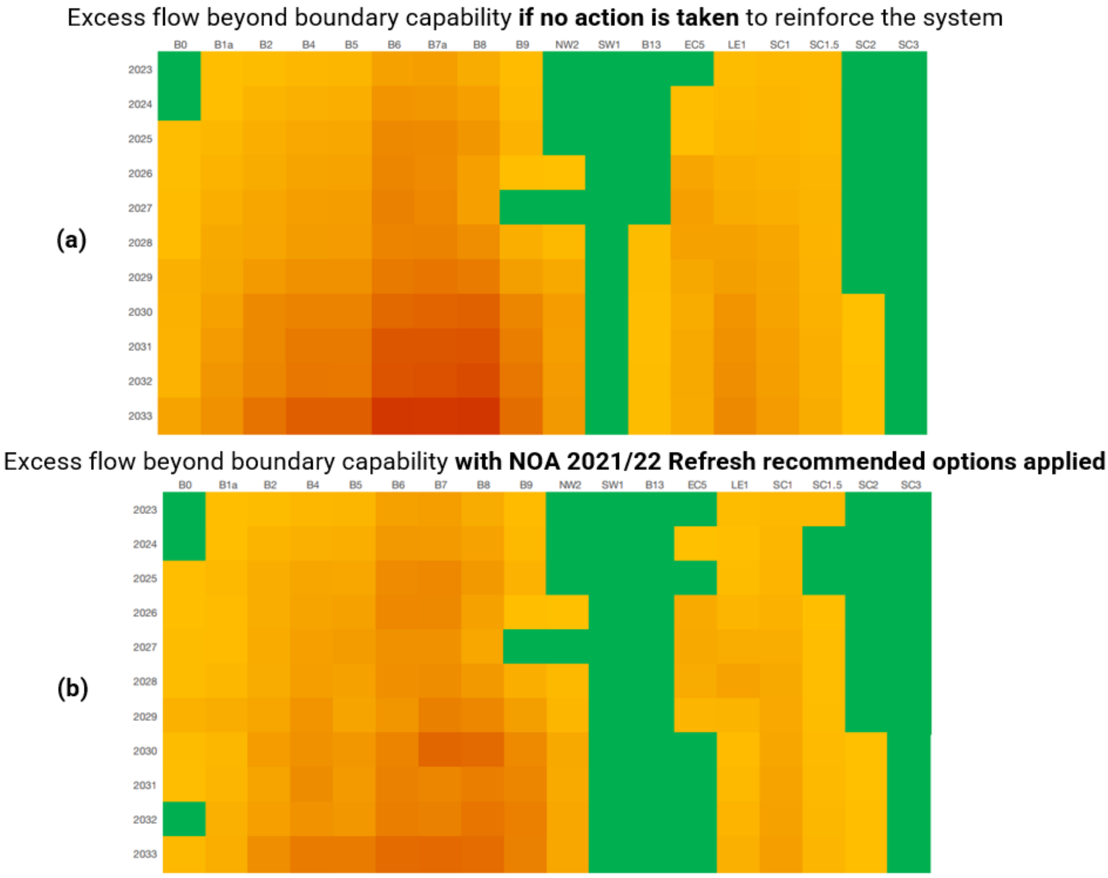


Source: IRENA, 2020

Using this approach, DLR allows for more efficient use of the transmission system, as the actual ampacity of the line can be monitored and adjusted based on real-time conditions, rather than relying on static ratings. This can help to reduce the likelihood of thermal overloads, which can cause power outages or damage to the transmission system, whilst maximising the use of existing transmission lines. There are a number of case studies which demonstrate the impact of DLR on power sector transformation. The implementation of DLR by Elia, the transmission system operator (TSO) in Belgium, and RTE, the French TSO, resulted in a 30% increase in the current their transmission lines could potentially carry[37]. The TWENTIES project was part of the EU FP7 research and technology programme and involved various stakeholders such as European TSOs, power technology and wind equipment manufacturers, and generators. This project found that the use of DLR forecasts led to an average increase in transmission capacity of approximately 10-15%[38].

[37] Elia Group (2022) Dynamic Line Ratings
[38] Pavlinic and Komen (2017) Direct monitoring methods of overhead line conductor temperature

Figure 5 - Heat map of constrained boundaries before (a) and after (b) NOA recommended options are applied



Source: NGESO, 2023. Note: The X-axis shows key transmission boundaries, the Y-axis shows time.

Overall, increasing network capacity can reduce constraints. However, even if all of NGESO's NOA recommendations were applied, there would still be significant constraints at congested transmission boundaries (Figure 5)[39]. Therefore, constraint management services will still be required. The following sections in this chapter assess how existing and proposed system services and market structures impact the management of constraints and identify the extent to which alternative or supplementary ancillary services could support this.

3.2 Assessment Criteria for Constraint Management Services

This section set outs the criteria against which constraint management services have been assessed in this report. The criteria are grouped into three sub-sections reflecting the Government's objective to deliver a reliable (operable), cost-effective and low carbon energy system.

[39] National Grid ESO (2023) Markets Roadmap

3.2.1 Reliability

- *Effectively manages short-duration constraints (<30 minutes):* Short-term (i.e., over several minutes) constraints may arise due to sudden increases in wind speeds, causing a sudden spike in renewable generation output. A gust of wind can cause a power spike in excess of 100MW in minutes, and effective measures need to be in place to deal with the constraints these sharp peaks cause at transmission boundaries Currently, system operators manage this by leaving capacity headroom at transmission boundaries, a mechanism that results in under-utilisation of available infrastructure.

  This criterion assesses the ability of a service/mechanism to maintain system reliability in response to short-duration constraints. It considers how the service/mechanism responds in real-time – not how the service might incentivise future network, generation, or demand investment decisions.

- *Effectively manages longer-duration constraints (>30 mins – days):* The median length of thermal constraints on the Scotland/England boundary is expected to be between 5-10 hrs in this decade, with a growing number of continuous +24-hr constraints. During these periods, conventional energy storage facilities will become fully charged and unable to absorb additional power, leading to prolonged curtailment actions. Longer periods of constraint could occur due to extended periods of high wind speeds but may also occur due to a significant transmission line fault event (resulting in reduced capacity across a boundary).

  This criterion assesses the ability of a service/mechanism to maintain system reliability in response to long-duration constraints. It considers how the service/mechanism responds in real-time – not how the service might incentivise future network, generation, or demand investment decisions.

- *Promotes efficient operation of assets:* During a constraint period, assets connected to the network can respond in ways which alleviate or exacerbate the constraint. This criterion assesses whether a service/mechanism provides a signal which promotes the efficient dispatch of assets (i.e., promotes system optimal behaviour). As above, this criterion is assessed based on how assets respond to a signal in real-time – not as an investment incentive.

3.2.2 Cost

- *Incentivises investment in location specific DSR/Storage:* Both long- and short-term thermal constraints will result in generation being curtailed if there is no option to store energy or increase demand within the constraint boundary. Flexible demand and energy storage can help reduce constraints by shifting demand or supply to different times of the day or year.

  This criterion assesses the extent to which a service/mechanism encourages investment in DSR and storage in locations which alleviate constraints.

- *Maximises the use of network assets:* Effective use of network assets is essential to operating the transmission system cost-effectively. This can be achieved through more accurate and timely operational signals that allow for more effective dispatch of assets, along with new technologies such as those outlined in Chapter 3.1, that allow for greater volumes of electricity to pass through the network while retaining safety and security standards. This criterion assesses the extent to which a service/mechanism enables asset utilisation to be maximised at a given point in time.
- *Delivers a service cost-effectively:* This criterion assesses the extent to which a service/mechanism is delivered cost-effectively at a given point in time.

3.2.3 Carbon

- *Promotes low carbon technologies:* This criterion assesses whether the proposed solution will enable greater use of low carbon generation that would otherwise be curtailed, reducing overall system carbon emissions.

3.3 Assessing the performance of current and planned services for constraint management

Thermal constraints are currently largely managed through the Balancing Mechanism (BM). In addition, as part of NGESO's 5-point plan for constraint management[40], two further approaches are currently under development:

- Constraint Management Intertrip Services: Intertrip schemes have been developed through NGESO's Constraint Management Pathfinder[41]. Two tenders have since been launched to address constraints in East Anglia (EC5 region) and between Scotland and England (B6 Boundary).
- Local Constraint Markets (LCM)[42] and MW Dispatch Service[43]: These schemes aim to provide flexibility from distribution-connected generation or demand, which is not participating in the BM. The LCM is currently being trialled by NGESO and Piclo to manage constraints across the England/Scotland border. The MW Dispatch Service is due to be launched this year in the southwest and southern regions of England.

This chapter describes each approach and assesses them against the criteria in Chapter 3.2 above.

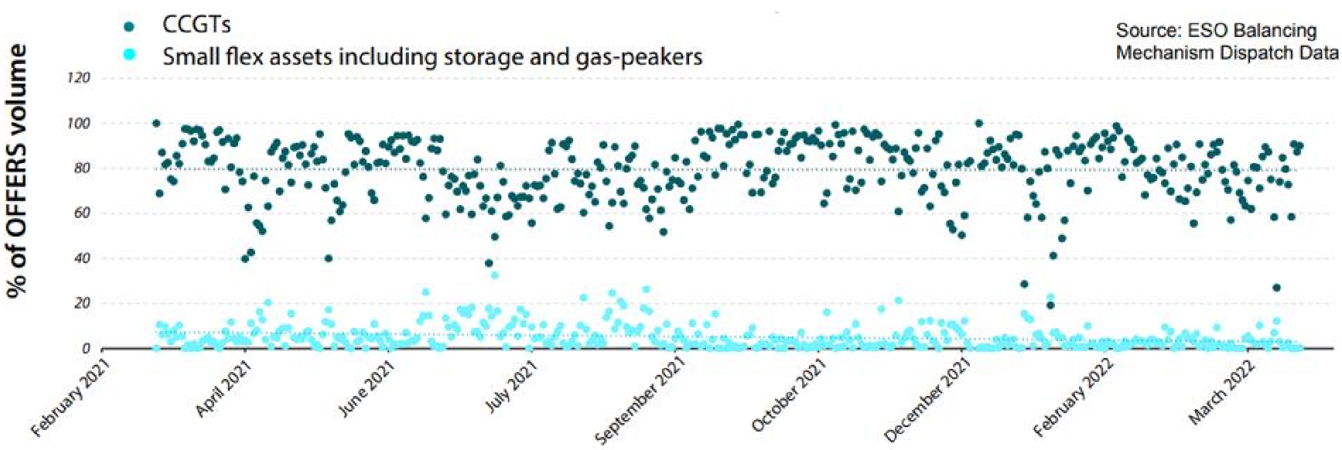
[40] NGESO (2021) 5-point plan to manage constraints on the system
[41] NGESO (2023) NOA Constraint Management Pathfinder
[42] NGESO (2023) Local Constraint Market
[43] NGESO (2023) Markets Roadmap

3.3.1 Thermal constraint management through the BM can be costly

The BM is the ESO's main means of balancing supply and demand across the energy system. All units in the BM (BMUs[44]) can provide bids (to decrease generation or increase consumption) and offers (to increase generation or decrease consumption), which can be called on by the ESO through Bid Offer Acceptances (BOAs) to balance supply and demand after gate closure.

Generation constraint management in the BM typically involves two distinct actions: (1) The 'turn-down' (curtailment) of a generation asset such as a solar or wind farm on one side of a constraint; and (2) the 'turn-up' of a generation asset, which is currently almost always a carbon-intensive Combined Cycle Gas Turbine (CCGT), to meet demand on the other. From April 2021 to March 2022, the costs associated with turn-up actions represented around 85% of the gross costs associated with thermal constraint management, with 94% of payments being made to large CCGT plants[45] (Figure 6)[46]. This is largely due to the price disparity between turn-up and turn-down services driven by high gas prices due to the war in Ukraine[47].

Figure 6 - Daily percentage of 'turn-up' electricity supplied March 2021 – April 2022



Source: Regen, 2022

[44] A BMU accounts for a collection of plant or apparatus, such as a generator or a consumer, that can be independently controlled and metered. Each BMU has a unique identifier and is registered with Elexon, the administrator of the Balancing and Settlement Code (BSC). There are different types of BMUs, depending on their connection and function.
[45] Drax (2022) Renewable curtailment report
[46] Regen (2022) Seven solutions to the rising cost of transmission network constraint management
[47] The average price of natural gas in Europe was $18.5 per million British thermal units (MMBtu) in 2022, compared with $3.9 per MMBtu in 2021. This represents a 374% increase in gas prices in one year.

3.3.2 Constraint management Intertrip Services allow the ESO to safely increase the power flows through constrained boundaries

The Constraint Management Intertrip Service (CMIS) employs a fibre network to enable the ESO to rapidly disconnect generation from linked assets (wind and storage) within 150 milliseconds during a significant system fault that may cause thermal thresholds to be exceeded. This provision enables the ESO to increase the volume of electricity transmitted through a constraint, reducing the costs associated with asset curtailment.

While the service will not become officially operational until October 2023, NGESO allowed six of the 15 generating units that were successfully awarded contracts for the B6 boundary (England-Scotland border) service to begin operation early, a measure NGESO state has saved consumers £80 million between April 2022 and January 2023[48]. NGESO is currently attempting to introduce a comparable contract for East Anglia (EC5).

3.3.3 The local constraint market and MW dispatch service are two markets developed by NGESO as part of its 5-point constraint management plan

Similar to the CMIS developed through NGESO’s Constraint Management Pathfinder, the local constraint market (LCM)[49] and MW Dispatch Service[50] constitute a part of ESO's 5-point constraint management plan (as part of the Regional Development Programmes). These novel schemes aim to reduce constraint management costs through ‘increased competition from new assets who currently have challenges with accessing the BM[51] such as distribution-connected generation or (for LCM only) demand.

The LCM will enable ESO to utilise more non-BM assets (both generation and demand) north of the B4-B6 boundary (Scotland to England) to tackle constraints, with distribution-connected wind comprising the majority of the capacity accessed in this market, effectively providing ESO with additional cost-effective curtailment alternatives. However, due to the low levels of available flexible demand in Scotland, the level of unlocked demand flexibility is likely to be low in the short-term.

The MW Dispatch Service will launch in the Southwest (NGED’s Distribution Network) and South Coast (UKPN) regions in late summer 2023. This will initially focus on generation turn-down only. Both the LCM and MW Dispatch Service will operate as pre-fault services, with the ESO paying providers a utilisation fee that they bid ahead of time. There are some differences in how assets are dispatched - the LCM will operate ahead in 30-min blocks (similar to the BM), while the MW Dispatch Service will constrain assets for the duration required by the system. Assets in the MW Dispatch Service must also be able to respond within two (2) minutes.

[48] Current News (2023) National Grid ESO uses constraint management to save consumers £80 million
[49] NGESO (2023) Local Constraint Market
[50] NGESO (2023) Markets Roadmap
[51] NGESO (2022) Local Constraint Market - Product and Service Design

3.3.4 Current approaches deliver a reliable network, but do not provide locational operational or investment signals to support thermal constraint management

Table 1 below assess NGESO's three existing and planned options for managing thermal constraints against the criteria set out in Chapter 3.2 using a RAG (Red, Amber, Green) rating where:

- *Red:* Does not deliver against the criterion
- *Amber:* Partially delivers against the criterion or has no impact
- *Green:* Delivers against the criterion

Overall, current measures allow NGESO to operate the network reliably. The introduction of the CMIS, Local Constraint Market and MW Dispatch Services enable NGESO to make better use of existing assets (both increasing utilisation of network capacity and making use of embedded generation and demand to manage constraints more cost-effectively).

Nevertheless, the current approach to thermal constraint management still relies heavily on curtailing renewable generation behind a constraint and on turning up CCGTs in front of a constraint. If GB is to deliver a net zero energy system by 2035, there is a need to invest in low carbon storage and flexibility (including dispatchable generation) in specific locations to manage thermal constraints. This could include short- and long-duration storage assets, and green hydrogen production (electrolysis) facilities in areas with frequent 'excess' generation. In addition, it is necessary to provide location-specific operational signals ahead of time to assets to encourage them to act in system-beneficial ways (e.g., storage charging in regions with excess generation and discharging in regions with limited generation). Finally, it is important to ensure that any operational decisions made in relation to thermal constraint management are transparent. Whilst not directly related to the assessment criteria, several stakeholders engaged for this project stated their belief that CCGTs are over-represented in the BM, which they think is in part due to inadequate digital infrastructure in NGESO's control room.

Since 2020, NGESO has opened up the BM to smaller assets, including battery energy storage systems (BESS). Despite this increased market access, owners of smaller assets feel like they are dispatched less than they would expect, with the pressures (limited time, resources and automation) in NGESO's control room assumed to favour larger assets due to their ease of dispatch when compared to the challenges of aggregating many smaller and diverse technologies. However, this assertion is challenged by NGESO, who argue that larger assets are often dispatched for other reasons such as assuring effective system operability (e.g., inertia or reactive power requirements). The need for investment and operational signals in low carbon flexibility and storage are also relevant to addressing capacity adequacy and will be revisited in Chapter 4 and the conclusion to this report. The following chapter examines proposed ancillary services for thermal constraint management.

Table 1 - Red, Amber, Green (RAG) analysis of current and planned ancillary services for constraint management. Services are assessed against the qualification criteria outlined in Chapter 3.2

	Reliability			Cost			Carbon
Current/planned services	Effectively manages short duration constraints	Effectively manages longer duration constraints	Promotes efficient dispatch of assets	Incentivises locational investment in DSR/storage	Maximises use of existing network assets	Cost Effectiveness	Promotes low carbon technologies
Balancing Mechanism (BM)	Although the BM manages thermal constraints in close to real time, it is not responsive enough to deal with very sudden spikes in generation or faults, relying on other system services and retained network headroom to ensure reliability	The BM manages prolonged constraint periods by turning down generation behind constraints, and turning up generation in front of them.	The BM re-dispatches assets post Gate Closure. It is a means of correcting inefficient dispatch, rather than incentivising assets to dispatch in accordance with network constraints.	The BM is not an efficient *locational* investment signal for DSR and Storage due to limited asset visibility at the distribution level, and a lack of locational information for aggregated assets. In isolation, the BM is unlikely to provide sufficient revenue certainty for investment decisions.	While recent reforms have allowed smaller assets to connect to the BM, there is limited locational information about many aggregated assets, and sometimes poor visibility of asset status. Some stakeholders felt that this resulted in low-carbon flexibility being overlooked in favour of larger fossil fuelled generators.	Limited asset visibility and a short-term focus (balancing in 30 minute segments) means the ESO may not select the most cost-effective means of managing constraints. In addition, the complexity of the BM and its interaction with the wider wholesale market and subsidy regimes has led to accusations of participants profiteering.	The service is technically technology & carbon agnostic. In practice, the vast majority of "turn up" payments are made to large CCGT plants. The service does not specifically promote low carbon technologies.
Constraint Management Intertrip Services (CMIS)	CMIS is good at addressing short term constraints, acting fast in post-fault periods to stop thermal breaches at specific boundaries.	CMIS is only set up to address post-fault periods until a fault is resolved. While these periods may be extended, there is little action taken after the initial intertrip signal.	The CMIS does not promote efficient pre-fault dispatch.	The CMIS is not an investment signal for new capacity.	The CMIS allows boundary capacity to be used as much as possible in post fault conditions, without exceeding thermal limits.	A core rationale behind the CMIS is a reduction in the need for new-build solutions. NGESO state the CMIS has already saved consumers £80 million between April 2022 and January 2023.	The CMIS is technology agnostic, but would allow for renewables to continue exporting power in post fault conditions. Most connected technologies are expected to be wind.
Local Constraint Market (LCM) / MW Dispatch Service	The LCM will operate ahead of gate closure and real time. Signals are day-ahead and within day. Post-gate closure balancing will still be achieved through the BM. However, it is not responsive enough to deal with sudden spikes in generation or faults. Assets under the MW Dispatch Service will need to be able to curtail generation within 2 mins.	The LCM and MW Dispatch Service share many similarities in design to the BM, but with distribution level assets unable to enter the BM.	Like the BM, the LCM and MW Dispatch Service do not address pre-fault dispatch.	The LCM and MW Dispatch Service do incentivise flexible assets in specific regions. The ability to stack this service with others will determine its impact on investment decisions.	The LCM and MW Dispatch Service unlocks access to new non-BM assets on the distribution network (mostly onshore wind) to resolve constraints.	The objective of the LCM is to reduce the annual B6 boundary cost through increased competition from new assets who currently have challenges accessing the BM. The MW Dispatch Service will provide a similar service in the South West and on the Southern Coast.	Whilst the LCM and MW Dispatch Service are technology and carbon agnostic, having better visibility of local assets may enable DSR and storage to play a greater role in managing constraints than curtailing renewable generation.

3.4 Supplementary or Alternative Ancillary Services

As outlined in Chapter 3.3, the current ancillary services have gaps in assuring cost-effective management of thermal constraints, particularly regarding incentivising investment in flexible assets and in incentivising system-beneficial operation of these assets.

Through our stakeholder engagement, the Carbon Trust has asked for views on alternative approaches to thermal constraint management. Zenobē have proposed two ancillary services which it feels address some of these gaps These are expanded on in its recent storage for constraint management report. This report summarises each service and assesses them against the criteria in Chapter 3.2. It should be noted that the ancillary services analysed in this section are examples, not an exhaustive list of all possible interventions.

3.4.1 A strategic cycling service could help cost-optimise curtailment actions

As noted in Chapter 3.2, the frequency of constraints lasting over 24 hours is set to increase. In these periods, it is likely that storage assets may become fully charged and unable to absorb additional demand, leading to curtailment of generation. Zenobē suggest there is an opportunity for NGESO to 'strategically cycle' storage behind constraints, which could help decrease the total costs associated with constraint management.

During a continuous 24-hour constraint, the level of excess wind generation varies, with peaks and troughs that can vary by gigawatts. The depth of the constraint increases with higher peaks of excess wind generation, resulting in greater wind curtailment. The most expensive curtailment action taken in any settlement period to alleviate the boundary constraint is known as the 'marginal cost of curtailment'. During a deep constraint, NGESO has to curtail increasingly high-gross-cost wind generation, such as wind farms in receipt of Renewable Obligation Certificates (ROCs) or with a high strike price Contract for Difference (CfD)[52]. In contrast, during shallow constraints, NGESO only needs to curtail the lowest-gross-cost wind generation, typically low strike price CfD contracts or 'merchant' wind with no support.

By strategically using storage to import power when the marginal cost of curtailment is highest and then exporting when it's lower, NGESO could reduce gross curtailment costs.

For example, during a multi-day constraint, the marginal cost of curtailment is assumed to vary between -£70/MWh (RO wind) and approximately £0-50/MWh (merchant/low CfD strike price wind). NGESO could use battery storage to import during the constraint, avoiding expensive wind curtailment (-£70/MWh in the BM), and then fully export a few hours later at £0/MWh, fully discharging the battery to allow it to import during the next costly period. This process could be

[52] It should be noted that CfDs expecting a high difference payment in any given hour would be expected to bid into the BM at around the difference payment value plus a small margin. As such, any cost (or payment, if market prices are above the strike price) is largely offset by a change in difference payments from the Low Carbon Contracts Company. Therefore, assuming that CfD generators are not profiting excessively from thermal constraints management in contravention of Generation Licence 20A, the net turn-down cost of CfDs on a system level is minimal.

repeated, enabling battery storage to cycle between peaks and troughs and ultimately helping NGESO reduce gross constraint cost. However, this would require a specific, pre-fault constraint management service that allows significant collaboration between NGESO and the battery operator.

While this service would reduce the gross cost of constraint, the net cost (assuming CfD generators bid their short-run marginal cost (SRMC) – see Chapter 5 for further details) may be small. If this service were therefore only used to optimise which CfDs are curtailed, this would have limited impact on overall customer costs. If this service were to lead to a shift from CfD curtailment to additional RO curtailment, net cost to consumers would rise.

However, as some CfD generators have not been bidding their SRMC in periods of thermal constraints, adding to end customer costs, a service such as this could serve as a 'disciplining tool' by giving NGESO additional options in constraint situations, increasing competition between providers of flexibility, and limiting gross curtailment costs for end customers.

3.4.2 A capacity 'shock absorber' reserve service could allow for improved management of short-term spikes in renewable output

As noted to in Chapter 3.2, during periods of thermal constraints, NGESO does not utilise the maximum boundary transfer capabilities to maintain reserve capacity or headroom on the network. This is necessary to prevent sudden increases in power flows across transmission boundaries caused by strong gusts of wind that could surpass safe limits. NGESO's Frequency Response ancillary services can act fast enough to address these issues, however such services do not contain a locational requirement to position demand assets behind constraint boundaries.

Zenobē states that through ongoing conversations with NGESO control room operators, it has concluded that battery storage assets may have a unique role to play as 'short duration constraint reserve'. These battery assets would act as shock absorbers, importing power during wind gusts and exporting power during periods of calm. This service, Zenobē argues, would significantly ease the task of balancing the system for NGESO during constraints and enable ESO to optimise the utilisation of existing network capacity.

Table 1 assesses both of these options against the criteria set out in Chapter 3.2. As demonstrated by Table 2, both a Strategic Cycling Service and a Short-Term Constraint Reserve Service (STCRS) 'shock absorber' deliver against a number of the assessment criteria outlined in Chapter 3.2.

Subject to further investigation, if implemented alongside the current and planned measures assessed in Chapter 3.3, the Strategic Cycling Service could provide an additional means of influencing asset behaviour to reduce the cost of curtailment, and the STCRS could complement the CMIS to reduce the extent to which renewable generation is curtailed. However, even when implemented alongside existing measures, these additional ancillary services do not significantly encourage location-specific investment in assets which will reduce constraints, nor encourage all system assets to operate in a way which benefits the system.

As a result, there is still a gap in regard to creating the investment and operational signals needed to incentivise storage and flexibility assets.

There are currently no clear investment signals for low carbon assets, particularly longer-duration storage technologies (including hydrogen) that could be used to absorb power needs in periods of extended constraint. An additional service to ameliorate the operational signal gap is discussed towards the end of this chapter.

The next chapter examines the impact a change in wholesale market structure could have on thermal constraint management.

Table 2 - Red, Amber, Green (RAG) analysis of ancillary services for constraint management proposed by Zenobē. Services are assessed against the qualification criteria outlined in Chapter 3.2

Supplementary or alternative services		Reliability		Cost			Carbon
	Effectively manages short duration constraints	**Effectively manages longer duration constraints**	**Promotes efficient dispatch of assets**	**Incentivises locational investment in DSR/storage**	**Maximises use of existing network assets**	**Cost Effectiveness**	**Promotes low carbon technologies**
Strategic Cycling Service	Strategic Cycling is focused on the peaks and troughs of long, continuous constraint periods. It is not intended to address short duration constraints	This service would support the management of longer duration constraints. Whilst it may not significantly reduce the quantity of renewable generation constrained overall, the costs associated with constraint management during longer duration constraints periods can be significantly reduced.	The service would encourage efficient dispatch of assets contracted to deliver under the service. However, it would not impact wider price signals which may lead to inefficient behaviour by other generation and storage assets.	This service would only be available to storage assets, potentially favouring shorter durations assets such as batteries. The nature of the service means that assets would need to be situated in specific locations to manage constraints. This service could form part of an investment case for storage/flexibility assets (providing it could be stacked with other services)	While this service would help optimise curtailment of generation assets, it does not impact the extent to which transmission assets are utilised.	This service would help reduce the marginal cost of curtailment during long duration curtailment events. However the impact on net costs would likley be minimal.	The Strategic Cycling Service is technology agnostic regarding what generation assets are curtailed. Under current circumstances, this service is likely to change which renewable assets are curtailed, rather than having a significant impact on the total amount of curtailed generation.
Short Term Constraint Reserve Service (STCRS)	The STCRS is specifically aimed at wind gusts that cause rapid (within minutes) 100MW+ generation spikes.	The STCRS is not aimed at long term constraint management.	The STCRS does not promote efficient pre-fault dispatch.		The STCRS allows boundary capacity to be used to a greater extent and acts as a 'shock absorber' when generation exceeds network capacity	This service would help maximise the use of network capacity, reducing the level of reinforcement required to transmit a given MW capacity.	The STCRS is technology agnostic, but would allow for some renewable generators to continue exporting power during short duration conditions. Most connected technologies are expected to be wind.

3.5 The impact of REMA wholesale market options on the role of ancillary services for thermal constraint management

The wholesale market options set out in DESNZ's REMA Consultation (July 2022) provide different operational signals to assets and could have a significant impact on the role of ancillary services for constraint management. This chapter provides a summary of those options and their impact on thermal constraint management before assessing them against the criteria from Chapter 3.2.

3.5.1 National Wholesale Pricing

Under current wholesale market arrangements, energy is traded bilaterally in national markets. Due to financially firm access (the connect and manage approach), there is no incentive for market participants to take into consideration the impact of their trades on thermal constraints. Generators and suppliers/demand customers provide NGESO with their final physical notifications (the amount of energy they expect to consume/generate within a half-hour period) at gate closure. NGESO then uses the BM to redispatch generation according to the thermal limitations of the network.

3.5.2 Locational Wholesale Pricing

Locational pricing is not a new concept on GB's energy network, with locational elements being present in use of system charges such as Transmission Network Use of System (TNUoS) charges for many years. In REMA, two options for locational wholesale pricing are suggested:

- *Zonal Pricing:* where the transmission system is split into multiple zones, (typically in the range of 2-15) and is usually demarcated along congested boundaries. The wholesale price of electricity differs between these zones, but within zones, prices remain uniform.
- *Nodal Pricing:* also known as locational marginal pricing (LMP), is the most granular form of pricing setting. The transmission network is divided into, potentially, hundreds of 'nodes', each of which has an individual and separate wholesale electricity price. While complex, nodal pricing more precisely reflects the variability of value of electricity due to location across the country.

For all three options (national, zonal, and nodal pricing), assets can be dispatched centrally by NGESO (central dispatch) or by the generators themselves (self-dispatch). The current National Pricing model uses self-dispatch whilst nodal pricing is typically paired with central dispatch due to the complexity of balancing the network over a large number of nodes. Zonal pricing sits between national pricing and nodal pricing in terms of its complexity and could adopt either approach. Table 3 summarises the impact of these options on thermal constraint management only. This report then highlights the wider points raised by stakeholders on the implementation and operation of these options which could impact the transition to net zero.

For the purposes of this analysis, we have assumed that NGESO continue to operate the system in 30-min segments, with generators and suppliers/large demand notifying NGESO of their position at gate closure.

Table 3 indicates that a move to zonal or nodal pricing could deliver several benefits relating to operational management of thermal constraints compared to the status quo. However, a change to the wholesale price market structure has a wider impact on the investment into, and operation of, the energy system. This was a point most of the stakeholders interviewed made. Further details on those potential impacts are set out below.

Table 3 - Red, Amber, Green (RAG) analysis of different REMA wholesale market scenarios. Services are assessed against the qualification criteria outlined in Chapter 3.2

<table>
<tr><th rowspan="2">REMA wholesale market options</th><th colspan="4">Reliability</th><th colspan="2">Cost</th><th>Carbon</th></tr>
<tr><th>Addresses short term constraints</th><th>Addresses long term constraints</th><th>Promotes efficient dispatch of assets</th><th>Incentivises locational DSR/storage</th><th>Maximises use of existing network assets</th><th>Cost Effectiveness</th><th>Promotes low carbon technologies</th></tr>
<tr><td>National pricing with self dispatch (BAU)</td><td colspan="2">A single national price provides no locational constraint management signals</td><td>Current national wholesale energy prices do not factor in the cost of thermal constraint management - this is achieved through other means such as the BM.</td><td>A national price can signal overall supply scarcity, but does not provide locational signals for assets which could alleviate thermal constraints</td><td>National pricing does not help to alleviate thermal constraints, nor maximise use of existing assets</td><td rowspan="4">Cost effectiveness is a function of:

- Prices paid for energy in the wholesale market
- Prices paid for bid/offers in the balancing mechanism (net of policy cost impacts)
- Costs involved in operating and participating in the wholesale market and balancing mechanism
- Impact of wholesale market structure on wider costs (e.g. investment WACC)

While this report comments on each of these, it is beyond the scope to assess which market structure will result in the lowest cost for consumers.</td><td rowspan="4">All pricing structures are technology agnostic. However the implementation and operation of different structures can impact investment in new assets, and it how different asset types can respond to signals.</td></tr>
<tr><td>Zonal pricing with self-dispatch</td><td>A zonal price for a half hour period will not reflect the timing or duration of very short-term constraints. These will continue to be managed by the ESO</td><td rowspan="2">Zonal prices provide clear signals for constraints management across congested borders. If there are material constraints within a zone which limit the extent to which generators/suppliers can adapt their behaviour to manage constraints, a new zone would need to be formed.</td><td>Zonal pricing would provide better operational signals than national pricing, improving efficiency of dispatch and materially reducing need for BM re-dispatch.</td><td rowspan="2">Zonal pricing provides a limited locational signal, but does not reflect where within a region assets could be best placed. Long term visibility of generation build out and network reinforcement, which will change constraints over time, are critical to inform investment decisions.</td><td>Zonal pricing provides more granular signals than national pricing and related IT changes and improvements should give the ESO better visibility of connected assets.</td></tr>
<tr><td>Zonal pricing with central dispatch</td><td rowspan="2">Similar to the row above, the zonal signal does not provide a sufficiently granular (temporal) signal to indicate short duration constraints. However, central dispatch may make it easier for NGESO to dispatch assets in response to short term constraints</td><td>Zonal pricing provide better signals than national pricing. Full control allows the ESO to dispatch assets according to system needs.</td><td rowspan="2">A central dispatch algorithm enabled by locational price data would allow for better use of available infrastructure by giving the ESO better control and visibility of connected assets.</td></tr>
<tr><td>Nodal pricing with central dispatch</td><td>Nodal pricing with central dispatch would provide accurate locational signals incentivising generation and demand to operate in ways which alleviate constraints.</td><td>Nodal pricing would give granular operational signals to feed into the central dispatch algorithm, allowing the ESO to unilaterally dispatch assets according to system needs.</td><td>Nodal pricing would provide the most granular locational signal, but also considerable uncertainty for investors due to local circumstances potentially chancing significantly when system is reinforced or other demand/ generation connects nearby.</td></tr>
</table>

3.5.3 The impact of nodal pricing on thermal constraint management

To date, analysis on the impact of nodal pricing in GB has been undertaken by NGESO[53], Octopus Energy[54], Frontier Economics[55] and Ofgem[56]. Some stakeholders, including NGESO in their SOF[57], have indicated that a move to nodal pricing is a possible means for cost-optimising the level of investment needed to address the problems associated with management of thermal constraints for two key reasons:

1. Nodal pricing with central dispatch facilitates co-optimisation of energy and ancillary services with the physical limits of the network while, sending accurate signals in operational timeframes for efficient dispatch of energy, rewarding customers and generators who react to highly granular market prices.

2. The locational price signals created by nodal pricing should incentivise generation to move closer to areas of high demand, and flexible demand to move closer to areas with plentiful supply, reducing the amount of network capacity that's needed to facilitate transmission and distribution of energy generated and assure cost-effective operability.

Nodal pricing would provide more accurate and granular price signals that reflect the locational value of energy, including the cost of losses and congestion on the network. This could allow assets to be dispatched in a more efficient manner without the need for BM interventions. However, the majority of stakeholders expressed scepticism as to the extent to which the location of most generation and demand assets would change in response to locational price signals and felt that the uncertainty caused by nodal (or zonal) pricing (both during its implementation and operational) would have negative consequences on investment certainty that outweighed any benefit.

Specifically, concerns were raised about:

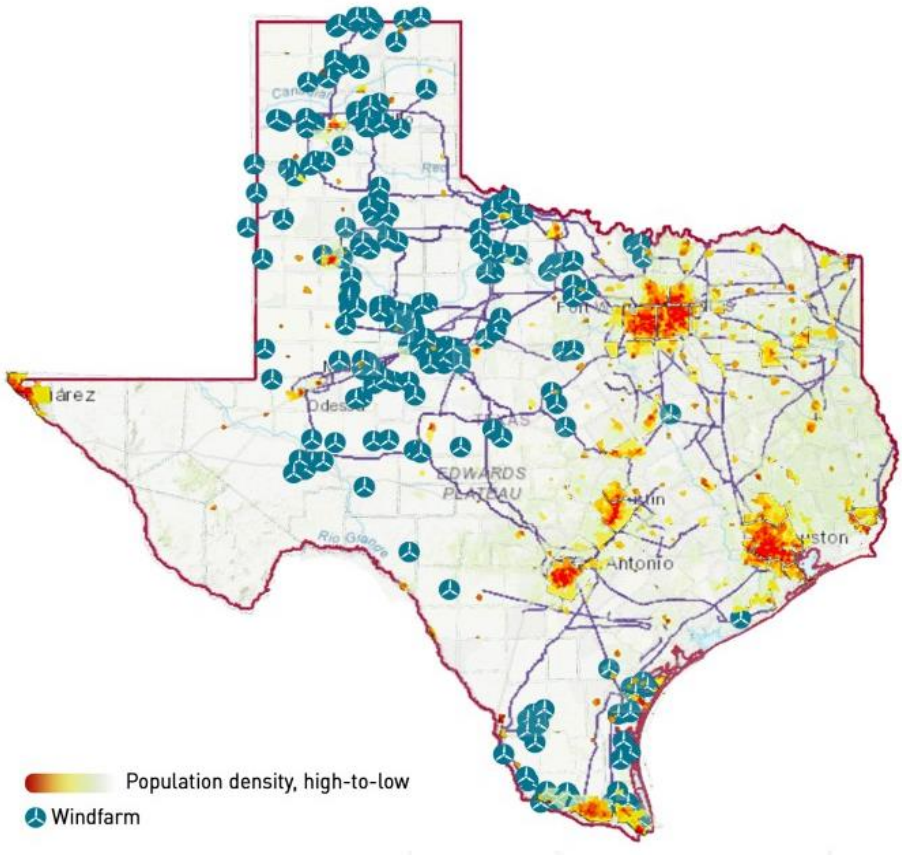
- *Limited impact of locational investment signals on generation and demand assets*. While zonal or nodal pricing could incentivise smaller fuelled generators, storage or demand assets to locate in response to thermal constraints, it is unlikely to have a significant impact on the location most low carbon generation. These assets must balance several constraints and considerations including:
- *Ability to obtain planning permission*, which is heavily dependent on environmental factors as well as local policy.
- *Resource availability*, particularly for wind and solar. The most optimum sites are typically located at network extremities, far away from demand.

[53] NGESO (2023) Net Zero market reform
[54] Octopus Energy (2022) GB Locational Pricing
[55] Frontier Economics (2022) An Assessment Framework for a move to LMP
[56] Ofgem (2022) Locational Pricing Assessment
[57] NGESO (2023) System Operability Framework (SOF)

- *Ability of demand to move*, as energy-intensive demand (industry) is typically co-located with logistics routes and available labour, affecting available supply chains.

In support of the above, studies of international markets, such as in the USA and New Zealand, which have introduced nodal pricing, have confirmed that natural resources play a critical role in determining the siting of generation facilities. For example, onshore wind facilities in Texas (Figure 7) have been situated in areas with high wind load factors, despite nodal pricing signals incentivising siting closer to demand in the east[58]. Similarly, in New Zealand, geothermal generation build-out has increased in regions with high resource potential, such as the Taupō Volcanic Zone, over the past two decades – rather than areas of high demand[59]. It is important, however, to recognise the international examples given above will inevitably have many varying aspects that do not fit neatly into the context and configuration of the GB network.

Figure 7 - Location of windfarms and population density in Texas



Source: Energy Information Administration, US Census data, 2020. Note: wind capacity in the Northwest is mainly located away from demand centres in the East

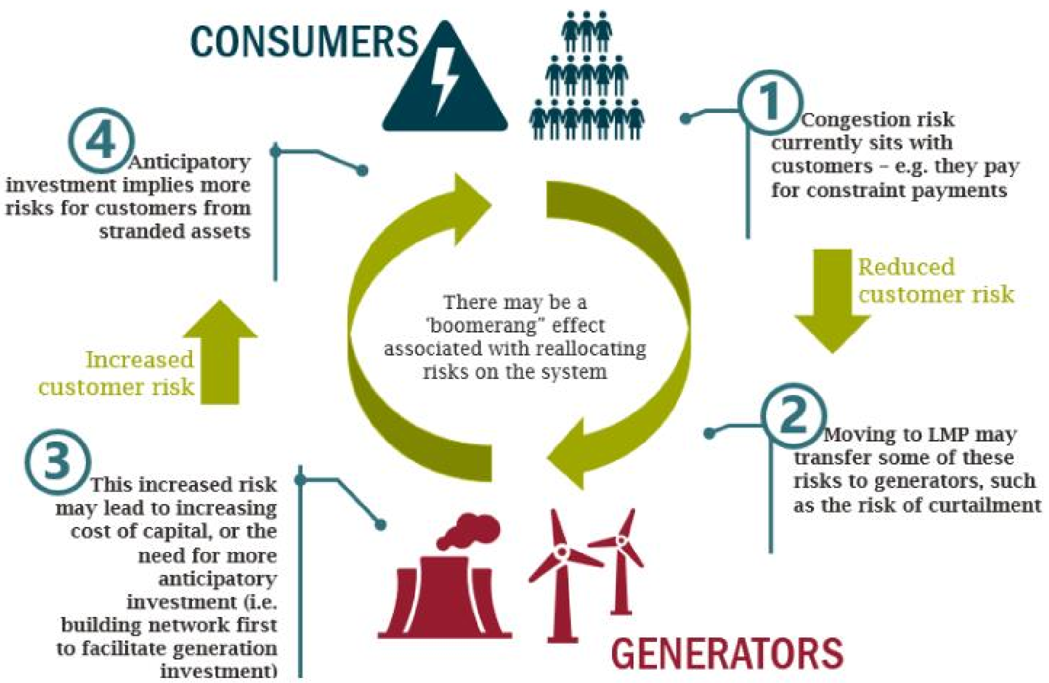
[58] Regen (2022) Wild Texas Wind: Insight Paper on Locational Marginal Pricing
[59] Ministry of Business, Innovation and Employment (2021) Energy in New Zealand

Impact of moving revenue risk from thermal constraints onto generators

Presently, investors in generation and flexible assets are shielded from the risk of network-related curtailment during operation. If the power grid is unable to accommodate their electricity output, they will receive compensation through the BM for any reduction of output. This cost is currently socialised through network charges (ultimately paid by the consumer). A move to nodal pricing would shift the risk of non-availability of transmission onto generators.

Investors interviewed for this project have stated that this uncertainty would increase the cost of capital for new assets, or potentially make it difficult for assets to secure funding at all. This could increase the overall cost of delivering a net zero energy system – a cost that ultimately is placed on consumers through higher energy prices. This 'boomerang effect' has been mapped by Frontier Economics[60] (Figure 8). In other words, while an implementation of nodal pricing would reduce the day-to-day costs associated with thermal constraint management for NGESO and potentially somewhat reduce constraint costs viewed from a system level (i.e., including netting off reduced difference payments from curtailment of CfDs), it could ultimately result in a more expensive net zero energy system for consumers.

Figure 8 - Risk reallocation 'boomerang' effect



Source: Frontier Economics, 2022

[60] Frontier Economics (2022) An Assessment Framework for a move to LMP

Investment in digital infrastructure is necessary to deliver operability improvements

Several stakeholders noted that, regardless of wholesale market price structure, investment is needed in digital infrastructure to enable market players to have better visibility of market signals, and for the NGESO to have better visibility of the status of system assets. Additionally, increased automation is needed to enable NGESO to respond to system needs quickly and cost-effectively as the network decarbonises.

Improvements to digital infrastructure are part of NGESO's Balancing Programme. Some aspects have already been introduced (such as the Power Available signal which provides real time information to NGESO on the potential generation from individual wind farms). Similar visibility of storage and flexibility assets (e.g., state of charge) has been identified as future requirement.

3.5.4 A move to central dispatch could reduce the need for constraint management services, but much of the associated costs would remain

Central dispatch is a model where the system operator determines the optimal generation and demand schedule based on bids and offers from market participants, network conditions and operational security constraints. A move to central dispatch could help with management of thermal constraints on the energy system by allowing the system operator to coordinate the output of different generators and demand-side resources in order to optimise the network flows and avoid overloading any equipment.

Central dispatch would require a significant increase in asset visibility for NGESO to ensure efficient dispatch. This might, therefore, also enable more efficient use of flexibility services from distributed energy resources (DERs) that are connected to the distribution network, such as storage or demand response, though this would depend on the degree to which the visibility of these assets extends beyond the transmission network and into the distribution networks. Central dispatch could also reduce the uncertainty and risk associated with variable generation sources such as wind, by using probabilistic methods to determine the optimal reserve requirements.

Nodal pricing is typically paired with central dispatch, as the granular signals created by LMP are essential data required for the central dispatch algorithms to be effective. If the correct technical and digital infrastructure were in place to allow effective central dispatch, the role of balancing and ancillary services to address thermal constraints would be greatly reduced, as NGESO would have unilateral control over system assets. However, it is probable that some aspect of real-time balancing would still be needed to resolve forecasting errors or intervene in specific localities where constraints arise. In US energy systems such as Texas that have adopted central dispatch, system operators run what they refer to as Real-Time Markets (RTM), a version of which would likely be needed in GB.

In self-dispatch scenarios, including those with national (BAU) and zonal pricing, there is still a significant need for ancillary services to manage thermal constraints, with their role increasing as key transmission boundaries, such as the B6, become congested.

3.6 Concluding points on thermal constraints

This chapter addresses the question: "*What is the scope for addressing thermal constraint management through an ancillary service, as an alternative or as a supplement to the wider structural reforms being considered in the wholesale markets chapter of the REMA consultation, which could in theory help mitigate thermal constraints?*"

In analysing this question, this report concludes that regardless of wholesale market or ancillary service structure, effectively managing thermal constraints in a net zero 2035 system will require:

- *Significant investment in network capacity:* This is vital, regardless of other changes to the wholesale market or ancillary services. Providing a clear plan for investment in network capacity is also important to enable generation, storage/flexibility and demand assets to better understand the long-term implications of thermal constraints.
- *Using the most accurate counterfactual cost assessments for thermal constraints:* Using net cost of curtailment rather than the gross (Net costs include the change in CfD costs from BM / ancillary service actions), the net costs should be used in all NGESO documentation and analysis, including the network options assessment. This will reduce the size of the thermal constraint costs compared to using gross costs and, all other things being equal, result in a lower optimal transmission build.
- *Investment in schemes that maximise network capacity utilisation:* According to NGESO, the Congestion Management Intertrip Scheme (CMIS) has saved consumers £80 million between April 2022 and January 2023 by enabling NGESO to curtail generation only when a thermal limit is reached, rather than having to leave unused headroom in case of system faults. This approach could be further supported by an ancillary service to rapidly turn up demand behind a constraint when reaching a thermal limit to further reduce curtailed generation (see Chapter 3.4.2).
- *Investment in digital infrastructure to improve asset visibility and system responsiveness:* In particular:
    - *Improved visibility of embedded generation*, flexibility, and demand to enable more cost-effective action to be taken by NGESO to manage constraints. This is currently being trialled through the Local Constraint Market and MW Dispatch Service.
    - *Improved visibility of the status of assets* (e.g., state of charge for storage or real-time generation potential for offshore wind). This also allows more cost-effective action to be taken by the NGESO.

- *Improved automation of balancing decision making* and instructions to enable NGESO to make better use of smaller-scale flexibility assets as the network transitions away from larger-scale fossil fuel generators. Responding more rapidly to system constraints should reduce the requirement to pre-emptively constraint generation assets.

These measures can help reduce overall constraints and assess the cost of remaining thermal constraints more accurately which will allow any residual constraints within a given system to be managed more cost-effectively (at a system level) by NGESO.

3.6.1 Ancillary Services needs under different wholesale market structures

None of the wholesale market design options provides short-term constraint management signals as short-term constraints manifest after gate closure. The current wholesale market structure (national pricing) and BM neither provide locational price signals nor provide signals to dispatch efficiently before gate closure, necessitating significant volumes of re-dispatch.

All three of the proposed REMA changes to the wholesale market design outlined in this report, i.e., zonal self-dispatch, zonal central dispatch, and nodal central dispatch, provide locational and dispatch signals ahead of gate closure – nodal to a greater and zonal to a lesser extent.

Therefore, the remaining gap to be filled by additional ancillary services depends heavily on the choice of wholesale market counterfactual. However, as change to either nodal or zonal pricing will likely take a number of years before being fully implemented (if selected at all), further consideration of additional services/support schemes outlined below and in Table 4 is recommended, with a possible phase-out at a later date if no longer needed under a new market structure.

3.6.2 A before-day-ahead auction constraint price signal

As discussed above, there exists a gap in operational signals for efficient dispatch, i.e. a signal that reduces the need for post-gate closure redispatch. This signal would not be needed if zonal or nodal pricing were to be implemented.

This signal could be a pre-day ahead forecast constraint cost added to generators in areas forecast to be constrained. By sending the signal before the day-ahead auction has taken place, the signal would, depending on the strength of the signal, reduce the requirement balancing actions to a greater or lesser degree.

This signal could be considered a ‘dynamic zonal price’ in nature as it would be a reduction in power price for constrained areas without the need for an ex-ante definition of zonal areas. Similarly to zonal pricing, a strong version of the signal would encourage generation reduction in constrained areas and therefore a higher unconstrained clearing price as additional dispatchable assets would need to run in the unconstrained areas.

A key difference to zonal pricing would be that the strength of the signal could be modulated by NGESO / DESNZ. It could be a simple £1/MWh cost to run in constrained areas to start with. Note that if the cost did not change ex ante dispatch, the £1/MWh cost would still increase the (likely negative) SRMC of supported plants. As such, assuming that supported generators are bidding reasonable margins during constraints as per the spirit of Generation Licence 20A, the turn-down costs of balancing would reduce by both the lower bids and the £1/MWh cost levied on all generators in constrained areas.

A worked example would be as follows:

- Without the constraint signal, market clearing price is assumed to be £30/MWh
- Generation in (a future) Scotland, at this time, is assumed to be 20GW, mostly wind, running at 80% load factor in this hour as it is a constraint situation, so 16GW of capacity.
- Assume that the marginal constraint is 1GW capacity of RO generation for 1 hour at cost of £60/MWh = £60,000/hour cost of constraint.
- With the £1/MWh constraint cost signal, assume that the dispatch stack does not change and clearing price remains at £30/MWh.
- 16GW of capacity is charged £1/MWh which is £16,000 for that hour.
- 1GW is turned down in the BM, but the cost of this curtailment action should now be £59/MWh, resulting in a £59,000 - £16,000 = £43,000 net cost.

Of course, the signal could be increased to £3.53/MWh in this example which, if this did not change dispatch, would result in a net NGESO cost of zero as the curtailment would cost £56,470 and the constraint signal would collect £56,480.

If the service were used at this relatively low level, the effect of this approach would be to socialise the costs of turn-down costs on the generators in the zones which need turning down, reducing impacts on customers.

Alternatively, the cost of the signal could be set at a higher level to discourage generation behind the constraint directly. In this case, the effect of the signal would be similar to the that of zonal pricing with a significant reduction in generator value for all generators behind the constraint and reduction in volume for some generators behind the constraint. This ‘pre-BM turndown’ would then lead to higher dispatch in front of the constraint and higher wholesale prices for customers in unconstrained areas.

If the total value of wholesale price reduction behind the constraint multiplied by the volumes (reduction in £/MWh x MWh generated) is lower than the increase in £/MWh x MWh in unconstrained areas, the total cost of energy for the average consumer will increase.

A worked example would be as follows:

- Without the constraint signal, market clearing price is assumed to be £30/MWh

- Generation in Scotland, at this time, is assumed to be 20GW, mostly wind, running at 80% load factor in this hour as it is a constraint situation, so 16GW of capacity.
- Assume 1GW capacity of RO generation is constrained off for 1 hour at cost of £60/MWh = £60,000/hour cost of constraint.
- Assume that the marginal constraint is 1GW capacity. To discourage this capacity to run, a £60.01/MWh constraint cost signal would be needed.
    - This would then lead to 1GW additional generation needed in unconstrained areas, assuming that the clearing price rises to £32/MWh for 30GW of demand in unconstrained areas, which would be a cost of £2/MWh x 30GW = £60,000
- 16GW of capacity is charged £60.01/MWh which is £960,160 for that hour.
- There is now no need to turn down the RO plant as it will have 'self-curtailed', so the net cost for NGESO is -£960,160 and the cost for the unconstrained customers would be £60,000.

The net reduction in NGESO costs would then, all other things being equal, lead to lower BSUoS costs. Overall, there would then be a net shift of £960,160 from generators in constrained areas to customers (£900,160) and unconstrained generators (£60,000), though some of the value transferred to unconstrained generators would be used to pay for higher fuel costs for the marginal plants and therefore cannot be considered fully a shift in net value.

It should again be noted here, as will be expanded upon in Chapter 5, that if the marginal plants turned down in the BM are either CfD or merchant, the net turn-down costs to customers, assuming that the plants bid their SRMC in a constraint situation, would be minimal in the first place. If the focus of the ancillary service signal were to minimise net cost of constraints to customers by shifting the costs on constrained generators, the level would have to be very low indeed when CfD and merchant are the marginal turn-down assets.

Impact on CfD Generators from Constraint Price Signal

As outlined above, the constraint price signal could be used to socialise turn-down costs amongst generators in the constrained areas or be used as a pre-BM signal to change dispatch. If used as the latter, it will be similar to zonal price adjustments and would shift significant amounts of value from constrained generators to customers and unconstrained generators.

The above calculations assume that the constraint price signal would not be included in the CfD reference price and therefore that it would reduce CfD value on par with other generators in the constrained areas. This could result in a claim of 'qualifying change in law' or QCiL, which would trigger a re-opener or compensation from the LCCC (ultimately paid for by customers) for lost value as a result of the new constraint price signal.

However, if the constraint price signal is included in the day ahead reference price, CfD generators would be unaffected so long as the combined day ahead and constraint price signal remained positive as the CfD payments would rise commensurately. If the total price signal would become significantly negative, this would again start to erode CfD value as the CfD difference payments are capped at the strike price.

Possible Implementation Challenges

As NGESO does not currently have the legal authority to levy a constraint price signal on generators in constrained areas, this power would need to be legislated for.

The purpose and intent of the constraint price signal would need to be very clearly defined, ideally by Ofgem and DESNZ, to ensure stakeholder buy-in.

The operational setting of the constraint price signal would have to be fully transparent and ‘formulaic’ so that market participants are able to produce their own forecasts and risk assessments to optimise their dispatch incorporating the signal.

The dynamic nature of constraints would make the signal less predictable for generators than a defined zone. A de minimis limit would probably be needed to be applied, above which the signal would be triggered, e.g. a total constraint of 1GWh over a day to avoid signal noise.

3.6.3 Ancillary service options to manage thermal constraints

Table 4 summarises the recommend additional ancillary services that could be developed to manage thermal constraint challenges. Please note that the "amber colouring” in the table highlights services where the impact on thermal constraint management will depend on the service’s detailed design, which is beyond the scope of this report.

Table 4 - Summary of proposed new schemes to efficiently manage GB thermal constraints under different wholesale market options

<table>
<tr><th>REMA wholesale market options</th><th>Addresses short term constraints --> Maximise use of network</th><th>Addresses long term constraints --> minimise cost/volume of constraints</th><th>Promotes efficient dispatch of assets --> minimise re-dispatch needed</th></tr>
<tr><td>National pricing with self dispatch (BAU)</td><td rowspan="4">[Proposed] Expansion of the Constraint Management Intertrip Service + [Proposed] Short Term Constraint Reserve Service</td><td rowspan="2">[Proposed] Strategic Cycling Service + [Proposed] Storage Level Signal (SLS)</td><td>[Proposed] Before Day Ahead price signal (BDA)</td></tr>
<tr><td>Zonal pricing with self-dispatch</td><td rowspan="3"></td></tr>
<tr><td>Zonal pricing with central dispatch</td><td rowspan="2">[Proposed] Storage Level Signal (SLS)</td></tr>
<tr><td>Nodal pricing with central dispatch</td></tr>
</table>

These services address the following challenges:

Maximise the utilisation of the available network capacity:

- *Expanding the CMIS beyond the current 1.6GW and utilise it for constraints other than B6*. NGESO has stated that the 1.6GW limit is in place as a larger capacity would result in an increase in largest infeed loss, which affects the requirement for inertia and frequency management. However, the significant additional value already demonstrated by the CMIS, the Carbon Trust recommends NGESO conduct a cost-benefit analysis of the trade-offs involved as the largest infeed loss is likely to increase beyond 1.6GW.
- *Further investigate the Short-Term Constraint Reserve Service (STCRS).* The network operator currently needs to keep headroom for sudden (timescale is in minutes) increases in generation due to wind gusts when a thermal constraint is already active. The STCRS, similarly to the CMIS, NGESO can reduce or eliminate the headroom, increasing the throughput of the transmission wires and reducing thermal constraints.

Minimise cost and volume of long-term constraints:

- *Further investigate the Strategic Cycling Service (SCS).* This service aims to reduce the depth of thermal constraints by charging when constraints are high and discharging when constraint costs are lower. This would not reduce the constraint volume but could reduce the cost of curtailment and discipline dispatch by CfD generators. In centrally dispatched wholesale market scenarios, NGESO could optimise the storage levels directly, making this service redundant.
- *Investigate a Long-Duration Energy Storage Support Scheme*. The CM is not designed to bring forward long-duration energy storage (including hydrogen, pumped hydro, and other technologies like Compressed Air Energy Storage (CAES) and Liquid Air Energy Storage (LAES)). Relying only on market arbitrage opportunities and/or ancillary service revenues would result in higher investment risk (and cost) than a fixed/capacity payment, coupled with a dispatch incentive in a self-dispatch scenario. The cost-effectiveness of this option would depend on the counterfactual cost of increasing transmission capacity. However, LDES behind constraint boundaries can contribute to adequacy as discussed in the following chapter.

  When considering the LDES support scheme, it is important to note that uncertainty around future wholesale market design would add to the investment risk. As such, the required support to bring forward the service would default to the highest required under all the wholesale market design scenarios, whether BAU, zonal, or nodal.
- *Investigate a possible Storage Level Signal (SLS).* The aim of this service would be to ensure that the storage levels of (in particular) LDES are managed in a cost-minimising manner for customers. LDES would already have a wholesale market price signal to

encourage charging or discharging. However, the price signals may not be fully reflecting of the thermal constraint or adequacy risk as seen from NGESO's point of view. As such, an incentive to keep a certain level of headroom (for thermal constraint absorption) or reserve (for adequacy purposes) should be investigated. The optimal setup of this incentive would need to be based on detailed modelling by NGESO, but could include both a payment to incentivise holding a given level of storage, with greater distance from the target attracting less payments and could include a direct utilisation payment when constraints or adequacy issues manifest if the further modelling suggests that this would be needed.

Encouraging efficient dispatch:

- *Further investigate pre-gate-closure constraint price signal (ideally pre-day-ahead):* Based on demand and weather forecasts, thermal constraints can be accurately predicted hours ahead of delivery and, less accurately, days ahead of delivery. This signal would be redundant in zonal and nodal price scenarios if these were selected but could be beneficial in a national pricing scenario (including to minimise thermal constraint costs until and changes to wholesale markets are fully implemented).

Table 5 - Red, Amber, Green (RAG) analysis of two proposed operability signals for thermal constraint management – a "before day ahead price signal" and "storage level service". Services are assessed against the qualification criteria outlined in Chapter 3.2

<table>
<tr><th rowspan="2">Service</th><th colspan="4">Reliability</th><th colspan="2">Cost</th><th>Carbon</th></tr>
<tr><th>Addresses short term constraints</th><th>Addresses long term constraints</th><th>Promotes efficient dispatch of assets</th><th>Incentivises locational DSR/storage</th><th>Maximises use of existing network assets</th><th>Cost Effectiveness</th><th>Promotes low carbon technologies</th></tr>
<tr><td>Before day ahead (BDA) price signal</td><td>A BDA price signal would not help manage very short thermal contraints (seconds / minutes) that may arise from sudden falls in demand or spikes in supply .</td><td colspan="2">By providing a moderate constraint signal, the BDA would help manage longer contraint periods by shifting part or all of the costs of redispatch onto generators in the constrained area.

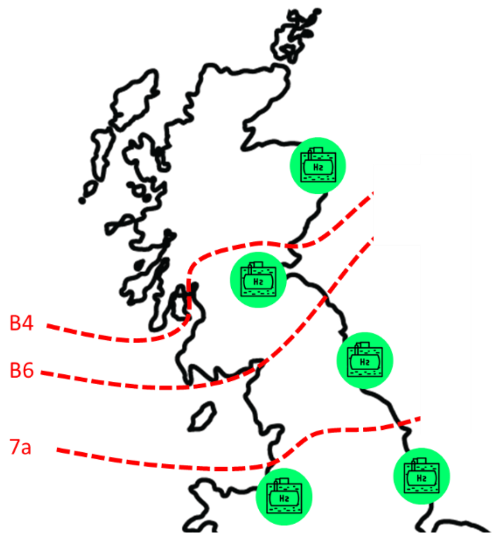
If a strong cost signal were to be implemented, this would have similar effects to a zonal price market reform.</td><td>A BDA price signal would incentivise storage in a similar way to nodal/zonal pricing by creating sharper locational price deltas which storage can take advantage of.</td><td>Assets would be no more or less utilised than the BAU senario.</td><td>Net contraint management costs would not be impacted, the risk of network constraints would be partially or fully moved from NGESO to generators in the constrained areas.</td><td>This service looks to re-allocate contraint costs to generators and is carbon agnostic.</td></tr>
<tr><td>Storage levels service (SLS)</td><td>SLS would be focused on longer-duration headroom would have minimal impact on short-term constraints (seconds / minutes)</td><td colspan="2">SLS would ensure headroom for longer periods of thermal constraints, reducing curtailment volumes and cost.</td><td>SLS would create an addtional revenue stream for LDES technologies located behind constraints.</td><td>SLS should result in additional, otherwise-curtailed generation being exported from constrained to unconstrained areas.</td><td>By encouraging retention of LDES headroom for thermal constraints, the amount of curtailed energy should be reduced. Assessing net costs for customers would need detailed modelling.</td><td>This service could allow for more otherwise-curtailed energy to be utilised in thermal constraint periods, likely displacing higher carbon generation in at least some periods.</td></tr>
</table>

3.6.4 Hydrogen options for thermal constraints management

Low carbon hydrogen-to-power has a number of advantages as an LDES technology that could assist in both thermal constraint management and securing capacity in periods of extended low renewable output. In a constraint period, curtailed energy could be used for the electrolysis of green hydrogen which could then either be utilised in hydrogen industrial clusters, stored, and converted back to electricity through fuel cells or turbines; it can also be transported across constraint boundaries with hydrogen distribution infrastructure. Among other advantages, hydrogen storage can be scaled for capacity (MWh) independent of power (MW) and has low self-discharge rates compared to other storage devices. According to the UK Government's Industrial Decarbonisation Strategy, there are six key industrial clusters in the UK that could fuel a significant proportion of hydrogen demand by 2030.

As illustrated in Figure 9, many of the planned hydrogen industrial clusters are located in key positions in relation to constraint boundaries, allowing the effective absorption and utilisation of curtailed energy. It is possible that hydrogen transmission infrastructure could be installed at a greater speed than electricity transmission infrastructure, particularly if gas infrastructure is retrofitted for this capability. Such a network could allow electrolytically derived green hydrogen to bypass constraint boundaries, helping cost optimise the level of network investment needed for the government to meet its 2035 and 2050 climate targets.

Figure 9 - Location of Northern hydrogen industrial clusters (green) against the B4, B6 and 7a constraint boundaries (red)



Source: Carbon Trust, 2023

However, hydrogen does have a number of hurdles to overcome in relation to its use as a LDES technology and further work is required to fully realise its role in the energy system. Chief amongst these hurdles is the low round-trip efficiency if hydrogen were to be produced via electrolysis of low carbon variable generation, compressed and stored, and used to fuel a hydrogen burning open cycle gas turbine (OCGT) or even in a more efficient CCGT. Assuming a high-efficiency electrolyser at 80% and a best-in-class CCGT at 60% efficiency, excluding compression, the round-trip efficiency would be 48% at best. In reality, the likely infrequent need for hydrogen fired generation as well as the fact even the best CCGTs would struggle achieving the maximum efficiency if not running at full load and/or flexible generation, the actual round-trip efficiency would be more likely to be around 35% (70% for a typical electrolyser efficiency including compression / storage and an assumed 50% efficiency OCGT).

Additionally, hydrogen has lower commercial maturity at the scales required for strategic LDES compared to other technologies and there are limited geological storage areas available in Scotland, potentially increasing the cost of using hydrogen storage to alleviate the B6 constraint. More work needs to be done to fully understand the role that hydrogen may plan in alleviating network constraints and providing capacity in periods of extended low renewable output

4 Ancillary Services for Capacity Adequacy

This chapter considers the question: "What is the role for ancillary services in creating the investment and operational signals for low carbon technologies to meet system needs during extended periods of low wind/sun?"

The REMA Case for Change indicates that a significant volume of new capacity is required to deliver a secure and low carbon system, with around 300GW of capacity needed by 2035, a significant increase from the current 100GW[61]. This implies that an average of over 15GW of new capacity must be added each year until 2035, which is more than the historical average of 5-6GW. As outlined in Chapter 2, the growth of renewable generation sources will mean the system is more significantly impacted by extended periods of low sun/wind, and it is important that the correct investment and operational signals are in place to encourage the low carbon technologies that could deal with these periods.

- *Investment signals* are price signals that incentivise developers to invest in new assets. A good investment signal should be stable, predictable, and transparent so investors have confidence in their revenue streams which reduces financing costs.
- *Operational signals* encourage the operation and dispatch of an asset in a way that compliments system needs. This could take the form of a locational price signal, or an ancillary service where specific operational characteristics are contracted by a system operator.

There are many low carbon technologies capable of meeting system needs in extended periods of low wind/sun which generally fall into three main categories:

- *Interconnectors* are high-voltage cables that connect the electricity systems of neighbouring countries. They enable power to be traded and shared between connected countries and smooth out variations in renewable energy output, reducing curtailment. However, as outlined in Chapter 4.2, an overreliance on interconnectors for power at times of low renewable output could pose a risk to energy security, as the anti-cyclonic periods of low wind and sun can simultaneously affect a large part of Continental Europe which would limit the amount of energy available to import to GB.
- *Dispatchable generation* such as nuclear, bioenergy, hydrogen, and conventional gas facilities with carbon capture, use and storage (CCUS) can generate independently of weather which makes them valuable from a system adequacy perspective.
- *Long-duration energy storage (LDES)* is a group of technologies that can store energy for long periods at scale (at least 24 hours duration at full load) and provide system flexibility by managing fluctuations in supply and demand. LDES technologies include

[61] DESNZ (2023) Review of electricity market arrangements

mechanical, thermal, electrochemical, and chemical storage. No LDES connected to the GB energy system can currently deliver power for a 24-hour duration at full load; the closest are the Foyers and Cruchan in Scotland, at around 20 hours at maximum output. Other emerging technologies such as hydrogen[62], liquid air energy storage (LAES) and compressed air storage (CAES).

This chapter will provide an overview of the role of ancillary services in creating the investment and operational signals needed to meet system needs in periods of extended low wind/sun. Chapter 4.1 will outline the current mechanism for securing capacity – the Capacity Market (CM) – and assess the extent to which it brings forward the low carbon capacity required to meet system adequacy. The following Chapters (4.2 to 4.4) will explore the operational and investment signals available for each technology relevant to system adequacy, outlining where any gaps remain and highlighting the possible role of ancillary services.

4.1 The GB energy system will become increasingly vulnerable to long periods of low renewable generation

In 2022, NGESO commissioned Afry to conduct a long-term study on power system adequacy aimed at evaluating the security of supply risks associated with a fully decarbonised power system and the necessary resources required to maintain adequacy in the 2030s[63]. The study focused on four potential resource portfolios which utilise different combinations of batteries, nuclear, hydrogen power generation, and CCUS. Afry found that the duration of the critical stress events in the GB system is set to increase over time due to additional wind, flexible demand and storage, especially in scenarios using short-duration batteries to ensure security of supply.

Anti-cyclonic conditions can cause low wind conditions and are expected to present the most significant challenges to adequacy for GB, particularly during the winter months. These meteorological systems can result in considerably longer periods of capacity shortfall in the future compared to those observed currently due to the increased future GB reliance on variable renewables. By 2038, the average length of ‘critically tight hours’ (periods where load loss will happen if demand increases) is set to increase from around 5 hours to 45 hours[63]. The report also indicated that storage duration will need to be tens of hours or days in duration to assist with critical stress events and that ‘tight hours’ can no longer be met by storage with a few hours’ duration.

4.1.1 The Capacity Market has been a success, but is focused on MW rather than MWh

The CM is the UK Government’s primary policy for ensuring security of electricity supply. It offers payments to power generators and demand-response providers for being available to

[62] Hydrogen storage and turbines for reconversion to electricity
[63] NGESO and Afry (2022) Resource adequacy in the 2030s

generate or reduce electricity demand at certain times. The CM was introduced in 2014 as part of the Electricity Market Reform package and has been subject to ongoing changes to improve the operation of the scheme[64] as well as a legal challenge[65].

To date, the CM has successfully secured enough reliable capacity to meet peak demand and maintain system stability. By design, the CM aims to secure sufficient capacity to meet a 'Capacity Market Event' (CME), at which time a signal is sent four hours in advance to participants in the CM to prepare to deliver, but to date, no CME has occurred. The expected length of a CME is not defined by regulation but can be inferred by looking at technology-specific 'derating factors', i.e., what percentage of the CM clearing price is paid to a winning capacity provider. Since the start of the CM, the average derating factors for fossil generators have risen from 85% to 93%, DSR has fallen from 90% to 72%, while storage (1-hour to 5-hour duration) has fallen to a range of 12 – 53%, see Appendix for further details[66]. This trend towards a lower derating factor for limited-duration assets is, in Carbon Trust's view, largely a reflection of the changing adequacy requirements which the CM is indirectly meeting through recalculation of derating factors to reflect declining value of short-duration storage to longer duration adequacy needs.

- *Indirectly:* The CM only indirectly deals with adequacy because there is no 'controlling mind' approach to favouring longer-duration assets; it is an output of probabilistic modelling for each CM auction. The impact of the indirect approach is that the realisation that short-duration storage is not going to be sufficient is only incrementally emerging. A more direct approach such a more targeted incentive for assets which help more with future adequacy (like LDES), could then 'get ahead of the curve' and deliver the assets most useful in 10 years' time rather than what is needed for 4 years' time (T-4).
- *Derating factor inefficiency:* The inefficiency comes from lower derating factors leading to, all other things being equal, a lower volume of derated capacity being bid into the CM and therefore higher clearing prices for all, especially the high derating factored fossil assets. The load factors of any new fossil assets (bidding for a 15-year contract) beyond 2030 are highly uncertain. As such, these assets are bidding in very high CM prices to ensure that their capex is essentially fully covered by the CM, regardless of earnings in the wholesale market – the 1,600MW Eggborough plant winning a £63/kW, 15-year contract in the 2023 CM auction being case in point[67]. Therefore whilst the de-rating factors reflect the need to ensure CM assets can provide capacity over a prolonged period, the current structure does not provide sufficient incentive for new LDES which typically have higher CAPEX costs, longer lead times and (excluding pumped storage)

[64] Elexon (2023) CM Change Proposals
[65] Tempus Energy challenged the compatibility of the CM with EU State Aid Rules in 2018 which led to a year-long hiatus for the CM until the EU commission re-confirmed that the CM is not in conflict with State Aid.
[66] NGESO (2022) EMR Delivery Body Electricity Capacity Reports
[67] EnAppSys (2023) T-4 Capacity Market Auction for the delivery year 2026-27

are often more nascent technologies.

A key ‘missing factor’ in the CM that would deal with both the indirectness and inefficiency of using only the derating approach the CM in delivering longer-duration adequacy would be if the CM (or an alternative mechanism) were to specifically incentivise new-build of low carbon, long-duration assets. As interconnectors and dispatchable generation have established investment signals (Chapter 4.2 and 4.3 respectively), this highlights a gap for LDES support[68], which includes hydrogen, pumped hydro, and more novel storage technologies such as CAES and LAES.

In addition to the issue of derating, another key factor limiting CM LDES investment signals may be a combination of the relatively short-term focus of the CM, i.e., only T-1 or T-4 (one and four years ahead of delivery respectively). Assets that take a long time to develop and build, such as pumped hydro, will therefore need to take more development risk than other assets to meet this deadline, reducing the attractiveness of the CM as an investment signal for some LDES assets. Finally, while up to 15-year contracts are available to new build assets, the contracts are relatively short for assets that have lifetimes easily exceeding 40 years such as hydroelectric facilities and, potentially, electricity generators co-located with geological storage facilities such as salt caverns (utilised for air and/or hydrogen storage).

For the reasons outlined above, the CM, as it stands, is not a sufficient investment signal to meet adequacy and to bring forward LDES. However, the shortfalls outlined above give a good indication as to what a successful investment signal is likely to look like, including:

- Values MWh storage capacity as well as the MW output.
- Stable income to fund debt (due to the high share of capex costs).
- Sufficient time between agreement of contract and delivery of capacity (procurement horizon).
- Contract lengths that are tailored/related to the asset life of the investment.

In addition to the above points, to minimise the support (and the costs for end consumers) needed to bring forward investment, assets ought to be able to earn income from participation in all relevant ancillary services markets as well as wholesale markets, as is the case with the CM today.

[68] While the term ‘Long-duration energy storage’ is used for in this report, the focus is on *electricity* storage rather than general energy storage as there is no route-to-market for heat-as-an-ancillary-service as it stands. A wider consideration, such as included in the CT GB Flex 2021 report, could include long-duration heat storage which could reduce the need for electrical heating in Dunkelflaute events and therefore ameliorate the adequacy challenge. Any recommendation in this report could be applied to heat storage as well as electricity storage if the necessary regulatory changes are made.

4.2 Investment and operational signals for interconnectors are well established

As noted to in the introduction to this chapter, interconnectors are a crucial technology for meeting system needs in periods of extended low renewable output. According to NGESO's Consumer Transformation Scenario, the level of interconnection could rise from 9.8GW in 2025, to 27GW in 2040[69]. Interconnectors use the difference in wholesale energy prices between countries as operational signals to adjust the flow of electricity. When the price is higher in one market than the other, importers can utilise the interconnectors to move electricity from the cheaper market to the more expensive market, earning the price difference as revenue. Conversely, when the prices are equal or very close, market participants have no incentive to trade electricity across interconnectors and may reduce or stop flow nominations. These wholesale market price deltas provide the operational signal for interconnectors to be utilised in line with system capacity needs.

With more interconnector capacity, GB can import electricity from neighbouring countries when the GB system is under stress; and export when power is plentiful, such as in high-wind situations, creating interdependency between countries. However, if there is low renewable output in the UK, there may also be low renewable output in interconnected countries, or high demand and high prices. Therefore, while a more coordinated approach to capacity adequacy involving coordination with other countries is necessary, it may not be sufficient in isolation to meet periods of extended low sun/wind – a notion supported by stakeholder discussions as well as analysis done by NGESO/Afry[70].

4.2.1 Revenue cap and floors have been successful in de-risking investment in interconnectors

In the UK, income floors and caps have been used to enable investments in interconnectors as part of the regulatory framework overseen by Ofgem. These measures are intended to provide a balance between ensuring that interconnector operators can earn a reasonable return on their investment while also protecting consumers from owners of interconnectors earning excessive revenues.

The income cap sets a maximum level of revenue that interconnector operators can earn from their operations, based on a formula that takes into account their costs and other factors. This ensures that prices for using the interconnector remain reasonable and competitive. The income floor provides a minimum level of revenue that interconnector operators are guaranteed, regardless of the level of usage of the interconnector. However, there are availability thresholds that interconnector operators have to meet.[71] This helps to provide a level of stability and certainty for investors and operators as they know they will earn a certain amount of revenue even if market conditions are unfavourable. The specific levels of the

[69] National Grid (2022) Future Energy Scenarios
[70] NGESO and Afry (2022) Resource adequacy in the 2030s
[71] Ofgem (2021) Interconnector Cap and Floor Regime Handbook

income cap and floor are determined through a process of regulatory review and consultation, considering factors such as the cost of capital, market conditions, and the needs of consumers. Ofgem has approved seven projects under the cap and floor regime since 2014, with a total capacity of 7.8 GW. These projects are expected to connect by 2025 and more than double GB's interconnector capacity. The consumer benefits from the cap and floor regime are estimated to be between £3.4bn and £6.9bn over 25 years, depending on the future scenarios of electricity demand and supply[72]. As investment signals for interconnectors are well established through revenue cap and floor regimes, the role/importance of ancillary services in creating investment signals for interconnectors is reduced.

4.3 Dispatchable high and low carbon thermal generation have established operational and investment signals

Dispatchable generation is needed for system adequacy during extended periods of low renewable output because it can provide reliable and flexible electricity supply when variable, renewable sources are insufficient. Dispatchable generation can be split into high and low carbon, both of which have established, but distinct operational and investment signals. Dispatchable high carbon generation is mainly CCGTs whereas dispatchable low carbon assets can utilise a range of technologies, including CCUS applied to CCGTs or Bioenergy with Carbon Capture and Storage (BECCS), nuclear, and hydrogen[73].

As with interconnectors (Chapter 4.3), dispatchable thermal generation has a part to play in assuring system adequacy in periods of low renewable output, but it is not enough in isolation. Most NGESO FES assume GB makes good use of its high wind load factors, a necessity for reaching a cost-effective, low carbon energy system. While there is a planned build out of dispatchable low carbon generation such as new nuclear and BECCS, these technologies are not expected to meet the minimum GB demand and a potentially delayed roll-out of these assets poses a risk to GB adequacy. Therefore, in addition to increased interconnection and dispatchable, low carbon generation, there is arguably a need for LDES, a notion explored further in Chapter 4.4.

4.3.1 CCGTs use the wholesale market as operational signals and the CM to de-risk investment

CCGTs earn their income through the difference in wholesale market price of electricity and its SRMC of production using natural gas. This 'spark spread' is a market term used to estimate CCGT profitability by using a reference plant efficiency of 49.13%[74]. The spark spread can be positive or negative. If it is positive, a CCGT with at least the efficiency equal to the reference

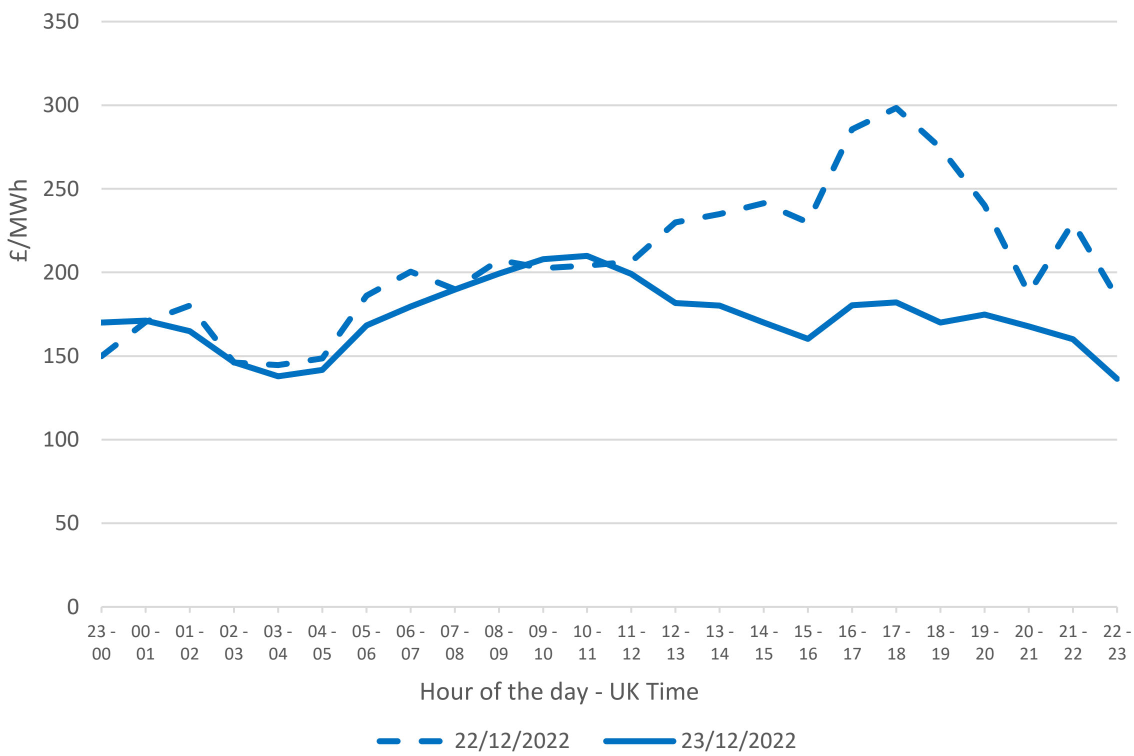
[72] Ofgem (2021) Interconnector Policy Review
[73] In this context 'dispatchable' can be interpreted as 'firm'. Nuclear plants can vary load up and down depending on need as is done regularly in France, but UK nuclear plants are not expected to provide significant flexibility services.
[74] ICE Futures (2023) Product Guide

generator earns a profit from dispatching in that hour, while if it is negative, the generator would lose money and therefore would not dispatch. This creates a strong price incentive to operate CCGTs according to system needs[75].

The spark spread depends on several factors, such as the price of natural gas, the price of electricity, the efficiency of the generator, and the cost of UK Emissions Trading Scheme (UK ETS) CO2 emissions certificates. In addition, CCGTs can earn what is referred to as 'the scarcity premium' which represents the premium which asset operators can charge above the marginal costs during periods of system stress. This setup means that when the demand for electricity is high and the supply is low, the price of electricity increases to reflect the value of flexible capacity that can help avoid power outages.

This is well demonstrated by the comparison of NordPool day ahead (D+1) prices just before Christmas 2022, see Figure 10. The power price was between £150/MWh and £200/MWh, set by gas plants through the two days (overlayed), but around 16:00 – 19:00 on the 22nd of December, prices rose to between £250/MWh and £300/MWh, which was significantly above the SRMC of any generation asset.

Figure 10 - Nord Pool D+1 Prices for GB 22/12/2023 - 23/12/2023



Source: Nord Pool, 2023

[75] Analogous price signals exist for coal (dark spread), biomass (bark spread) and even nuclear (quark spread)

While scarcity premiums can provide investment signals for dispatchable capacity, including storage, an important de-risked investment signal currently comes from the CM. As detailed in Chapter 4.1.1, the CM provides investment signals for new capacity by offering long-term contracts (up to 15 years) for new assets that have not yet been built or are being substantially refurbished. These contracts provide a stable and predictable revenue stream for investors, reducing their exposure to market risks and lowering their financing costs. The contracts also create an incentive for new capacity to be built in time for the delivery year, as there are penalties for failing to deliver when required.

4.3.2 In isolation, the CM is not enough to provide the investment signals for dispatchable low carbon generation

While, as outlined in Chapter 4.3.1, CM contracts may be enough to successfully de-risk investment in gas-powered assets such as CCGTs, the CM's current iteration also provides a moderately successful investment signal for short-duration storage (limited by decreasing derating factors for short-duration storage) but has yet to bring forward any LDES. Currently, the CM is largely[76] carbon agnostic[77] meaning that comparatively nascent low carbon technologies have to compete against well-established CCGTs where the investment risks are well understood.

To compensate for this, a number of other support schemes are available to dispatchable low carbon generation. Contracts for Difference (CfDs) support investment in low carbon renewables by providing developers with a guaranteed price for the electricity they produce[78]. CfD awards contracts through a competitive auction process, where different renewable technologies bid for a fixed price per unit of electricity. CfD contracts are well known for their success in incentivising offshore wind but are also applied to dispatchable generation such as bioenergy. The previous round (AR4)[79] awarded contracts to 11GW of renewable capacity across 11 technologies.

In addition to CfDs, the regulated asset base model (RAB) has been used to fund a number of capital-intensive projects with long asset lifetimes, e.g., transmission and distribution infrastructure. The Secretary of State for DESNZ has the authority to grant a RAB licence to an organisation, which enables the company to recover its regulated 'allowed revenue' over the duration of the project. This period can encompass the whole of the project's design, construction, commissioning, and operations. Ofgem would then determine the allowed revenue at regular intervals in accordance with the company's licence conditions.

[76] The CM is starting to eliminate some peaking diesel plant and less efficient OCGTs
[77] The CM does have emission limits, but is agnostic provided the asset meets the criteria
[78] Note that for the older CfDs that do not have 'negative price provisions', if the reference price is negative, the achieved price is the fixed strike price less the negative reference price, the total of which would be less than the strike price. CfDs therefore provide a degree, but not absolute, price stability for the CfD generators. More recent CfDs have provisions that remove CfD payments if the reference price is negative for a set number of hours, which weakens the price stability aspect of the contract, but also removes the incentive to keep generating when the power is not needed by the system.
[79] DESNZ (2022) Contracts for Difference (CfD): Allocation Round 4

The Nuclear Energy (Financing) Act 2022[80] received Royal Assent on 31st March 2022 and allows for the implementation of a Regulated Asset Base (RAB) model for nuclear energy generation. While the RAB model has yet to be applied to any specific nuclear project in the UK, it is expected that it will be used for at least one large-scale nuclear project (Sizewell C) in this parliament, subject to clear value for money and all relevant approvals[81]. The government is also considering whether the RAB model could be applied to other low carbon technologies, such as small modular reactors (SMR)[82].

While the above support schemes serve as a means of ‘fixing’ some of the problems associated with competition in the CM, an alternative approach could be amending the CM to run split auctions for low and high carbon assets, a method proposed as part of the REMA consultation and further explored in Chapter 6.

4.3.3 Operational signals for dispatchable low carbon generation currently encourage ‘must run’, with other arrangements suggested for CCUS technologies

Due to the CfD providing fixed prices (when reference prices are non-negative), low carbon assets can be considered ‘must run’, i.e., will always dispatch at the maximum available capacity. High carbon assets are incentivised to run during periods of capacity shortages both by a high market price and, if triggered, the CME signal. If not responded to, the CME results in a penalty for the non-performing assets.

The proposed Dispatchable Power Agreement (DPA) business model for CCUS projects in the UK mimics the market signals for CCGTs, ensuring that the CCUS plant runs ahead of CCGTs, but not in periods where a CCGT was not needed in the absence of CCUS. The DPA business model therefore does not incentivise the power CCUS project to generate at all times, but rather to react to market prices and provide dispatchable output when needed, without displacing lower cost and lower carbon sources of generation such as renewables and nuclear. The DPA business model consists of two components: an availability payment that is decoupled from dispatch, and a variable payment that accounts for the increased running costs of a power CCS project compared to an unabated competitor.

[80] UK Parliament (2022) Nuclear Energy (Financing) Act
[81] DESNZ (2022) Kwarteng advances plans for funding new nuclear projects, including Sizewell C
[82] DESNZ (2019) Innovative funding models for new low carbon energy

4.4 Interconnectors and dispatchable generation will support the system during periods of Dunkelflaute but there is still a need for LDES

Increasing the level of interconnection is vital, but still leaves GB exposed to continent-wide Dunkelflaute periods and the associated adequacy issues. Dispatchable high carbon assets can provide adequacy, but their use (beyond de minimis) makes them incompatible with the net zero targets. Dispatchable low carbon generation could have a significant role to play in assuring system adequacy, however deployment at the scales needed would not make cost effective use of the excess/curtailed energy from renewable generation projects and would arguably not represent a least regrets approach.

Some amount of LDES is likely to be cost-effective in a net zero energy system. LDES can function as a key enabler of a renewable-based system, providing long-term flexibility, complimenting the role of interconnectors and dispatchable generation. LDES can store excess renewable energy when supply exceeds demand and release it when needed, reducing curtailment, and increasing the cost-effective penetration of low carbon renewable assets and help cost-optimise thermal constraint management actions, a notion discussed further in Chapter 7. LDES technologies could also provide a range of ancillary services such as frequency regulation, voltage support, spinning reserve, and black start capability.

4.4.1 Operational signals for LDES will primarily come from the wholesale market, but there may be a need for ancillary services to ensure capacity

A key revenue stream for LDES is likely to be wholesale price arbitrage. Theoretically, LDES may be able to take advantage of the most extreme wholesale market fluctuations, charging up on the cheapest energy in periods of high renewable output/constraint and discharging when wholesale prices are expected to be highest, in extended Dunkelflaute periods. Price arbitrage naturally complements the operational profile of storage, incentivising the asset to be dispatched according to system needs. However, to assure effective capacity when needed, storage operators would have to accurately forecast periods of low wind/sun. While forecasting of weather patterns is improving, granular (both spatially and temporally) accurate forecasts longer than a couple of weeks remain challenging[83].

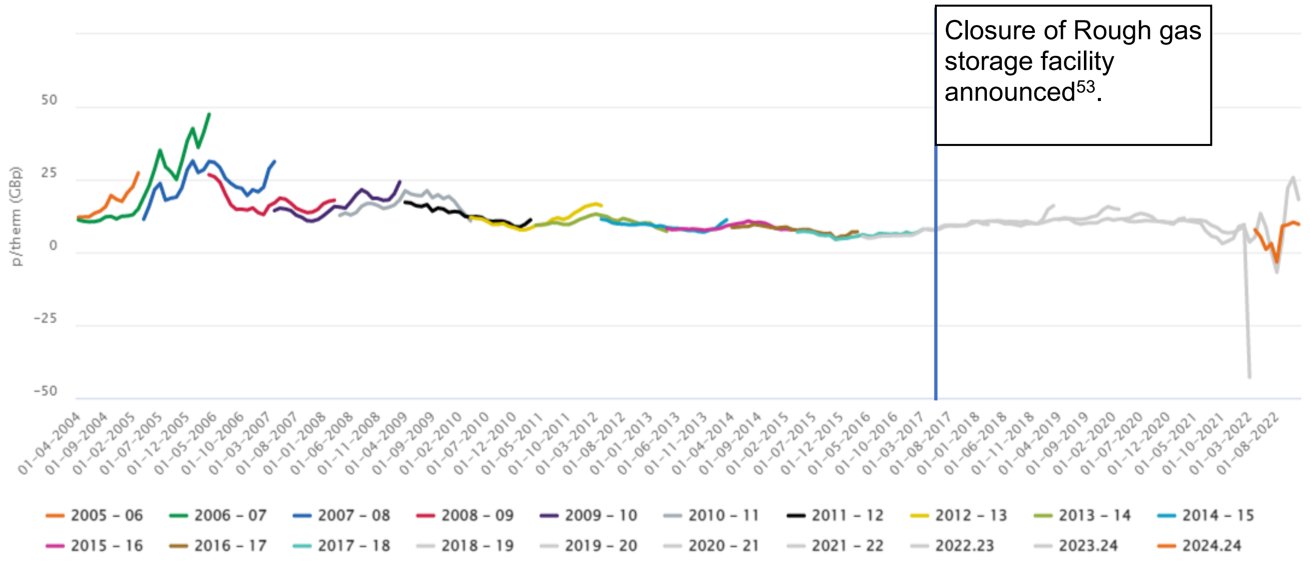
Therefore, while any new LDES would have strong, short-term price signals which would be aligned with system needs, the limited ability to forecast Dunkelflaute events sufficiently in advance would not provide an incentive to hold significant energy in reserve to manage adequacy issues. As such, a there is a need for new signals that incentivise availability in (infrequent) times of system need, similar to the incentive provided by the CM to generate during CMEs. A useful comparison asset could be the Rough natural gas storage facility off the Yorkshire coast. The Rough facility was a seasonal gas storage, injecting during the summer and withdrawing during the winter, with the value earned being optimised by injecting when

[83] Met Office (2023) Forecast Information

short-term gas cost was lower due to transient market conditions and, vice versa, withdrawing gas when prices were higher, e.g., when demand was particularly high.

Rough was closed in 2017, after more than 30 years of operations, when the cost of repairs and upgrades were such that the proprietor did not expect to earn a sufficient return on investment on the summer-winter spread on natural gas, which had declined steadily over the previous decade, see Figure 11[84].

Figure 11 - Gas summer-winter spreads at the national balancing point



Source: Ofgem, 2023. Note: The closure of the rough gas storage facility is marked by the blue line

In 2022, Rough reopened as the spreads had risen due to the market impacts of the Russian invasion of Ukraine[85].

The circumstances around the closure of Rough is a good case study for the need to provide stable conditions for low carbon LDES in the UK – having high capital costs, long lifetimes and a dependence on variable income from wholesale markets. Rough did not have a signal separate from the wholesale market to inject and withdraw gas as this was not needed at the time nor another form of support to ensure gas adequacy as this was not considered necessary given UK's access to a deep and liquid natural gas market[86].

Following on from the Rough example, if there is a partial investment case for LDES, but one that is not sufficient to bring forward low carbon capacity to deliver adequacy, some form of

[84] Centrica (2017) Closure of UK's largest gas storage site 'could mean volatile prices'
[85] Centrica (2022) Centrica re-opens Rough storage facility
[86] Reuters (2013) UK decides against intervening to boost gas storage

support may be needed either through capital investment or mechanisms which sufficiently value the services provided.

A key service to be provided by LDES is to discharge energy in periods of low renewable generation. However, the storage asset owner may not end up delivering the energy when it is needed the most by the system on a risk-adjusted basis. This may be due to the LDES asset owner wanting lock-in value in the near-term rather than hold back and potentially achieve higher, but riskier, profits later should adequacy issues materialise due to a Dunkelflaute.

Therefore, an ancillary service targeting specific energy storage reserve levels could be considered. The storage provider could receive a fixed fee for making this capacity available and potentially a variable fee for delivery.

- The availability fee could be based on the size, duration, and location of the storage capacity, as well as the market conditions and system needs.
- The (possible) variable fee could be based on the amount, timing, and price of the energy delivered.

A service of this sort could have synergies with ancillary services for thermal constraints, allowing LDES operators to be paid to charge up their assets in constraint periods, increasing the arbitrage price delta, a notion explored further in Chapter 7.

Nodal and zonal pricing could be effective means of sharpening the operational signals for LDES

Nodal and zonal pricing could be an effective means of incentivising asset operation through accurate price signals. A locational price would allow storage assets to be located where they can make best use of variable energy prices, increasing arbitrage revenues. For example, constraints at congested grid boundaries such as the B6 could encourage storage assets to locate behind the constraint and charge up on low-cost energy that would otherwise be curtailed, exporting it when there’s an extended period of low sun/wind.

While locational pricing, if selected as the wholesale market options for REMA, may be enough to change the daily operational profile of arbitrage activities, there would likely still be a need for a mechanism/ancillary service that encourages storage assets to hold capacity in reserve for periods of extended low renewable output.

4.4.2 Ancillary services will play an essential role in creating the investment signals for LDES, but are unlikely to de-risk revenues to the levels needed in isolation

‘Capacity reserve’ ancillary services such as the one described in Chapter 4.4.1, along with other ancillary services for frequency, black start, voltage etc., could make up a significant portion of storage asset revenue streams and, particularly for LDES, they have a role to play in creating the investment signals that financiers need. However, there are a number of problems which current ancillary services have which do not make them suitable instruments (in isolation) for de-risking the long-lifetime high-capex assets that are needed to address capacity adequacy:

- *Contract durations:* Investing in LDES assets can be challenging due to short ancillary service contract durations which typically last around a year, with some longer contracts such as those in the stability pathfinders, extended to as long as 15-years for assets such as synchronous condensers (SynCons) providing fault-level contributions[87]. However, hydroelectric facilities have been known to be operational for as long as 100 years and, if maintained properly, the civil engineering infrastructure could last almost indefinitely. The discrepancy between contract duration and asset lifetime can result in a lack of return security for investors, making longer-duration energy storage projects riskier and less bankable.
- *Procurement horizons:* LDES can take years/decades to plan and construct. If the aim is to minimise investment costs, it is important that revenue is secured as much as is reasonably possible before investment decisions are made. Current ancillary service contracts only procure services a few years ahead of delivery. The CM, the current investment signal for capacity, procures capacity four years ahead of delivery, which may be short for some forms of LDES.
- *Stackable services:* Procurement of ancillary services is often done separately, and contracts for various services cannot always be effectively combined. The result is a lack of clear pricing signals for investments in technologies that can provide multiple services at a lower overall cost.
- *Unpredictable revenues:* Compared to the wholesale energy market, ancillary services markets are relatively novel, illiquid, and shallow. These markets are often location-specific, such as for reactive power, and are subject to specific technical requirements set by NGESO. This lack of track record over a longer period of time and standardisation means that investors cannot confidently predict how the value of these services will evolve over long periods.
- *Regulatory uncertainty:* The ancillary services market has been undergoing a number of reviews over the last several years, and the constant changes in regulations and policies make it challenging for investors to accurately forecast the value that can be obtained from these services throughout the lifespan of an asset.

Therefore, when designing ancillary services for LDES that de-risk investor revenues, it is important that services are designed with contract lengths that compliment asset lifetimes, with procurement horizons that allow stable, de-risked revenues to be assured before investment decisions are made. While long-term contracts have the advantage of lowering project weighted average cost of capital (WACC), there would be a relatively strong incentive for NGESO to err on the side of caution to account for forecasting errors, which may lead to over-

[87] NGESO (2022) NOA Stability Pathfinder

contracting some services. The advantages and disadvantages of long- and short-term contracts are reviewed at a high level in Chapter 6.

While an ancillary service designed with the above characteristics would allow for de-risked investor revenues, a principal revenue stream for LDES is price arbitrage, which ancillary services are not designed to de-risk. Due to the primary need for operating LDES at times of extended system stress, the resulting, infrequent price 'extremes' will be less certain and predictable for asset owners, leading to higher financing costs. Therefore, as with offshore wind, there may be a need for additional support schemes that reduce cost of capital. Ideally, an effective investment support scheme would:

- *Encourage efficient dispatch* of the storage asset by allowing the asset operator to respond to the operational signals outlined in Chapter 4.4.1.
- *Reduce the WACC for investors* by effectively de-risking LDES revenue streams across an asset's lifetime.
- *Provide value for money for taxpayers*, making sure energy prices are kept low and excessive benefits are not made with public capital.
- *Promote a range of low carbon technologies*, making sure the wide range of LDES technologies are able to compete with one another.

Table 6 below compares several support schemes that have been applied to interconnectors (Chapter 4.2) and dispatchable generation (Chapter 4.3) against the criteria above for their applicability for incentivising investment in LDES:

- *The CM* is a mechanism that works to ensure electricity supply continues to meet demand in the long-term. Participants bid for contracts 1 or 4 years ahead of delivery (see Chapter 4.1.1).
- *Revenue Cap and Floor* is a system that sets a maximum and minimum limit on the amount of revenue a company can earn in a particular period. This has been successful for other long-lifetime, high-capex assets like interconnectors (see Chapter 4.2.2).
- *A RAB model* is a regulatory framework that determines the rate of return a company can earn based on the value of its assets, which are subject to regulation (see Chapter 4.3.2).
- *CfDs* are government-backed financial instrument that guarantee a fixed price for the electricity generated by renewable energy projects, providing certainty for investors and supporting the development of renewable energy (see Chapter 4.3.2).
- *Ancillary investment signal*: An ancillary service could also play a role in creating investment signals for LDES, given appropriately designed procurement horizons and contract durations that complement the planning, construction, and commissioning of LDES assets. An example of how ancillary services can drive investment in low carbon assets is the Stability Pathfinders which successfully enabled investment in low carbon

assets including synchronous condensers to provide inertia and short circuit level (SCL) by providing a fixed, 10-year annual income to de-risk the upfront capex.

However, applying this approach to LDES would be analogous to providing a revenue floor (guaranteed income) with no cap. As LDES have significantly greater opportunity than synchronous condensers to earn revenues in the wholesale and other ancillary services markets, the risk of over-rewarding would be high.

Table 6 - Red, Amber, Green (RAG) analysis of different investment support schemes for LDES. Services are assessed against the qualification criteria outlined in Chapter 4.4.2

Incentive	Compliments operational signals	Reduces WACC for investors	Cost-effectiveness for taxpayers	Promotes a range of low carbon technologies
Capacity Market (CM)	CM participants are free to respond to operational signals in the wholesale market and most ancillary services. In its current form, the CM provides no additional incentive for flexibility or prevision of ancillary services, although this is under review as part of the REMA consultation.	While the CM is effective at de-risking investment in lower capex assets such as OCGTs and batteries, contract lengths are not suited for some LDES asset lifetimes, and T-4 auctions are not far enough in advance to de-risk investment decisions in some, longer lead-time LDES .	While the auction design of the CM enables robust competition for provision of necessary capacity, the nature of adequacy is changing, i.e. the need to be low carbon and that the duration of a CME is increasingly weather driven. As such, the cost-effectiveness of the CM in delivering adequacy will likely continue to decline from current, moderate levels overall.	While the CM is technology agnostic, it implicitly favours low capex, high opex projects. Low carbon technologies have to compete with carbon intensive projects, reducing their representation. This aspect of the CM is currently under review as part of the REMA consultation.
Revenue cap and floor	Dispatch is still driven by operational price signals in the wholesale market and ancillary services. A cap and floor should incentivise LDES operators to operate their asset according to system needs to maximise revenue.	A revenue floor could, depending on the level it was sett at, significantly de-risk investment for LDES by providing investors with certainty of a minimum return on investment while at the same time allowing possible upside to be captured.	A revenue cap and floor, depending on where they would be set, could deliver value for taxpayers by reducing investment costs and capping potentially very high returns from LDES operations. Unlike RAB models, construction risks would still sit with the developer who is best at managing this risk.	A cap and floor approach could be set at a general adequacy service level, allowing different LDES technologies to compete. Alternatively, a technology or asset specific approach, while more onerous on the regulator / contractor, could also allow a wide range of low carbon technologies to compete.
Regulated asset base (RAB)	RAB models generally have incentive structures which encourage efficient behaviour. with price signals that incentivise effective asset operation would be challenging/complex, but not impossible. RIIO has been successful at producing KPIs that encourage effective asset operation, but where challenging to design and implement.	De-risks investments more than other mechanisms, providing a high degree of certainty of revenues even before project construction.	Consumers may face risks from cost overruns (e.g. nuclear power plants), but could potentially be mitigated depending on the specific methodology proposed by the regulator	Similarly to the cap and floor approach, RAB is being used to enable new nuclear and could be used to deliver LDES and/or other low carbon assets. Probably not possible to make a "service level" RAB - i.e. every contract will likely need to be bespoke.
Contracts for difference (CfD)	The CfD incentivises maximum generation output which does not align well with LDES operation. It is not clear what the correct reference price / market would be for an LDES asset to calculate the difference payments on.	Should an appropriate reference market be identified, then CfDs reduce the uncertainty of investor returns for energy generation projects. CfDs would likely not work well for LDES because they only cover energy revenues and not other sources of income (ancillary services) from storage.	Should an appropriate reference market be identified, then CfDs reduce the uncertainty of investor returns for energy generation projects. The CfD has demonstrable success in providing value for taxpayers who share in revenues above strike prices. The CfD auctions in 2021 resulted in record low prices for offshore wind (65% lower than the first auction in 2015).	The CfD has facilitated competitive bidding among projects. It also has the potential to create competition among various technologies, depending on how it is designed as well as target low carbon over high carbon.
Ancillary Services for Adequacy	The core rationale of ancillary services is to provide an operational price signal.	Ancillary services with long contracts that match LDES asset lifetimes will provide a de-risked revenue stream for LDES. However it is highly unlikely that such a service would cover project capex, and asset operators would still be exposed to merchant risk in other areas - e.g. arbitrage, other ancillary serves, balancing services etc.	While ancillary services are designed with cost effective procurement in mind, the long duration contracts needed for LDES would lead to more inaccurate procurement volumes. To assure capacity is available when needed, it is likely that services will be over, rather than under procured.	Ancillary services can be designed to be technology agnostic, and many current ancillary services are open to a wide variety of technologies.

As can be seen from Table 6 above, the CM, while efficient at encouraging investments in (particularly lower cost) generation capacity, it is not suited in its current form for LDES due to procurement horizons, contract durations, and inefficient incentivisation of LDES for adequacy purposes. However, some amendments to this scheme, which might make it more attractive for LDES, are under consideration as part of the REMA consultation (explored further in Chapter 7).

While RAB models are excellent at de-risking investor revenues, they would be complex to manage. Additionally, there is a greater risk burden for taxpayers under RAB models, as they are likely to shoulder the cost of any construction overruns/delays.

CfDs have been successful in de-risking investment in a range of low carbon technologies but are not suited for LDES in their current form as they incentivise maximising output. This is not going to be the most efficient signal for LDES operation for adequacy where the need to ensure sufficient energy is retained to meet infrequent, extended shortfalls in variable low carbon energy generation. Furthermore, CfDs work by awarding a fixed price for a service which in turn needs a liquid market for the asset operator to hedge in. It is not clear at this point that a relevant, liquid market exists for storage risk to base these CfDs off.

While an ancillary service, similar to that used in the Stability Pathfinders, could be used to enable LDES, this approach carries a significant risk of excessive benefits, being essentially equivalent to a cap and floor approach without the floor which can result in higher-than-necessary costs for end customers.

Finally, a revenue cap and floor approach allows asset operation in relation to operational signals in the wholesale, balancing, and ancillary services markets while effectively de-risking capital returns for investors. Returns to assets are controlled by the revenue cap, providing value for money for end customers.

4.5 Ancillary services have a significant role to play in creating the operational signals for LDES, but their impact on investment signals is limited

Alongside price fluctuations in the wholesale market, ancillary services could assist in creating the operational signals that allow LDES to keep capacity in reserve, ensuring capacity is available to respond to extended periods of low renewable output.

This capacity could be secured with availability and dispatch payments, with procurement volumes varying on a monthly (or more frequent if needed) basis depending on forward-looking weather analysis. This would allow storage operators to use a larger/smaller part of their capacity of daily operation, depending on system needs.

It should be noted that if cap and floor or RAB were to be used to de-risk the investment, the storage level ancillary service payment would have to be excluded from this cap and floor/RAB, or at the very least, a gain-share would need to be allowed. If not, the responsiveness of the asset to the ancillary service signal for storage level could fall to zero when total incomes are far below the floor or above the cap.

Finally, ancillary services that complement the operational and investment signals of LDES for capacity adequacy could have synergies with ancillary services for thermal constraint management (Chapter 3), a notion that is explored further in Chapter 7.

5 REMA Support Scheme Reform to meet DESNZ objectives for operability

As agreed with DESNZ project managers, the focus of this report is primarily on thermal constraints management (Chapter 3) and capacity adequacy (Chapter 4) with a secondary focus on Support Scheme Reform.

The main guiding question for this chapter is: *What is the scope for reforms to support schemes like the CfD and CM to meet our objectives for operability?*

As the future GB electricity system becomes increasingly dominated by variable, low carbon generation (largely supported by either the RO or the CfD) the incentives that support schemes provide for generators to manage system operability are important. This chapter will give a summary of the key mass low carbon electricity support schemes. It will also restate, at a high level, the operability challenges that exist and the technical solutions. This chapter also considers other ways in which the CfD could be modified, including the use of multipliers, for encouraging investment in operability. Separately, the chapter will briefly consider options for CM reform with only high-level comments on auction reform, which is being looked at separately by DESNZ.

Bringing the understanding of the support scheme function and the operability challenges together, this chapter outlines the mechanics of how support schemes disincentivise supported generators from participating in flexibility and ancillary services provision, as well as the investment in the required assets to enable this provision. The proposed REMA reforms to the CfD will then be assessed in terms of likely impact on the:

- short-term cost of providing the service,
- short-term impact on system cost,
- the wider impact on investors and their cost of capital, and,
- the degree to which a reform is likely to incentivise investment in operability-related assets.

The chapter will conclude with a look at remaining gaps and Carbon Trust recommendations for alternative reforms to ameliorate the impacts of CfD support on CfD generators' incentives to provide flexibility and ancillary services.

5.1 CfD Scheme Context

The primary obstacle that investors face when considering investment in mass low carbon power in GB is that low carbon power tends to have, to a lesser or greater degree, high upfront capital costs and low operating costs. As the asset revenue (including electricity and services income) is uncertain at the point of investment (absent support schemes), investors have a higher risk of getting insufficient returns on investment. This risk is reflected in a higher WACC that makes these projects costly, limiting the number and volume of projects taken forward. To ensure sufficient mass low carbon power was brought forward, the UK government has implemented three main schemes, beginning with the RO in 2002 (closed in 2017)[88] and the Feed-in-Tariffs (FiTs) in 2010 (closed in 2019)[89]. These schemes were replaced by the Contracts for Difference (CfD) scheme with the first CfD plant, Charity Farm PV, commissioning in June 2016.

Although both the FiT and RO schemes are closed, generation from these schemes will continue until the mid-2035s and still accounts for 78TWh in 2021/22 at a cost of £6.4bn[90] compared to CfD generation of 22.2TWh with costs of £0.3bn for CfDs[91]. These two main schemes for low carbon power, RO and CfD, reduce generator risk in different ways, though both have reduced the incentive that supported generators have to provide some ancillary and flexibility services. RO generators are awarded a (semi) fixed uplift to the price that they achieve for their energy while CfD generators are paid a variable uplift to take the value of the energy generated up to a pre-agreed strike price[92]. RO generators are awarded ROCs when generating electricity and suppliers are obligated to submit one ROC per MWh for a percentage of eligible supply volumes, set in advance of the delivery year by DESNZ[93]. The RO scheme has (overall) the same cost to end consumers regardless of how many ROCs are submitted[94]. CfD generators, on the other hand, are paid a premium which varies hourly with the wholesale price, resulting in a (nearly) fixed price for CfD output. This means that, for the end user, the price of the power produced by a CfD is (nearly) fixed, functioning like an insurance against price rises. It is important to highlight that RO generators retain market price and generation volume risk while CfD generators primarily have volume risk.

5.2 What is the operability challenge and how do mass low carbon support schemes exacerbate it?

Building on Chapters 1 and 2, the operability vectors and their related markets are outlined in Table 7.

88 Ofgem (2023) Renewables Obligation (RO)
89 Ofgem (2023) Feed-in Tariffs (FIT) - Scheme Closure
90 Ofgem (2023) Renewables Obligation (RO) Annual Report: Scheme Year 20 (2021-22)
91 LCCC (2023) Actual CfD Generation and avoided GHG emissions data set
92 If reference prices are above the strike price, the CfD generator has an obligation to pay the LCCC the difference. This would create a negative opportunity cost and should see CfD generators paying to provide services that reduce their active power output. However, not all ancillary services can currently accept negative bids – NGESO is developing this capability.
93 The forecast which sets the volumes is produced by DESNZ (as of 2022) and approved/implemented by Ofgem. Therefore, DESNZ in effect sets the cost of the RO scheme.
94 If more ROCs are produced than the target requires, the value of additional ROCs drops to zero. There will then be no recycle benefit. Depending on the impact on the market price of ROCs (which is likely to fall below the face value), given that the overall cost to suppliers and end consumers is set by number of ROCs (fixed) x price (variable) and that the recycle benefit cannot be negative, the cost of the scheme would fall. Scheme cost stability was the reason for the 10% headroom being introduced in the first instance – details can be found here.

Table 7 - Operability areas and related markets

Operability Area	Service	Current Markets
(Low carbon) Adequacy	MW over a given period of time	CM
Thermal Constraints	MWh – Turn-down of excess generation behind a constraint and turn-up of replacement generation in front of the constraint	BM, CIMS
Voltage	MVAr	BM Voltage Pathfinder, (future) Reactive Power Markets
Stability	GVAs of inertia	Stability Pathfinders and BM
Frequency	Fast acting turn-up or turn-down of services to manage frequency	Range of markets, including: Stability Pathfinders, Dynamic Containment, Dynamic Moderation and Dynamic Regulation
Within-day Flexibility	Demand Flexibility Service (DFS) (note that it is unconfirmed if this is intended to run beyond winter 2022/23)	n/a
Restoration	Black-start capability	Tenders and/or bilateral contracts[95]

Note: Within-day Flexibility has been included but it can be seen as a newly established solution area (versus a conventional NGESO operability area). Demand Flexibility Service (DFS) can be regarded as a nascent ancillary service.

The two key concerns relating to CfDs and operability raised in the REMA consultation were barriers to participation in flexibility and ancillary services and to investment in operability supporting assets. As the lack of investment is likely driven by barriers to participation, the latter will be looked at first.

[95] NGESO (2020/21) Black Start Allowed Revenue Report

5.2.1 Why do CfDs and RO plants participate less than desired in flexibility and ancillary services?

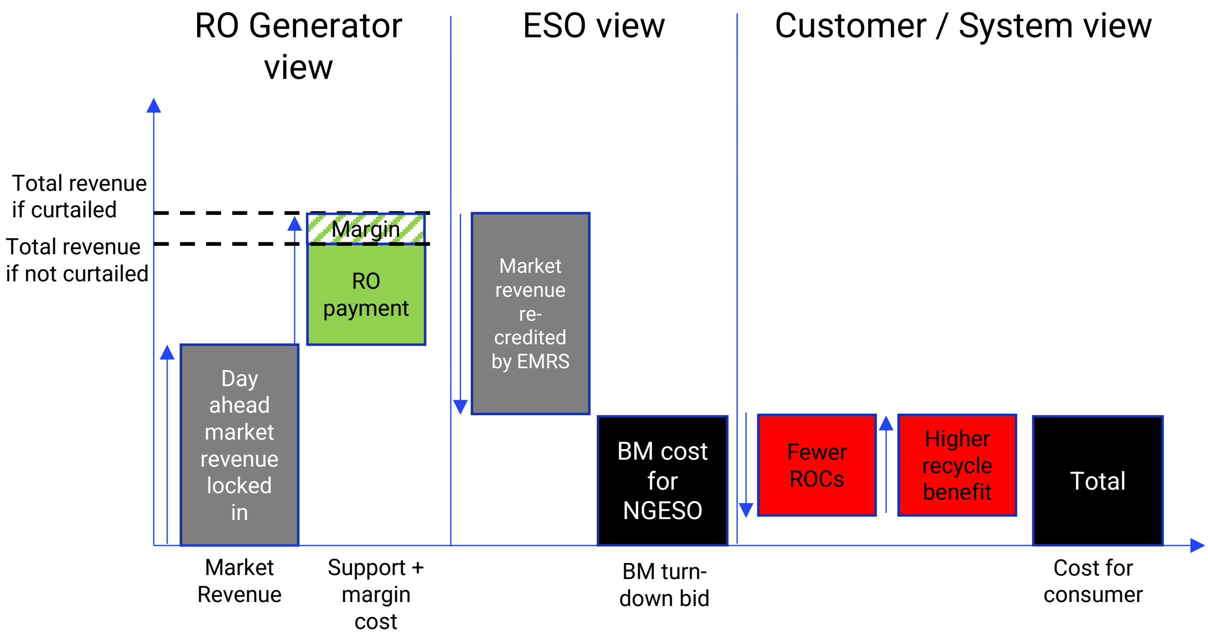
As stated in the original REMA consultation and confirmed in the consultation responses, CfD (and RO) discourage participation in flexibility and ancillary services markets because participation in these services often reduces the metered, active power output on which CfD difference payments and ROs are calculated. As such, this creates an opportunity cost to participation in these markets which will need to be paid by the purchaser of the service, i.e., NGESO, giving rise to increased gross costs of these services. However, it is clear from looking at market data that when the opportunity cost is appropriately accounted for, supported plants actively participate in the flexibility services that they are well-suited for, e.g., turn-down for thermal constraint management. As a result of these, sometimes high, difference payments for CfDs (and stable ones for RO plant), NGESO will dispatch the plants that need the least amount of payment to switch off in a thermal constraint scenario. The ‘disincentive to participation’ in this case is therefore not a capability issue, but an issue of minimising the cost of the service provision for NGESO. The REMA consultation outlined a number of potential changes to future CfDs to ameliorate the above disincentive – the effectiveness of which will be assessed in the following chapter.

5.2.2 Net versus Gross costs of flexibility and ancillary services

Before assessing the likely effect of reform proposals, it is necessary to clearly lay out the impact of CfDs and RO support payments on higher flexibility and ancillary services costs from a customer or system view. The customer cost of the RO, as outlined above, is fixed by design. This means that when an RO plant is considering whether to participate in a flexibility or ancillary service, it will need to consider the opportunity cost of doing so in terms of lost support payments. A worked example is shown in Figure 12 and Figure 13 below.

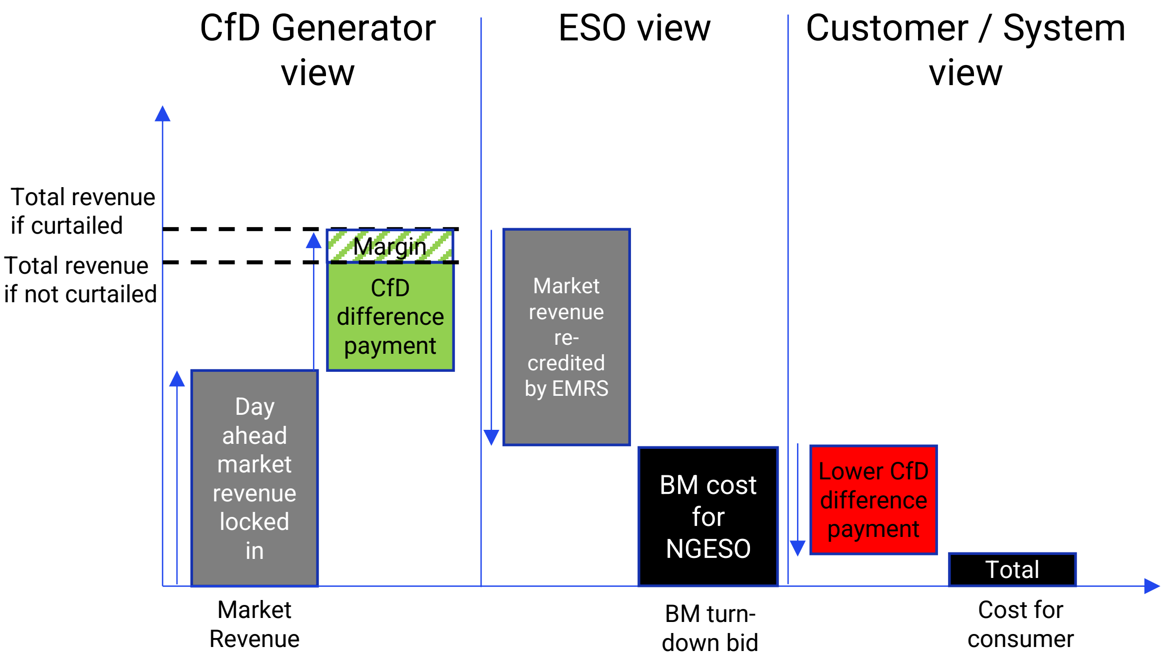
Therefore, the opportunity costs paid to RO plants are not offset elsewhere in the system and the costs, in this case BM, that NGESO sees is an accurate reflection of the cost to the end customer for dispatching this asset.

Figure 12 - Conceptual outline of cashflows for RO generators, NGESO and end customers if the generator were to be curtailed



Source: Carbon Trust, 2023

Figure 13 - Conceptual outline of cashflows for CfD generators, NGESO and end customers if the generator were to be curtailed



Source: Carbon Trust, 2023

Turning to CfDs (Figure 13) the exact situation results in a very different cost outcome for customers. This is because the opportunity cost faced by the CfD generator is fully offset by a reduction in CfD difference payments, resulting in a potentially significant difference in costs that NGESO pays for (in this case BM actions) and the cost to end customers.

This difference in customer versus NGESO costs means that despite the CfD generators looking to NGESO when considering them for flexibility and/or ancillary services, the net cost to the end consumer for CfD generators providing these services is low. The latter assumes that CfD generators bid their actual short run marginal cost with a small margin.

5.2.3 Do CfD generators bid their SRMC?

If CfD generators are bidding close to their SRMC, the CfD generator provision of flexibility and/or ancillary services is low. This situation would manifest in the market as follows: when day-ahead market reference prices are above a CfD generator's strike price, the generator should pay NGESO to be turned down in the BM and/or provide certain forms of ancillary services like reactive power (over and above the regulatory minimum). This behaviour has been seen occasionally in the BM, with an example being a CfD windfarm paying NGESO over £700/MWh to be switched off in early January 2021 when prices spiked[96].

Negative bids into ancillary services are currently not possible, but this capability is being developed and launched as part of Enduring Auction Capability. For example, the Dynamic Regulation High (DRH) prices have conformed to £0/MW/hr because it's effectively free charging for batteries. Negative pricing in response services will allow batteries to pay NGESO to charge (and provide high response) providing the cost is less than what it would cost to declare unavailable and then restore state of energy through an energy trade[97].

Reactive power is paid for through ORPS ex-post based on the amount of reactive power which is provided (as per grid code obligations or Mandatory Services Agreement). The price of ORPS is determined in reference to Month Ahead power and oil prices[98],[99]. This means that any ORPS volumes requested from CfD generators when D+1 prices are above strike price would be paid for based on a month ahead index while the actual of reactive power provision by the CfD generator would be negative, arguably resulting in excess returns for the CfD generator.

The answer to the question of whether CfD generators bid their SRMC into flexibility and/or ancillary services during high price periods is, 'no' or at least not universally. It follows that the cost to end consumers in these periods will be higher than they need to be. The Carbon Trust

---

96 However, there are also several examples of CfD generators bidding substantially less than their SRMC to be turned down in the BM (analysis of publicly available BMRS data produced by NGESO and provided privately to DESNZ – not in the public domain), resulting in excess profits in a constraint situation, something that Generation Licence 20A is designed to prevent. Ofgem, NGESO, and DESNZ are aware of these ongoing issues
97 CT conversations with NGESO
98 NGESO (2021) Enduring Auction Capability: Project Launch
99 NGESO (2012) Schedule 3 of the CUSC

has recommended changes in this report to help remedy this situation which are outlined at the end of this chapter.

5.2.4 Investment Signals

The operability services outlined above require differing levels of investment for CfD and RO plants to participate. Some, like flexibility services through the BM, require no additional investment while some, like inertia and reactive power provision (beyond the regulatory minimum) do require investments.

As NGESO is faced with high gross costs for dispatching supported generators in flexibility and ancillary service provision, as outlined above, it is reasonable to assume that NGESO supplants the high gross cost/low net cost dispatch of CfD generators in favour of services provided by alternatives. These alternatives can include high carbon plants, batteries, as well as more specific transmission assets such as Synchronous Condensers. If a generator (CfD or otherwise) therefore does not expect to be dispatched for these services sufficiently frequently, this discourages investment in otherwise cost-effective flexibility and/or ancillary service capabilities. An exception to this is the recent announcement that NGESO has awarded Dogger Bank C a £25m contract to provide 200MVar reactive power from 2024 to 2034[100]. Dogger Bank C has a current strike price of £49.47/MWh[101].

5.3 REMA CfD reform options and their impact on CfD operability

The original REMA consultation outlined a number of options for CfD reform. DESNZ considered that the majority of these could possibly encourage investment in capabilities to provide flexibility and ancillary services. They also considered that the reforms might encourage participation in these markets by increasing market price exposure for CfD generators[102]. While the proposed CfD reforms would all reduce the gross flexibility and ancillary services costs faced by NGESO – which is one of the goals that Ofgem has set for NGESO[103] – the actual impact on net cost to customers from these services is likely to be minimal.

In the event of any of the reforms, NGESO would likely dispatch CfD/Cap and Floor generators more often for flexibility/ancillary services (where the choice exists) and this may provide a limited market signal for investors to build/upgrade assets to increase stability. This signal could be significantly enhanced through targeted tenders like the Pathfinders for Stability,

[100] Dogger Bank Press Release (2022) Dogger Bank C in UK offshore wind first to provide reactive power capability
[101] LCCC (2021) Dogger Bank C P1
[102] Invitation to Tender for REMA Operability Project
[103] NGESO Annual Report 2019/20: "Our goal is to provide fair energy competition across vital services like energy balancing, to keep costs for consumers as low as possible and provide society with reliable, affordable and clean electricity today and in the future."

Thermal Constraints, and Reactive Power that NGESO has recently enacted. Tenders under a reformed CfD should have lower gross costs due to the lower opportunity costs for CfD/Cap and Floor generators relative to the current CfD.

Table 8 includes the high-level impact on investors in mass low carbon assets, in general, based on a combination of stakeholder feedback and Carbon Trust analysis. The aim of the table is to highlight the differences in drivers for the two investment signals. I.e., it is possible to make mass low carbon more attractive as an investment class while at the same time only providing limited incentives for co-location or upgrading of assets to provide increased operability capability.

The table intentionally does not comment on whether reforms would reduce overall system costs for GB end customers as this is not possible to establish with any degree of confidence without doing detailed modelling which is outside of scope for this report.

Table 8 - Red, Amber, Green (RAG) analysis of the effect of different CfD reform options on ancillary services costs and investment signals

REMA Options	Description	Impact on gross flexiblity / ancillary services costs for NGESO	Impact on net cost of flexibility / ancillary services from CfDs for customers	Encourages investment in ancillary service infrastructure	High level impact on CfD investor certainty / WACC
BAU CfD	The existing CfD scheme provides certainty to investors in low carbon projects, by guaranteeing a pre-determined 'strike price' for every MWh generated. If the reference market price is below this, they receive a top-up. If it is above, they must pay back into the scheme. In each round the strike price is set through a competitive auction and contracts are awarded for 15 years. Under latest rules, generators will not receive payments under a CfD when the reference price is negative	Baseline			
Nodal/Zonal CfD	Under a zonal or nodal wholesale market the Market Reference Price would need to be changed to either the respective locational index or a national system price. For the purposes of the assessment we assume the former, but that would remove the locational signal. Therefore it is assumed that some degree of locational differentiation in allocation rounds may be included in the Evolved CfD approach, for example to align with the Holistic Network Design.	Overall, nodal/zonal pricing would substantially reduce combined gross flexiblity and ancillary services cost, particularly constraint management. The pricing for ancillary services would still be priced price on opportunity cost of CfD payment. Could rise for constrained areas (lower wholesale price) and fall for unconstrained	If net CfD costs were considered for flexiblity and/or ancillary services dispatch, this would be equivalent to merchant plant cost, assuming CfDs dispatch according to SRMC	Constrained areas likely reduced incentive to invest due to higher opportunity costs, though frequent (self-) curtailment may encourage investments in assets to gain value from these hours. Lower CfD opportunity cost in unconstrained areas may lead to higher utilisation and invesment signal.	Nearly all stakeholders interviewed believed that a move to nodal or zonal would increase investor risk. Specifically, volume risk would rise for CfD generators as the sharper local price signals would lead to (un-remunerated) self-curtailment for generators in constrained areas.
CfD with Price cap and floor	Instead of a single strike price, generators are guaranteed a maximum and minimum price per MWh output, with market exposure within that range.	When prices are between cap and floor, gross opportunity costs are zero for CfDs, reducing costs for NGESO. When below floor, some gross opportunity cost remaining while above cap should be gross negative cost, i.e. paying to reduce active power		Increased utilisation for ancillary services by NGESO due to lower gross costs may provide limited investment signal for increased capablity.	An increase in market exposure would, all other things being equal, increase the risk for investors and their WACC over a fixed CfD. To reduce investor uncertainty, an asymmetric floor and cap prelative to a counterfactual strike price could encourage investment, e.g. a fixed strike price of £50/MWh or a floor of £40/MWh and a cap of £70/MWh. This would lead to higher costs to consumers.
Revenue cap and floor	Generators would be guaranteed a minimum revenue in each period. They would compete in the full range of markets (capacity, wholesale, balancing, ancillary services), and if they do not meet their minimum revenue, then they would be topped up. Above the cap, a proportion of the excess revenue would be paid back. There would be no transfer if revenue was between the floor and the cap. In addition to supporting mass low carbon power this option could be available for certain forms of low carbon flexibility.	As there would be no hourly support payment and assuming that the wholesale price allows for a D+1 auction locking in revenues, the gross cost of providing flexiblity and ancillary services ought to be equivalent to merchant plant.		The floor could be set at such a level as to guarantee a minimum return to operability enhancing assets located on the generator substation. Low gross costs should lead to higher utilisation of ancillary services provision and a higher likelihood of earning more than the floor.	Assuming that the floor would be set at a sufficient level, the reduction of volume risk would make a cap and floor more certain than the current CfD, increasing investor certainty. However, a "soft cap" could lead to higher overall costs for consumers, though still likely lower than a "hard cap" as this would incentivise participation in ancillary servcies and flexiblity.
Deemed output CfD	Generators are paid based on their potential to generate in a particular period, rather than their actual generation output. Generators would not have to export energy to receive their CfD top-up payment, as they do currently. This aims to remove dispatch distortions by decoupling support from output.	Deemed output would allow generators to participate in flexiblity and ancillary services markets similarly to merchant plant, reducing gross costs.		Increased utilisation for ancillary services by NGESO similarly to nodal/zonal. However, the simplicity / clarity of deemed would likely give greater confidence in increased utilisation and a storngher investment signal for increased ancillary services capablity.	Deemed CfDs would allow *at least* the same return for generators as the current CfD. In addition, some CfD generators may be able to optimise further (or even game) CfD so as to increase returns above the standard CfD.
Non Price Factors / Multipliers in CfD auction to encourage investment in ancillary services provision	CfD auction design can be reformed to encourage non-price factors such as supply chain investment, sustainability and/or ancillary services capabilities. The purpose would be to ensure that the assets supported by the CfD minimise negative operablity impacts of adding non-syncronous generation	The impact on gross costs from additional capablity being added by CfDs is unclear. If the difference payment opportunity cost remains in place, the cost of flexibility and ancillary services provided by CfDs would not be materially lower cost.	As above, but may have minor benefits from increased competitive pressures reducing margins on ancillary services.	Adjusting strike prices up to fund investment in increased capability to provide ancillary servcies would clearly drive investment. However, this could lead to oversupply of ancillary services capability and crowding-out investments, which could be more cost effective than linking to CfDs. Further analysis will be needed on this.	Assuming that the adjusted CfD would clear at a sufficient level to fund the investment in enhanced ancillary services capability, this would give CfD generators an alternativeC8:G10 revenue stream to CfD payments which should lead to higher returns, all other things being equal.

5.4 Remaining gaps and Carbon Trust recommendations

NGESO's current approach to procuring flexibility and ancillary services is based on gross costs. While dispatching on gross costs are largely accurate for RO supported plants, CfD supported generators, having low net costs, are almost certainly sub-optimally utilised in flexibility and ancillary services markets. This underutilisation, in turn, leads to sub-optimal investment signals and, quite probably, a higher-than-necessary operability cost for end customers.

As outlined above, the underlying driver of sub-optimal utilisation of CfD is that the net cost to consumers is not considered when flexibility and ancillary services are procured. The (significant) reform options for the CfD outlined in the REMA consultation would, to varying degrees, reduce the gross cost of dispatching CfDs for flexibility and ancillary services and provide limited investment signals for investment in operability-improving assets. However, significant reforms to a well-established and successful investment vehicle for mass low carbon carries high risk and may result in higher overall system costs even as operability costs are reduced. Further work is therefore necessary to accurately assess this challenge.

Of the options outlined in the REMA consultation, the outcome that would most strongly drive investment in increased ancillary services capability would likely be the (soft) Cap and Floor, though the exact impact on investment signals for operability assets would depend on whether the floor would include an allowance for operability assets. The Carbon Trust recommendations focus on eliminating the impact of the gross/net cost split directly and in such a way as to be applicable to existing as well as future CfD assets. These recommendations are grouped into:

- Improvements in reporting.
- Reform of the BM/Ancillary services procurement approach.
- (if needed after the above changes) Reform CfD difference payment rules to make allowance for approved flexibility and ancillary services provision.

5.4.1 Reporting improvements and Generation Licence 20A reform

Some existing CfD generators have been profit-optimising during periods of transmission constraints, contrary to the intent of Generation Licence 20A. This has led to higher-than-necessary costs for GB consumers from thermal constraints management and possibly some ancillary services. To establish the magnitude of the problem, the Carbon Trust recommends that a daily report be published (this can be done on already public EMRS data) on gross and net costs impact on consumers from current and historic flexibility and ancillary services from CfD generators.

After a possible grace period for generators to adjust, the Generation Licence 20A should be amended to specifically require Ofgem, NGESO, and CfD generators to include CfD difference payments when considering flexibility and ancillary services bids in periods of thermal constraints. Ideally, this requirement should be extended beyond periods of thermal constraints. The result of these changes (relatively simple and quick-to-implement reforms) would be to give CfD generators a strong incentive to bid closer to their SRMC into the relevant flexibility and ancillary services markets, which would help ensure that the net cost of dispatching CfD generators for these services is minimised.

5.4.2 BM and other ancillary markets' procurement reform

The Carbon Trust recommends that Ofgem allow and require NGESO to move from the current 'gross cost' approach to a 'net system cost' approach to BM and ancillary services procurement. This would necessitate using (publicly available) data on all CfD generators and creating a 'net cost supply curve' rather than the current gross cost curve used when deciding which assets to dispatch for flexibility and ancillary services.

Further recommended change would be for the Obligatory Reactive Power Service (ORPS) calculation methodology[104] which, being based on month-ahead electricity and oil prices, implicitly assumes that the cost of provision of reactive power value is related to wholesale prices (gross costs) which, for CfD generators, is not the case. NGESO should be encouraged to include in all future cost benefit analyses, both the current gross and net cost approach, to assessing value for money.

While the above changes would result in a higher gross cost (and therefore higher BSUoS) costs than the status quo, this would almost certainly reduce the system costs (i.e., BSUoS net of CfD difference payments). The Reporting Reform outlined in the previous chapter would be sufficient to quantify the cost savings prior to the BM reform and build wider support for the reform.

5.4.3 (If needed after the above changes) Reform CfD difference payment rules to make allowance for approved provision flexibility and ancillary services

While the changes outlined in the previous chapter should result in a more optimal utilisation of CfD assets and a lower net cost for consumers, a simple, retroactive change to CfDs could be implemented to allow for CfD payments to continue to be paid on Loss Adjusted Metered Output (LAMO) except for cases when the metered output was reduced as a result of participation in approved flexibility and/or ancillary services.

This would be similar in nature to the current imbalance exemptions for participation in the BM. If a generator has committed to generating and has sold its generation on the day-ahead market but is subsequently instructed by NGESO through the BM to turn down capacity, the generator has the BM volume credited to their imbalance account and is paid for the delivery of

[104] NGESO (2023) Obligatory Reactive Power Service Data Portal

the energy. Giving an CfD allowance for volumes bid into relevant balancing services has similarities to the 'deemed generation' CfD, but is not open to gaming. This is because "deemed generation" volumes would only be those volumes bid into NGESO flexibility and ancillary services markets. It follows, therefore, that as the CfD generator would still be paid (or has to pay) the difference payments for the energy volumes utilised for flexibility and/or ancillary services, the gross cost bid into the relevant service should be equivalent to that of a merchant plant, i.e., for variable generators such as wind, limited to a minimal reasonable margin.

5.4.4 NGESO should continue engagement and increase transparency of system-optimal analysis through the 'Virtual Energy Network' work

A view that has been often repeated by regulators, government, and stakeholders alike, is that variable renewable generators will need to provide a number of flexibility and/or ancillary services as GB decarbonises electricity. Whilst this may be true in many cases, the optimal solution is often not clear. An example would be that it is possible for interconnectors to provide reactive power services. However, given the network impact on the other side of the interconnector, investment costs and externalities, this option may not be cost-optimal. The same could be the case for provision of flexibility and/or some ancillary services for other assets.

NGESO should define the theoretical optimal system dispatch e.g., under a central dispatch approach, modelled using the 'Virtual Energy System' of the GB power system that NGESO has started to develop[105]. The transparency that this approach provides would combine well with the reporting reforms outlined above and will be critical to build trust between electricity market stakeholders and clarity on what an efficient future system looks like. Any material differences between this optimum and the actual outcome could be monitored and transparently communicated on an ongoing basis. Further reforms could then be enacted to bring the system closer to the required state while managing any unintended consequences in other parts of the system, e.g., on investment certainty.

5.5 CfD Conclusion

In summary, the impact from CfDs on gross costs has led, and continues to lead, to a lower-than-optimal dispatch of CfD generators for flexibility and ancillary services. The reforms considered in the REMA consultation solve, to a lesser or greater degree, the issue of sub-optimal dispatch. However, the other reforms would only apply to future CfDs and not solve the high gross price issue of already contracted CfDs. Whilst some changes could be applied retrospectively, significant changes like a move to Deemed Generation or a Revenue Cap and Floor would likely not be possible. Furthermore, any change that would be seen as disadvantageous for existing CfD generators would likely not be accepted by CfD generators

[105] NGESO (2023) Virtual Energy System

without generous compensation for the change, limiting the net value of such a change to GB customers.

The Carbon Trust recommendations therefore focus on solving the gross/net cost disparity directly, for both existing and future CfD generators, through improved reporting and monitoring. These changes could be complemented by BM/ancillary services procurement reform and, if considered necessary, a reform to allow a limited form of ‘deemed generation’ to apply to reduced metered volumes as a result of provision of approved flexibility and/or ancillary services – similarly to the adjustments for generators if they are out of balance due to following instructions from NGESO.

Finally, the Carbon Trust recommends that NGESO is allowed prioritise and resource the critical, confidence-building work a robust and transparent ‘Virtual Energy System’ of the GB network. Alongside the modelling, NGESO should be given the specific, and well-resourced, role of educating stakeholders with the aim of building trust and buy-in for any necessary, future reforms to enable the GB journey to net zero.

5.6 CM Scheme context

Chapter 4 on Adequacy has introduced the CM in some detail. As such this section only provides a high-level overview of what the CM is intended to do and what are the related operability challenges.

The CM was introduced as part of the EMR reform alongside a carbon price floor, CfDs, and an emissions performance standard which effectively banned new-build coal. The role of the CM is to ensure that, up to a standard set by DESNZ and NGESO, there is enough generation capacity available to meet demand in all but the most extreme cases.

The target capacity for each of the T-1 and T-4 auctions is proposed by NGESO, possibly adjusted and finally approved by DESNZ. Each technology is given a derating factor, discussed in Chapter 4, which reflects NGESO’s modelled average contribution of the technology to periods of constrained capacity. As the GB system composition changes over time, the adequacy needs will change as well, with the derating factors changing as a result.

Capacity providers that successfully bid into the CM receive a payment in return for an ‘auction acquired capacity obligation’ or AACO, which is the derated capacity. This obligation means that capacity providers must be ready to generate within 4 hours if ordered to by NGESO through the Capacity Market Notice.

Outside of this notice, obligations under the CM include:

- Confirmation that capacity providers are available by outputting at a minimum of 100% of AACO three times during the November to February (inclusive) period each year.

- Provide relevant data to the Electricity Settlements Company (ESC) responsible for settling the CM, including required data flows, emissions declarations and, for new-build providers, milestone achievements to ensure that the capacity commission on time.
- Provide the required credit cover.

CfD contract holders are excluded from CM to avoid stacking of support. With the changing grid mix, the operability challenges in addition to adequacy, discussed in Chapters 1 and 2, are changing as well. The question that this section is looking to answer is what reforms to the CM, apart from auction reforms, could be implemented to help deliver sufficient operability capability for a low carbon electricity system.

5.7 Does the CM incentivise or disincentivise capacity providers' participation in ancillary services?

The vast majority of ancillary services and thermal constraint turn-up services are provided by capacity contracted under the CM. NGESO publishes a list of 'relevant balancing services'[106] which, if providing these results in reduced generation output provide by capacity providers, this reduction is re-credited to capacity providers as if they delivered the full output. Capacity providers are therefore free to provide flexibility or ancillary services at any time, including during capacity market event (CMEs) when the CM obligates capacity providers to maximise output.

As such, the CM as it is currently set up, does not present any obstacle to capacity providers' participation in flexibility and ancillary service markets – in other words, the CM income is 'stackable' with other revenue streams. On the other hand, the CM does not directly incentivise capacity providers to include greater capability to provide ancillary services.

5.8 How can the CM incentivise capacity providers' investment in enhanced operability capabilities?

In this context, the concept of incentivising investment in capability is assumed to be for new-build only. When considering the T-1 auction, which exclusively provides 1-year contracts, the share of new-build is for the past 3 years was 17%, around 2/3rds small gas generators and 1/3rd batteries[107]. However, given that the T-1 is held around 6 months before the start of the delivery window (November – February), it is highly likely that a significant portion of what is classed as new-build in T-1 has, by the time of the T-1 auction, have already fixed their design and taken Final Investment Decision (FID). As such, any reform to add an ancillary service capability in the T-1 would likely have very limited effect.

[106] NGESO (2023) Relevant Balancing Services Guidelines
[107] NGESO-EMR Delivery Body (2023) CM Auction Result Database.

For the T-4 auction and with up to 15-year contracts available, a signal in the CM could encourage capacity providers to build more capable – and more costly – plants if, for example, a multiplier were added to the clearing price for different services. At a conceptual level, this could be done as ‘CM clearing price in £/kW derated capacity * 1+ Reactive Power factor x 1 + Inertia factor’ etc. These factors would likely need to be known in advance to reduce uncertainty for bidders as otherwise it could become prohibitively complicated for capacity providers to decide which plant configuration to bid into the auction. The downside with pre-selected factors, on the other hand, would be reduced competitiveness and the need for NGESO/DESNZ to ascribe a value to the services in advance, which could lead to over- or under-procurement of any given service.

If the multipliers were known in advance, the new-build capacity providers could consider the (uncertain) CM clearing price for capacity only (current case) and evaluate if adding additional capability to the plant would be worthwhile. While this multiplicative approach would amplify the volatility of the clearing price for capacity providers, it would at the same time make it less complicated for capacity providers to compare their options. Investors could compare a ‘basic’ capacity for which the value from the additional factors were zero. If the clearing price were to be below this number, the plant would not be built.

This approach could be compared to a set of enhanced options with varying capabilities for ancillary service provision. As long as the cost of enhancement were at least the minimum clearing price x the relevant factor, this would be the plant that would be bid into the CM.

Worked example

Battery (de-rating not included for simplicity):

1. Minimum clearing price for basic battery needed: £40/kW
2. Bonus for providing a given amount of enhanced reactive power: e.g., 10%
3. Cost of increasing reactive power capability: £3/kW
4. Battery bid in the ‘basic’ CM could then either be:
    - £40/kW OR
    - (£40/kW + £3/kW)/1.1 = £39.09/kW

In the above example, if the battery bid in £40/kW and set the price, the basic battery would be built. If the battery decided to include the enhanced reactive power capability, it could drop the bid to £39.09/kW as with the 10% bonus, the actual clearing price it would be awarded would be £39.09/kW * (1 + 10%) = £43/kW, which would be the cost of building this enhanced battery. The downside of this approach would be that some ancillary services like Reactive Power and SCL, are very locational in nature. A CM that added rewards for such services would need to be locational in nature to reflect the true value of the service.

The addition of multipliers would, all other things being equal, add to the complexity of the CM, which would reduce the accessibility for smaller, less sophisticated capacity providers. Adding a locational element would add further complication, further undermining the simplicity that is one of the CM's features.

Finally, there would be a significant uncertainty around whether this approach would provide value for money for end consumers. Considering NGESO's approach to ancillary services, the preference has been for separate markets for different services to encourage competition[108]. If the CM became a bundled scheme to provide additional ancillary services capability, the specific procurement volume of each type of service would be difficult achieve and would likely require additional sub-rules, further complicating the CM.

5.9 Streamlining bureaucracy and qualification for a wide range of services could bring further assets to market

Beyond bundling with CM itself, other actions can be taken to streamline the procurement and transparency of CM and ancillary services procurement. A key concern raised in stakeholder conversations around the CM and ancillary provisions is the total complexity of participating and the administrative burden to prove availability. Individual schemes – the CM requirements listed in the section above – are not considered onerous, but the fact that participants need to prove qualification for each service separately and possibly provide credit cover multiple times, raises barriers to participation, particularly for smaller, less sophisticated providers of capacity and ancillary services.

One way of managing this complexity is for the small service provider to sign up with an aggregator and/or supplier to pool the assets. While this can reduce the complexity barrier, adding layers to the participating results in higher costs through mark-up. Alternatively, standardisation and joint accreditation/qualification for service providers could streamline the participation process and reduce the need (and cost) for aggregation.

The concept could work as a 'platform' approach for which service providers will need to qualify. From this platform, they could then decide which markets to bid into, i.e., CM, Frequency etc. The benefit of this approach is the value of participation for smaller assets for a given amount of bureaucracy/credit cover while at the same time increasing the visibility of assets and their capabilities to those procuring these services.

A downside of moving from service-specific assessments would be an increased risk that some assets could be awarded service contracts for which they might not be fully suitable.

[108] NGESO response to a question around bundling services was "Bundling services potentially introduces barriers to participation as assets might not be able to provide all the services asked for which might preclude them from any tender." : NGESO (2023) Pathfinders Markets Day – Unanswered Questions

However, this would be (at least partially) offset by the increased participation and competition of assets, which should reduce prices for the wider services.

5.10 CM conclusions

The CM has been successful at retaining sufficient capacity on the GB system since it started in 2015, though it has brought forward limited efficient gas plants and no LDES. The changing nature of capacity provision, with wind and solar resource shortfalls becoming an increasingly significant driver of capacity adequacy concerns, means that the CM will need to change to meet these challenges. Its primary purpose of ensuring adequacy of capacity was discussed in Chapter 4, which recommended a separate signal for LDES. This could be done through the CM, though this report concludes that a Cap and Floor regime is likely to be more cost-effective for end consumers.

This chapter has, at a high level, considered the use of multipliers in conjunction with the standard CM auction to indicate preference for assets that can provide a greater range or degree of ancillary services. However, this ‘bundling’ of services would both increase the complexity of the scheme and appears contrary to the NGESO preferred approach for separate markets on ancillary services to encourage competition and cost-effective procurement.

Finally, this chapter briefly outlined options to standardise qualification and verification requirements across the CM and ancillary services markets. A ‘platform’ approach with reduced bureaucracy would likely increase the participation rate of, particularly for smaller and less sophisticated capacity providers, increasing competition and reducing the cost of procurement for necessary ancillary services.

6 Ancillary Service Procurement Strategies

Beyond the main focus of Chapters 3 and 4, we have included high-level commentary on the TOR Questions 5 (Impact of different contract lengths for ancillary services), 8 (Local vs national ancillary services) and 9 (Co-optimisation of energy and ancillary services) in this chapter. These questions are considered at a high-level and are intended to provide a starting point for future research rather than being a definitive examination.

6.1 TOR Question 5

Under what conditions might a long-term contract be appropriate for the provision of ancillary services rather than closer to real-time markets?

Ancillary services are procured by NGESO in two ways, long-term contracts and close-to-real-time markets. The means through which a service is procured is dependent on both characteristics of the ancillary service in question and the technologies providing it. It is important to note that across most operability areas, both long-term contracts and close-to-real-time markets are used to procure services and there is no set cut-off in terms of timeframes that separate the two.

Long-term contracts are bilateral legal agreements between NGESO and ancillary service providers such as generators. These agreements will set an amount, type, or duration of service that generators must or be able to provide. Such contracts typically operate in timeframes of over a year and are contracted years in advance by NGESO. The operability areas in which NGESO typically operate long-term contracts for procurement services are Stability, Voltage, and Restoration. Long-term contracts can be characterised as affording certainty and stability, an important precondition and enabling factor for investment. If a generation asset is contracted to provide an ancillary service such as inertia for a set number of years, the contract guarantees a level of stability and certainty of revenue. As such, developers can ensure a minimum return on investment for their assets through entering into these long-term contracts. This is particularly important for assets or investments with a high capex cost, where a longer-term contract reduces WACC and makes investment possible. Equally, such a contract provides certainty for NGESO, who can confirm adequate levels of inertia for the set period and ensuring future system operability. We identify assets that provide one service exclusively (e.g., synchronous condensers providing inertia) as also benefiting from procurement through long-term contracts. In the absence of a long-term contracts, there would be little to no incentive to build out and operate such an asset as there will be reduced guarantee on revenue, irrespective of system needs. Under these conditions, a long-term contract is more appropriate for ensuring system operability.

There is a potential risk associated with the locational dynamics of ancillary services and long-term contracts. Certain ancillary services, e.g., voltage management or SCL, are provided from assets that typically have long operational lifecycles and high capex costs but respond to locational needs. If these needs change, there is a risk of these assets becoming stranded. In this case, long-term contracts may facilitate the investment need for assets that provide ancillary services, but may also risk incentivising assets to build in the 'wrong' location.

Close-to-real-time markets typically encompass frequency and thermal constraint services that operate from real-time to day-ahead, although we can include procurement markets operating in the time frame of up to a month (e.g., Firm Frequency Response and certain voltage services[109]). In such markets, assets make offers to provide certain ancillary services and NGESO are able to procure as needed based on close-to-real-time network needs and conditions. Providing ancillary services through close-to-real-time markets is more suitable for operability areas which are more flexible and subject to short-term changes, e.g., frequency, where faults can develop and manifest in seconds. These markets also provide NGESO with a diverse set of assets and options from which to procure ancillary services, granting additional flexibility, theoretically at the lowest cost. Additionally, close-to-real-time markets are particularly important with regards to balancing services and thermal constraint, where having access to diverse options for service providers across different locations is key to ensuring effective system operation. Close-to-real-time markets allow NGESO to procure more accurate levels of service based on the latest data and system requirements, and while in theory this should lead to lower costs as a better service is procured, it can also lead to situations in which NGESO must procure services at a high cost. As is currently the situation under the BM, the costs of procuring services to resolve issues may be costly, however, to maintain the safe operation of the electricity system, NGESO must procure these services. This price volatility is a key characteristic that must be taken into account, and while short-term procurement may be typically cheaper as long-term risk is not priced in (due to external conditions), prices may spike on occasion and lead to periods where ancillary services are costly to procure. The latter occurs in the current BM through which NGESO resolves thermal constraints.

An optimal system should include both long-term contracts and close-to-real-time markets, where investment in assets that provide 'certainly needed' ancillary services are de-risked through contracts, and fluctuating ancillary services are procured in the short-term. Relying solely on long-term contracts risks over/under procurement of ancillary services if conditions change beyond what was forecast, while using exclusively close-to-real-time markets may lead to unsustainable price volatility. A contract could combine elements of both procurement strategies (e.g., a small availability payment for assets able to provide a service combined with a close to real-time market which pays for utilisation). This could be particularly suitable for assets such as offshore wind which may be able to provide system services with small adjustments, but cannot guarantee constant availability. Similarly, from a whole system cost perspective a combination of approaches may be suitable for some services, whereby a

[109] NGESO (2023) Market Roadmap 2023

baseline need could be met through long-term contracts, with close to real-time markets used for daily and/or seasonal variation.

Our research supports the recommendation that NGESO continue using a combination of both methods to procure ancillary services from a diversity of assets to ensure the safe and reliable operation of the electricity system. Conditions such as degrees of certainty regarding the optimal location and level of service, the types of assets used to deliver services, whether services require new assets to be built, and long-term revenue sources, can be used to assess whether long-term contracts or close-to-real-time markets are more appropriate.

6.2 TOR Question 8

What is the case for local ancillary markets for the provision of services such as frequency response, reserve, and inertia etc., to both local and national systems?

At present, ancillary services are managed at the national level by NGESO, who as electricity system operator have responsibility for ensuring the operability of the network. DNOs currently have a limited engagement with operability and ancillary services beyond thermal constraints and post-fault network management and restoration. However, as the electricity system decarbonises and DNOs make the transition to DSOs, there is potential scope for greater operability services being provided and ensured at a more local level; this is an option being considered within REMA.

The key rationale underpinning a move to local ancillary markets stems from the increasingly locational nature of operability and the electricity system more broadly. Changing patterns of consumer demand, in particular, the electrification of heating and transportation, are increasingly leading to scenarios where there is a contrast between local and national signals, and as such, there may be operability requirements that vary significantly across different locations. In this latter scenario, providing ancillary services and ensuring operability locally may be more effective and efficient, taking pressure off NGESO at the national level. For example, procuring reactive power services close to where there is a voltage constraint is more efficient as reactive power becomes less effective with distance[110]; it is also more cost-effective as less reactive power will be lost and so less will have to be procured in the first instance. DN(S)Os, who would likely take responsibility, would have greater visibility of local operability requirements. An example of what such a scenario would like would be the expected increase in DERs across GB, where operability requirements, such as balancing and voltage control, will increase in complexity as a greater number of small-scale generation assets (roof-top solar etc.) are connected[111]. A DN(S)O would likely have greater access to data of such resources and therefore be better placed to more quickly and accurately respond to challenges. Additionally, local ancillary markets allow smaller-scale generation sources such as batteries,

[110] NGESO (2018) Operability Strategy Report
[111] TradeRES (2020) Design of ancillary service markets and products: Challenges and recommendations for EU renewable power systems

wind, and solar assets, to participate in providing operability services. Currently, ancillary services are typically procured from large, centralised thermal assets but the procurement of ancillary services from diverse local sources could lead to greater cost-effectiveness through greater competition. The potential for a greater diversity of assets used for ancillary services was highlighted during our stakeholder engagements, where NGESO was criticised for dispatching high carbon thermal assets at a higher cost and 'out of merit', rather than using low carbon assets to address operability issues[112].

We identify system flexibility as an adjacent, but relevant area due to its ability to assist with managing operability issues, in particular adequacy and thermal constraints. In our Flexibility in Great Britain report (2021)[113], we found that system flexibility will increasingly be deployed at a more local level. Initiatives such as the DFS scheme are designed to help with managing peak demands during periods of system tightness but also aid in balancing and contributing to system stability (through frequency response and reserve etc.). This also relates to the proliferation of DERs and the increasing diversity of variable electricity generation sources.

Despite the theoretical increase in efficiency and cost-effectiveness that can be gained through local markets for ancillary services, our work has identified significant implementation challenges that may undermine any potential gains. If local markets are created with DN(S)Os taking responsibility, we foresee a significant level of inconsistency between levels of readiness, leading to varying degrees of successful management between locations. This problem is exacerbated by the scale of operability issues between locations, where challenges can be much greater, and therefore, more difficult to manage, between different districts.

Stakeholders have raised concerns regarding the current levels of data transparency and communication between NGESO and DN(S)Os, which was noted as hindering the management of operability at the local level; without access to granular data on network conditions and supply/demand dynamics, delivering accurately targeted ancillary services will be challenging. A move to local ancillary markets would require a greater integration of data, and given NGESO would likely need to retain a central co-ordinating role in some capacity, this as a potential issue. It is also important to ensure conflicting signals aren't given to assets operating in both DN(S)O and NGESO markets.

An additional complication arises over determining the optimal size of the local markets; taking into account the wholesale changes proposed within REMA, we expect markets that are zonal in size to be most appropriate as nodal markets are likely too small to be feasible. However, different operability areas can benefit from differently sized markets, where, for example, an ancillary service market for reactive power and SCL is more suited to a smaller market in

[112] Presented with this challenge through engagements, NGESO stated that (almost) all dispatches could be fully explained as efficiently resolving the operability requirements of the time. This gap in trust and understanding highlights the need for greater simulation capability, e.g. through the virtual system, and education of stakeholders.
[113] Carbon Trust (2021) Flexibility in Great Britain

comparison to managing thermal constraints or frequency which operate at bigger geographies and so may benefit from a larger pool of services.

In considering a potential move to local ancillary markets, questions must be asked on capability and efficiency: who is best placed to procure ancillary services, and where is it most efficient to provide them? This is an area of work in which more research is required to determine the case for local ancillary markets, and if so, what such markets would look like. A key conclusion drawn from stakeholder engagement is that it was thought that DSO-led markets were valuable for services which affected the DSO directly, for example thermal constraints, but that balance between local and national signals and services must be maintained.

6.3 TOR Question 9

How would co-optimisation of ancillary services with energy dispatch work under: a) a central dispatch model with a single national wholesale price, and (b) a Locational Marginal Price model, also with central dispatch?

The GB electricity system operates under a self-dispatch model, in which electricity operators and buyers engage in bilateral contracts with limited input from NGESO until Gate Closure (up to 1 hour ahead of Settlement Period), at which point, NGESO take actions under the BM in order to balance supply and demand across the network, taking into account its physical limitations (thermal constraints)[114]. While this model remedied problems of system gaming[115],[116] from generators that were present under the previous central dispatch model ('the pool'), it is increasingly suggested that self-dispatch may not align with the wholesale market reforms suggested within REMA, in particular, nodal pricing reforms. Under central dispatch, generators would inform NGESO of the prices at which they will potentially supply or consume electricity. NGESO would then calculate the most cost-efficient way of matching supply and demand while taking into account system constraints and operability, dictating a planned schedule for each generator to follow. Within each period, NGESO would also determine the system price and have access to technical plant information.

Central dispatch could be a more attractive means of market organisation in a decarbonised system as it enables greater efficiency in network balancing and operability, as well as facilitating enhanced transparency and positioning of low carbon assets[117]. Under a centrally dispatched model, NGESO would have more granular and accurate data and this enhanced network visibility can allow for operability challenges to be identified and therefore resolved sooner, lessening the magnitude of faults and reducing the level of action needed. This model

[114] UK Government, Appendix 5.1: Wholesale Electricity Market Rules
[115] The natural monopoly of PowerGen and National Power as effectively the two competing generators led to manipulation of the Pool Selling Price, in which non-binding offers were withdrawn at short notice leading to the dispatch of more expensive plants at a higher system marginal price.
[116] Green, R. (1999) Draining the Pool: the reform of electricity trading in England and Wales, Energy Policy
[117] NGESO-Baringa (2023) Assessment of Investment Policy and Market Design Packages

can also reduce the scale of costs as NGESO will also benefit from having access to a greater diversity of assets, leading to more effective targeting of operability issues with potentially more challenges being resolved using low carbon generators and capabilities. It has been noted that central dispatch would likely involve assets and resources being scheduled further in advance of real-time. In such a scenario, there is greater 'whole systems' visibility that would allow NGESO to better position low carbon assets where they see spare capacity. We see this whole systems perspective as an enabling factor that can be beneficial to the decarbonisation of the GB electricity system.

Central dispatch is however recognised as more inflexible due to the individual commitments as dictated by the systems operator; the 'running order' of generators and supply is generally fixed and trading or adjustments between producers is discouraged[118]. This would likely be the case under both a national wholesale price model and nodal pricing models. The Carbon Trust also foresee potential problems associated with potential gaming of the system under a central dispatch model. Central dispatch algorithms can be seen as 'black box' models unless there is adequate transparency of decision making[119], and this transparency is important for NGESO to be trusted and seen as fair, as well as encouraging long-term investment signals. On the other hand, if the decision-making process is well-understood, generators may be able to predict procedures, gaming the system to gain an unfair advantage. Any move to central dispatch would require significant upgrades in the technical and IT systems of the NGESO's control room in order to effectively manage and coordinate dispatch; this was noted repeatedly during stakeholder engagements and was raised a key concern of a move to central dispatch. Despite this, stakeholders and industry players seemed confident in the ability of NGESO to engage with the improvements necessary. Additionally, a move to central dispatch will involve a lengthy and complex implementation process at a high cost, and this is a significant challenge to be overcome.

Our work has also considered the degree to which a move to central dispatch would allow for the co-optimisation of ancillary services, in which NGESO would procure ancillary services and energy together, rather than the in the current system in which they are procured separately. Assets would submit bids for both energy and ancillary services which they are able to provide, and markets would be cleared simultaneously. At the Carbon Trust, we foresee such a system working most effectively under a centralised dispatch model where NGESO have greater co-ordinating authority and visibility of network conditions and capabilities. Proponents of co-optimisation, including NGESO, argue that this is a far more effective means of procuring ancillary services as it reduces the administrative and technical burden on both the systems operator and assets[120]. On the contrary, co-optimisation of ancillary services under either wholesale pricing or nodal pricing is likely to be difficult due to numerous pricing considerations and the complexity of coordination and implementation, although there is already precedent for the stacking and provision of multiple services through the BM. A key question needing further

[118] Ahlqvist, V & Holmberg, P & Tangeras, T (2019) Central- versus Self-dispatch in Electricity Markets, Cambridge EPRG Working Paper in Economics, 1902
[119] FTI Consulting (2022) Operation market design: Dispatch and Location: Industry Workshop
[120] NGESO (2022) Net Zero Market Reform Phase 3 Conclusions

consideration is whether to have a combined price for all services, or multiple prices for different services.

NGESO are beginning to move towards co-optimisation of frequency response via its Enduring Auction Capability platform[121]. Under this development, where a unit can provide more than one service, it will be able to put itself forward to provide each service and a clearing algorithm will allocate the unit to the service it can best provide. Frequency Response and Reserve are potentially the operability area and ancillary service most amenable to co-optimisation due to its close relation with active power supply. Co-optimisation of energy supply with Frequency Response and Reserve is already observed in electricity markets in the US and is generally seen as a success[122]. Greater challenges of co-optimisation in other operability areas are probable, in particular, localised operability issues such as voltage, SCL, and thermal constraint. An additional challenge associated with voltage is its disconnect from the provision of 'real-power', making its procurement synchronous with energy supply challenging. Australia, which operates under central dispatch and zonal pricing, is currently considering an optional mechanism/market for co-optimising constraints through the concurrent consideration and dispatch of both energy and 'constraint relief' bids[123]. Our findings suggest that this space should be monitored for further developments, and we also see challenges around the co-optimisation of ancillary services operating over longer timeframes such as with restoration and inertia. These ancillary services are typically procured through long-term contracts that might not be interchangeable or harmonious with the wholesale market.

A key difference between national and nodal markets both operating under central dispatch could be in that assets may require different payment levels depending on the nodal price. However in both market structures, the ancillary services that are most amenable to co-optimisation remain the same: primarily frequency, voltage, and inertia as these are provided by generators during normal operation and are unlikely to be significantly impacted by pricing.

[121] NGESO (2023) Enduring Auction Capability
[122] Pollit, M. G & Anaya, K. L (2019) 'Competition in Markets for Ancillary Services? The implications of rising distributed generation', University of Cambridge EPRG Working Paper 1928
[123] Market Wide Solutions & Clean Energy Council (2022) The Modified Congestion Relief Market model

7 Conclusions

Due to the large growth of renewable generation in GB, constraints at key boundary points are increasing. In 2020/2021, the B6 boundary was under constraint for 1,878 hours (21.4% of the time) resulting in constraint costs of £1.1billion for that year. Due to the availability of wind load factors in Scotland, this is set to increase into the 2030s as more renewable generation is deployed, even if NOA recommendations are applied. Currently, on the constrained side of transmission boundaries, the primary mechanism for constraint management is the curtailment of renewable generation sources through the BM. Flexible demand that can absorb or utilise otherwise curtailed energy is essential to assist in the cost-optimisation of curtailment actions and keep the carbon factor of the grid low.

This report therefore draws 6 key conclusions targeted at developing solutions for thermal constraints and extended periods of low renewable output, detailed below. These include:

1. Developing LDES capabilities behind constraint boundaries could allow for more effective management of both thermal constraints and extended periods of low renewable output.

2. Low carbon hydrogen could play a role in longer-duration flexibility but has a number of hurdles to overcome.

3. There is scope for addressing thermal constraints through ancillary services as a supplement or alternative to REMA wholesale market reform.

4. There is a role for ancillary services in creating operational signals for low carbon technologies able to meet system needs during extended periods of low sun/wind.

5. There is some scope for reforms to support schemes to meet the governments objectives on operability.

6. There is scope for ancillary service procurement reform to focus on net cost rather than gross cost.

7.1 Developing LDES capabilities behind constraint boundaries will allow for more effective management of both thermal constraints and extended periods of low renewable output

LDES has a key role to play in managing thermal constraints and extended periods of low renewable output, especially in longer periods of extended constraints where shorter-duration assets may reach maximum state of charge.

At a conceptual level, thermal constraints and Dunkelflaute periods have an inverse relationship that can be reconciled with long-duration flexibility. LDES technologies can charge up their assets on low or even negatively priced energy (if payments for constraint management are included) and discharge in periods of extended low renewable output where wholesale energy prices should reach their theoretical maximum.

While variations in wholesale market prices should be enough to incentivise this behaviour, operating LDES at the most extreme price highs and lows would require asset operators to accurately predict constraint/Dunklefluate periods, and investors would be exposed to the risk of forecasting errors. Moreover, the traditional operating model of energy storage has been to charge/discharge as much as possible to maximise the assets rate of return. Therefore, there may be a requirement for stronger, co-optimised price signal that both incentivises effective asset operation in accordance with operability requirements for thermal constraints and for capacity adequacy, the Storage Level Signal (SLS).

LDES operators could receive availability payments to retain a state of charge at the NGESO's discretion, assuring assets are available to respond to system needs at any given time. An alternative or supplementary approach would be the introduction of locational (nodal or zonal) pricing, which would create larger price deltas behind constraint boundaries, as the marginal cost of constraint management would be reflected in the wholesale price. While locational pricing with central dispatch may reduce the need for ancillary services for constraint management, some form of ancillary services may be needed to resolve any residual constraints. Additionally, locational wholesale pricing does not address adequacy needs, and some form of ancillary service would still be needed to provide certainty to system operators. These services would need to be designed in a way that compliments the planning, construction, commissioning, and operational lifetimes of LDES technologies.

While one or both options would increase the commercial rationale of LDES, as with offshore wind, there will be a need for support schemes to protect investor revenues as technologies and business models mature. While reforms to support schemes under consideration in REMA may address problems, the current iterations of the CM will need reform of supplementation to incentivise investment in LDES. Based on the analysis in Chapter 4, the Carbon Trust recommends that a cap and floor approach could be an effective means of de-risking investment while assuring assets are operated based on operational signals in the wholesale market.

7.2 Low carbon hydrogen could play a role in longer-duration flexibility, but has a number of hurdles to overcome

Low carbon hydrogen could assist in both thermal constraint management and securing capacity. Many of the planned hydrogen clusters are located close to key constraint boundaries and, during periods of constraint, could be used to produce hydrogen, either being utilised as an energy source or stored for future periods of low renewable output.

Hydrogen, however, has a lower commercial maturity at the scales required for strategic LDES compared to other technologies and more work is required to advance hydrogen capabilities.

Additionally, the round-trip efficiency of hydrogen when used for LDES energy storage can range between 18% to 46%[124], much lower than other forms of LDES which can have round trip efficiencies of 75% to 90%.

7.3 There is scope for addressing thermal constraints through ancillary services as a supplement or alternative to REMA wholesale market reform

While significant network investment is needed to meet the UK government's targets on climate change, there will always be a cost-optimum level of constraint management.

There are options outside of ancillary services that could help cost-optimise the level of network investment required, including introduction of dynamic line ratings and increasing cable capacity through improved emissivity.

The de facto mechanism for controlling thermal constraints (the BM) can be costly and carbon-intensive. NGESO has historically favoured larger assets for a complex set of reasons, including wider contribution to system operability, ease of dispatch, and lack of visibility of smaller assets. Reforms to the BM under consideration as part of the REMA consultation (BAU+) that alleviate some of these concerns could reduce the need for ancillary services for thermal constraint management.

NGESO's Constraint Management Intertrip and Constraint Market Services both reduce the costs associated with constraint management and encourage more effective use of existing network capacity. However, there is a gap regarding effective dispatch of assets pre-BM and incentivising DSR/storage. More work is required to understand the impact they might have on net curtailment/constraint costs.

A 'Strategic Cycling' service would help cost-optimise the most costly curtailment actions over longer constraint periods, but would do little to mitigate total constraints or utilise curtailed energy. While this service would reduce the gross cost of constraint, the impact on net cost would likely be minimal, though it could serve as a 'disciplining' signal to CfD and battery generators to align their operation closer to system needs.

A 'short-term constraint reserve service' could allow for effective management of short constraint periods with demand assets, enabling increased headroom at transmission boundaries which could theoretically reduce constraint/curtailment costs.

[124] University of Strathclyde, Round-trip Efficiency

There is a gap in operational and investment signals for demand assets (e.g., LDES) capable of absorbing curtailed energy in extended constraint periods. More work is needed to develop service and procurement specifications that are beyond the scope of this project.

Ancillary services for sustained constraint periods have synergise with requirements for capacity adequacy, providing a potential revenue stream for nascent long-term energy storage technologies, including hydrogen storage.

If the appropriate digital infrastructure is in place, a move to nodal pricing (LMP) would allow for more accurate and timely operational signals, improving the efficiency of dispatch. This would likely result in more effective management of thermal constraints.

A move to nodal or zonal pricing is unlikely to significantly impact generation and demand geographies in GB, which are more heavily influenced by a number of independent factors such as availability of load factors. Some smaller assets, such as battery storage and solar, are more flexible regarding location and may be able to respond to locational price signals

Nodal and zonal pricing would likely improve investment/operational signals for LDES which could allow for more effective management of thermal constraints and utilisation of curtailed energy.

A move to central dispatch would optimise the procurement and dispatch of ancillary services and the BM, as NGESO would be able to dispatch assets unilaterally. However, it is probable that some aspect of separately procured real-time balancing would still be needed to resolve forecasting errors or intervene in specific localities where constraints arise.

7.4 Ancillary services have a role in creating operational signals for low carbon technologies able to meet system needs during extended periods of low sun/wind

The technologies that can meet system needs in extended periods of low sun/wind include interconnectors that can import energy from areas of plentiful resource, both low and high carbon dispatchable generation and LDES (including hydrogen).

For interconnectors, the operational signals come from differences in wholesale market price between connected countries. Since 2014, UK interconnector investment has been supported by revenue cap and floor regimes.

Ancillary services will make up a significant portion of revenue for LDES, but are not enough in isolation to sufficiently de-risk investment to the levels needed to encourage substantial deployment of LDES.

Revenue cap and floors would likely be a suitable investment support scheme for LDES, as it complements the operational profile of storage, protects end consumers from investors earning excess returns effectively de-risks investment.

While the CM has been successful (along other market signals, to bring forward short-duration storage assets, the procurement horizons (max four years in advance) are not far enough in advance for LDES technologies, which typically have longer construction times than CM assets.

Wholesale market price deviations will be effective at providing day-to-day operational signals for LDES, but ancillary services may be needed to ensure that assets retain sufficient headroom to absorb energy during extended periods of constraints as well as retaining sufficient energy to be available in extended periods of low wind/sun.

Further research is required into a Storage Level Service for LDES which provides availability payments for standby capacity of long periods of low wind/sun. The procurement levels for this service could vary depending on seasonal variations in adequacy risk.

Ancillary service contracts for LDES should complement asset lifetimes and be procured in advance of investment decisions. This is less important if the asset has a revenue cap and floor, as investor returns would be de-risked irrespective of ancillary service design.

7.5 There is some scope for reforms to support schemes to meet the governments objectives on operability

The key reason for less-than-optimal dispatch of CfD generators for flexibility and/or ancillary services is that NGESO considers the gross cost rather than the next cost to consumers when dispatching the services.

Most of the proposed CfD reforms will lead to a lower gross cost of dispatching CfD generators for flexibility and/or ancillary services, but will have very little impact on net costs and can have unintended consequences such as leading to higher cost of capital or possibly excess returns to investors.

Most of the proposed CfD reforms would only apply to future CfDs and not resolve the dispatch distortion for already contracted CfDs. Retrospective changes that lead to greater investor risk are unlikely to be possible.

A combination of greater monitoring and reporting of CfD generator bidding behaviour alongside procurement reforms for BM and ancillary services, including changes to the ORPS pricing methodology, would reduce or remove the CfD dispatch distortions for both existing and future CfDs.

If reforms are still considered necessary after the reporting and BM/ancillary services procurement has been reformed, a limited, 'deemed generation' option under which any output lost due to participation in flexibility and/or ancillary services still attracts a CfD difference payment, similarly to allowance for 'applicable balancing services' under the BM.

NGESO's Virtual Energy System work should be accelerated, and the educational element expanded so as to build transparency and trust amongst stakeholders. This will be important to facilitate any future reforms.

The changing nature of capacity provision means that the CM is not efficient at encouraging investment in LDES – a separate investment signal will likely be needed.

The use of multipliers for CM would be similar to bundling ancillary and/or flexibility services. This can lead to inefficient procurement, added complexity and is out-of-step with NGESO's aim to create separate markets for services.

CM and ancillary services participation and qualification should be standardised. A 'platform' approach with reduced administration could increase competition and reduce costs for ancillary services.

7.6 Ancillary service procurement strategies

7.6.1 Long vs short-term contracts

Ancillary service contracts for LDES should complement asset lifetimes and be procured in advance of investment decisions. This is less important if the asset has a revenue cap and floor, as investor returns would be de-risked irrespective of ancillary service design.

Procurement of ancillary services through close to real time markets affords NGESO greater flexibility in addressing operability issues. In areas such as frequency and thermal constraint management, having access to a diverse set of options allows more accurate and faster procurement of service, this may confer cost-savings in some instances.

Longer-term contracts have the advantage of providing certainty of revenue, but procurement volumes may be more inaccurate as a result. Volumes will often be higher than needed to assure system security, which could lead to higher costs.

Utilising a combination of both long-term contracts and close-to-real-time markets is likely the most effective means of procuring ancillary services, because of conditions such as the characteristics of the ancillary service in question and the technologies providing it.

7.6.2 Local ancillary service markets

As the electricity system and associated operability challenges become increasingly localised, there is greater scope for providing and procuring ancillary services through local markets.

Operating ancillary services at a local level may relieve pressure at the national level and may lead to greater cost-effectiveness due to enhanced data visibility and more accurate procurement levels/strategies.

In operability areas such as voltage control, there are likely to be significant efficiency savings due to the locational nature of voltage and reductions in losses associated with transportation over long distances.

The management of local ancillary service markets would likely be undertaken by DNOs, who are expected to have an increased responsibility in the management of the energy network moving forward.

There are likely to be implementation challenges associated with local ancillary markets, including different levels of readiness between districts, and differing needs and challenges based on location that may hinder efficient operation. Additionally, there are concerns regarding data availability and transparency between the local and national authorities.

7.6.3 Co-optimisation of ancillary services

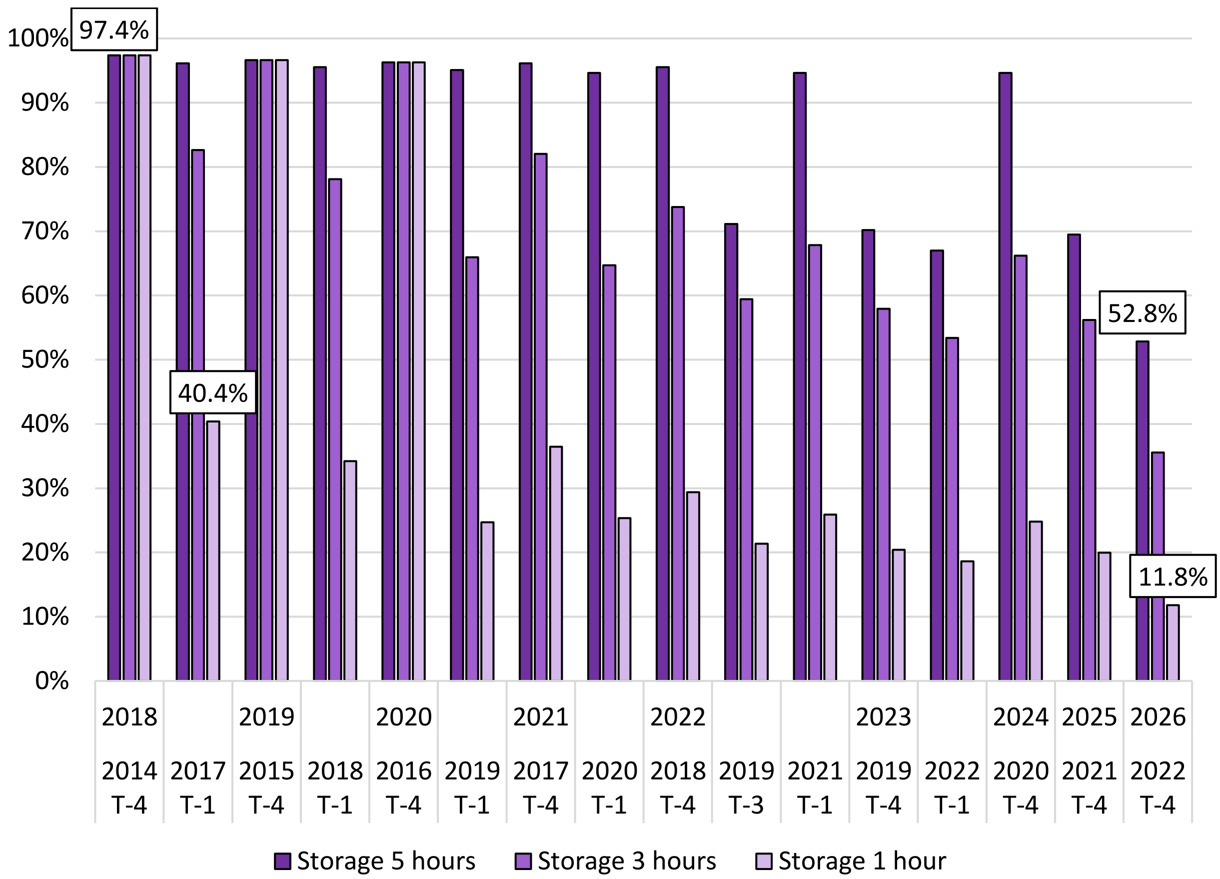
Central dispatch could be introduced under both a national pricing market and nodal pricing market; a move to a form of nodal pricing would likely require central dispatch, however, due to the increased complexity of pricing and associated need for greater coordination.

Central dispatch could lead to greater efficiencies in network balancing and operability as NGESO would have greater visibility of network conditions and therefore be able to take faster and more targeted actions to resolve operability challenges. It may also allow for the more effective positioning and utilisation of low carbon assets that are able to provide ancillary services, aiding in the decarbonisation of the electricity system.

Central dispatch is recognised as more inflexible as the running order and decision on when to operate is dictated by the systems operator, who discourages changes or trading of commitments. Additionally, a move to central dispatch requires a significant increase in the technical and data capacity of NGESO, which at present, may not be sufficient. These advantages and disadvantages of central dispatch would likely remain the same under both national and nodal pricing. Central dispatch would also allow for the co-optimisation of ancillary services alongside energy dispatch, under either a national wholesale or nodal pricing model. Beyond frequency response and reserve. However, there are a limited number of ancillary services that could be easily co-optimised. These include services to address locational challenges such as voltage, SCL, and thermal constraints.

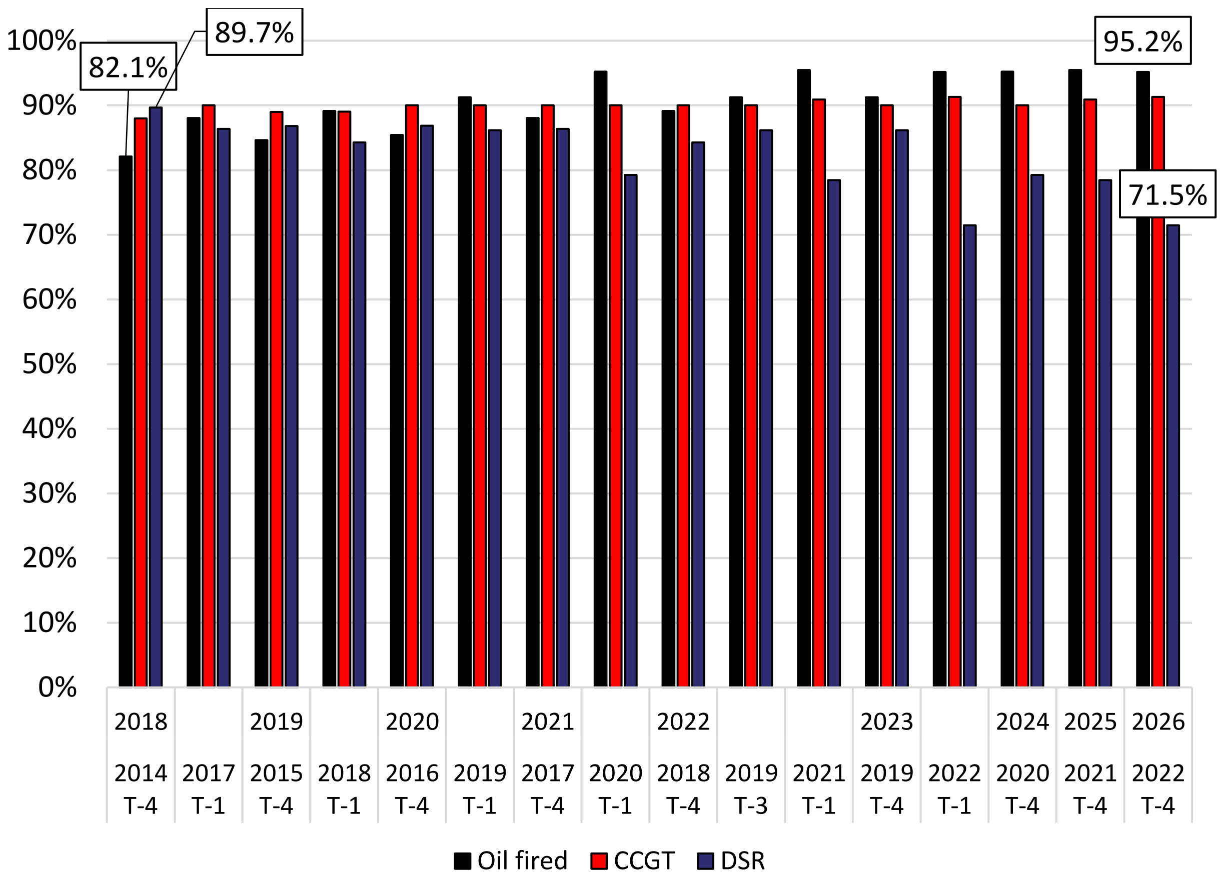
8 Appendix

Figure 14 - Capacity market derating factors for selected storage durations



Actual storage durations range from 0.5 hours to 9+ hours in 0.5 hour increments. Storage for T-4 2018 and 2019 were not separated out in technologies, but according to NGESO communications were pumped storage. However, these are still kept in for completeness.

Figure 15 - Capacity market derating factors for oil fired generators, CCGTs, and DSR



De-rating factors for most technologies reflect past seven (7) years of data on reliability.

Figure 16 - Capacity market derating factors

Auction year	2014	2015	2016	2017 T-1	2017 T-4	2018 T-1	2018 T-4	2019 T-1	2019 T-3	2019 T-4	2020 T-1	2020 T-4	2021 T-1	2021 T-4	2022 T-1	2022 T-4
Oil fired	82.10%	84.61%	85.44%	88.04%	88.04%	89.13%	89.13%	91.26%	91.26%	91.26%	95.22%	95.22%	95.47%	95.47%	95.18%	95.18%
OCGT	93.61%	94.54%	94.17%	94.81%	94.81%	95.14%	95.14%	94.98%	94.98%	94.98%	95.22%	95.22%	95.47%	95.47%	95.18%	95.18%
Recips	93.61%	94.54%	94.17%	94.81%	94.81%	95.14%	95.14%	94.98%	94.98%	94.98%	95.22%	95.22%	95.47%	95.47%	95.18%	95.18%
Nuclear	81.39%	82.31%	84.36%	85.24%	85.24%	84.20%	84.20%	81.22%	81.22%	81.22%	81.43%	81.43%	80.44%	80.44%	78.25%	78.25%
Hydro	83.86%	84.87%	86.16%	87.92%	87.92%	90.09%	90.09%	89.65%	89.65%	89.65%	90.99%	90.99%	91.15%	91.15%	91.13%	91.13%
CCGT 2017/18	88.00%	89.00%	87.60%													
CCGT 2018/19	88.00%	89.00%	88.00%	88.54%	88.54%											
CCGT 2019/20	88.00%	89.00%	89.00%	89.00%	89.00%	89.05%	90.00%									
CCGT	88.00%	89.00%	90.00%	90.00%	90.00%	89.05%	90.00%	90.00%	90.00%	90.00%	90.00%	90.00%	90.92%	90.92%	91.31%	91.31%
CHP/auto-gen	90.00%	90.00%	90.00%	90.00%	90.00%	90.00%	90.00%	90.00%	90.00%	90.00%	90.00%	90.00%	90.92%	90.92%	91.31%	91.31%
Coal	87.64%	87.86%	86.92%	87.58%	87.58%	86.56%	86.56%	85.81%	85.81%	85.81%	84.80%	84.80%	80.11%	80.11%	80.40%	80.40%
Biomass	87.64%	87.86%	86.92%	87.58%	87.58%	86.56%	86.56%	85.81%	85.81%	85.81%	84.80%	84.80%	88.55%	88.55%	87.99%	87.99%
EfW	87.64%	87.86%	86.92%	87.58%	87.58%	86.56%	86.56%	85.81%	85.81%	85.81%	84.80%	84.80%	88.55%	88.55%	87.99%	87.99%
DSR	89.70%	86.80%	86.88%	86.34%	86.34%	84.28%	84.28%	86.14%	86.14%	86.14%	79.21%	79.21%	78.45%	78.45%	71.45%	71.45%
Storage 0.5 hours	97.38%	96.63%	96.29%	21.34%	17.89%	17.50%	14.91%	12.26%	10.59%	10.21%	12.75%	12.38%	12.94%	9.98%	9.30%	5.95%
Storage 1 hour	97.38%	96.63%	96.29%	40.41%	36.44%	34.21%	29.40%	24.70%	21.36%	20.43%	25.32%	24.77%	25.87%	19.96%	18.60%	11.81%
Storage 1.5 hours	97.38%	96.63%	96.29%	55.95%	52.28%	50.00%	43.57%	36.96%	31.94%	30.83%	37.71%	36.97%	38.62%	29.94%	27.90%	17.77%
Storage 2 hours	97.38%	96.63%	96.29%	68.05%	64.79%	62.80%	56.68%	48.66%	42.53%	41.04%	49.17%	48.62%	50.63%	39.73%	37.02%	23.63%
Storage 2.5 hours	97.38%	96.63%	96.29%	77.27%	75.47%	71.96%	66.82%	58.68%	52.18%	50.51%	58.23%	58.78%	60.61%	48.97%	45.95%	29.58%
Storage 3 hours	97.38%	96.63%	96.29%	82.63%	82.03%	78.09%	73.76%	65.93%	59.43%	57.94%	64.70%	66.18%	67.82%	56.18%	53.39%	35.53%
Storage 3.5 hours	97.38%	96.63%	96.29%	85.74%	85.74%	81.57%	77.78%	70.38%	64.07%	62.77%	68.76%	70.98%	72.25%	61.54%	58.79%	41.11%
Storage 4 hours	97.38%	96.63%	96.29%	96.11%	96.11%	95.52%	80.00%	72.98%	67.04%	65.93%	71.35%	73.76%	74.84%	64.86%	62.32%	45.86%
Storage 4.5 hours	97.38%	96.63%	96.29%	96.11%	96.11%	95.52%	95.52%	75.03%	69.27%	68.16%	73.20%	75.79%	94.61%	67.45%	64.74%	49.48%
Storage 5 hours	97.38%	96.63%	96.29%	96.11%	96.11%	95.52%	95.52%	95.08%	71.13%	70.20%	94.64%	94.64%	94.61%	69.48%	66.97%	52.83%
Storage 5.5 hours	97.38%	96.63%	96.29%	96.11%	96.11%	95.52%	95.52%	95.08%	95.08%	95.08%	94.64%	94.64%	94.61%	94.61%	69.02%	55.81%
Storage 6 hours	97.38%	96.63%	96.29%	96.11%	96.11%	95.52%	95.52%	95.08%	95.08%	95.08%	94.64%	94.64%	94.61%	94.61%	95.25%	58.97%
Storage 6.5 hours	97.38%	96.63%	96.29%	96.11%	96.11%	95.52%	95.52%	95.08%	95.08%	95.08%	94.64%	94.64%	94.61%	94.61%	95.25%	61.95%
Storage 7 hours	97.38%	96.63%	96.29%	96.11%	96.11%	95.52%	95.52%	95.08%	95.08%	95.08%	94.64%	94.64%	94.61%	94.61%	95.25%	64.92%
Storage 7.5 hours	97.38%	96.63%	96.29%	96.11%	96.11%	95.52%	95.52%	95.08%	95.08%	95.08%	94.64%	94.64%	94.61%	94.61%	95.25%	67.99%
Storage 8 hours	97.38%	96.63%	96.29%	96.11%	96.11%	95.52%	95.52%	95.08%	95.08%	95.08%	94.64%	94.64%	94.61%	94.61%	95.25%	70.88%
Storage 8.5 hours	97.38%	96.63%	96.29%	96.11%	96.11%	95.52%	95.52%	95.08%	95.08%	95.08%	94.64%	94.64%	94.61%	94.61%	95.25%	73.85%
Storage 9 hours	97.38%	96.63%	96.29%	96.11%	96.11%	95.52%	95.52%	95.08%	95.08%	95.08%	94.64%	94.64%	94.61%	94.61%	95.25%	76.64%
Storage 9.5 hours	97.38%	96.63%	96.29%	96.11%	96.11%	95.52%	95.52%	95.08%	95.08%	95.08%	94.64%	94.64%	94.61%	94.61%	95.25%	95.25%
Storage 10 hours	97.38%	96.63%	96.29%	96.11%	96.11%	95.52%	95.52%	95.08%	95.08%	95.08%	94.64%	94.64%	94.61%	94.61%	95.25%	95.25%
Onshore wind								8.98%	8.20%	7.42%	8.01%	7.81%	7.81%	6.25%	8.20%	6.74%
Offshore wind								14.45%	12.30%	10.55%	12.11%	11.13%	11.33%	8.59%	11.33%	8.30%
Solar PV								2.34%	3.13%	3.22%	2.54%	2.34%	2.15%	3.32%	3.32%	4.98%

Source: Electricity Capacity Reports (raw data) Source: Electricity Capacity Reports (raw data)

This publication is available from: www.gov.uk/government/publications/review-of-electricity-market-arrangements-rema-technical-research-supporting-consultation