EXPLORING OPTIONS FOR CONSTRAINT MANAGEMENT IN THE GB ELECTRICITY SYSTEM:

THE POTENTIAL FOR CONSTRAINT MANAGEMENT MARKETS

A report for Scottish Renewables

January 2024

TEL the energy landscape

Author: Dr Simon Gill, The Energy Landscape

Contact details: simon@energylandscape.co.uk

Front cover image courtesy of Scottish and Southern Energy Networks

Table of Content

Executive summary

Background

Constraints happen when the electricity network is unable to safely and securely facilitate the outcome of the electricity market. This is particularly important where constraints lead to the curtailment of wind and solar generation and the need to replace that generation with alternatives, often fossil fuel generation, elsewhere in the country.

Constraints add costs to consumer bills and there is uncertainty over future constraints, both their volume and the cost of resolving them, which creates risk for future consumers. Constraint costs are expected to rise over the coming decade, although there is significant uncertainty over the scale of that increase.

Future constraint costs will be affected by the speed at which we deliver renewables, develop batteries, pumped storage, hydrogen electrolysis and other forms of flexibility, and the delivery or delay of new transmission capacity.

Even in a well optimised system, constraint costs will be higher in the future. Reducing the volume of constraints ultimately involves investment in network capacity or flexibility. As the fraction of renewables in our system grows, the level of economically efficient constraints – where the cost of constraints is balanced against the cost of investment to reduce them – is also likely to grow.

In the financial year 2022-23, constraint costs were £1.5 bn, under current arrangements these costs are socialised across demand which equates to around £5.70 / MWh. That adds about £15 / year to a typical domestic consumer's bill[1].

Figure ES1 shows the range of forecasts of constraint costs presented recently by National Grid Electricity System Operator (NGESO) alongside recent historical outturn. It shows that there is the potential for constraint costs to grow further during the late 2020s, potentially reaching as high as £3 bn before dropping back as new transmission capacity is commissioned in the early 2030s.

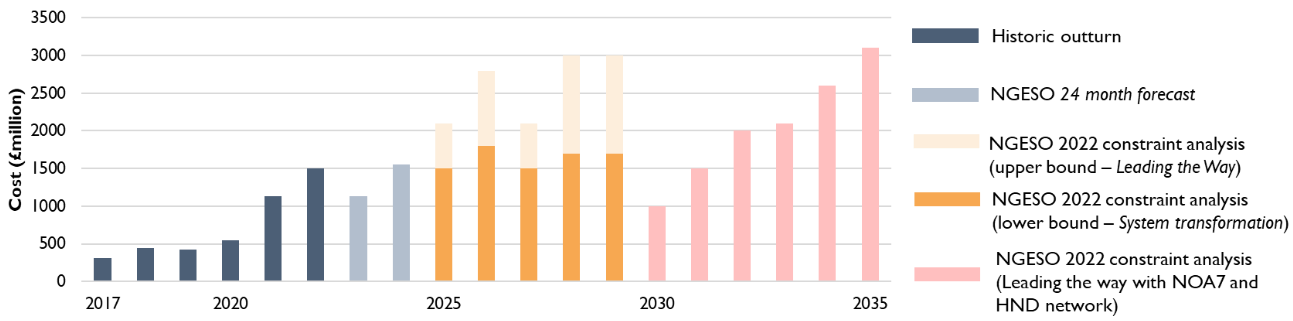


Figure ES1: Constraint costs in GB covering recent historical outturn and available forecasts out to 2035.

Today constraints are largely managed by two mechanisms:

- Firstly, on planning timescales, constraints are reduced through investment in new transmission capacity, with decisions made on timescales of a decade or more.
- Secondly, on operational timescales, constraints are resolved through the use of the Balancing Mechanism (BM) to adjust the electricity market outturn on timescales of an hour or less.

[1] Historic constraint costs available here. https://www.nationalgrideso.com/data-portal/constraint-breakdown. Calculation assumes a typical domestic consumption value of 2,700 kWh / year: https://www.ofgem.gov.uk/information-consumers/energy-advice-households/average-gas-and-electricity-use-explained

Transmission planning has, for the past decade, been carried out through NGESO's annual process: producing credible future market led scenarios in the *The Future Energy Scenarios* (FES), identifying transmission system needs for each scenario in *The Electricity Ten Year Statement* (ETYS), and a cost-benefit analysis of specific transmission investments in *The Network Options Assessment* (NOA). The NOA balances the cost of a particular project against the benefits it is expected to create, the most important of which is the reduction in constraint costs it will facilitate. Those transmission investment decisions have explicitly considered the uncertainty around future constraints through the use of all four of the FES scenarios and a 'least-worst-regrets' approach.

On operational timescales, the majority of constraints are resolved following 'gate closure', one hour ahead of delivery, using the BM. This tends to involve instructing wind and solar power stations to reduce their output behind a constraint, and instructing flexible units in front of a constraint to increase their output.

Over the past few years NGESO has been increasingly using some actions ahead of gate closure. This includes electricity market trading, countertrading over interconnectors, and the use of non-market regulatory frameworks for limiting interconnector flows. However, this remains an ad-hoc approach without a well defined and socialised strategy.

In the intervening period between transmission planning decisions and gate closure there are no formal mechanisms to allow NGESO to react to changes in forecast constraints.

This report recommends that a portfolio approach to constraint management should be explored with tools available to NGESO over all relevant timescales. Transmission planning and the BM would become the bookends to the portfolio. Within those bookends, both long-term and short-term constraint tools should be developed including constraint management markets, competitive tendering, and regulatory approaches.

One design criteria for a constraint management portfolio is that it should be capable of evolved over time. Straightforward reforms can be implemented quickly whilst more complex reforms should built on those initial interventions later.

A it transitions into the Future System Operator (FSO) NGESO needs a clear articulation of the objective of constraint management that goes beyond a high level statement to minimise constraint costs.

This report recommends that the objective of a constraint management portfolio should be to maximise consumer value, including a balance between minimising expected costs under specific scenarios and explicitly managing the risk associated with future uncertainty.

There are a number of pathfinders and exploratory projects which could form the basis of new tools to integrate into a constraint management portfolio. One example is a local constraint management market that has been operating in Scotland since April 2023, procuring day-ahead and intra-day constraint management.

The potential for constraint management markets

This report discusses a number of competitive mechanisms that can be used as part of a constraint management portfolio. These mechanisms are collectively referred to as Constraint Management Markets (CMMs) and are contrasted with non-market regulatory approaches. CMMs can include tendering for long-term contracts and short-term 'spot market' auctions.

The report uses the following working definition for a CMM:

> Any market-based approach operating ahead of gate closure through which the FSO can buy or sell flexibility, or related products such as 'availability' or 'options' in order to relieve constraints on the transmission network. This includes both downward flexibility behind a constraint and upward flexibility in front of a constraint. They can include options for contracting for constraint management over several months or years (long-term CMMs) or running auctions for constraint management days, hours or minutes ahead of delivery (short-term CMMs).

In addition to the definition there are several characteristics that are likely to be present in any CMM design:

- A central buyer market with the FSO defining the quantity and characteristics required.
- Non-mandatory participation by market participants, following the approach used with ancillary service products.
- Open to a wide range of potential providers and not unduly restricted by size and type of provider (e.g. it should not require being a BM participant).

CMMs can provide options to manage both turn down costs incurred in reducing generation or increasing demand behind a constraint, and turn up costs incurred in increasing generation or decreasing demand in front of a constraint. Today, turn up costs paid to schedulable generators represent the vast majority of constraint costs, and increases in turn up costs on a £/MWh basis are the main reason for the increase in constraint costs over the past few years (see box ES1).

Box ES1: Rising constraint costs - volumes and prices in front and behind a constraint

Constraints lead to curtailment of cheap generation, particularly wind and solar, behind an export constraint and more expensive generation, often from fossil fuel power stations, being used to replace the curtailed output.

There are costs associated with turning down renewables which are largely related to lost support mechanism payments. These are called turn down costs. There are also costs associated with turning up the replacement generation. These are called turn up costs.

Total constraint costs have risen significantly over the past few years. This has been driven almost exclusively by increases in turn up prices. Constraint volumes have stayed largely constant and the cost per MWh of curtailing renewables has also remained steady. However, typical turn up costs have increased several times. Overall, the average cost of relieving constraints has risen from £109 / MWh in 2018-19 to £366 / MWh in 2022-23 with almost all that cost increase relating to turn up costs[2].

As gas prices fall from the peak of 2022, unit costs and total costs are expected to fall again in the short term. In the slightly longer term, volumes are expected to rise, leading to a further increase in overall constraint costs.

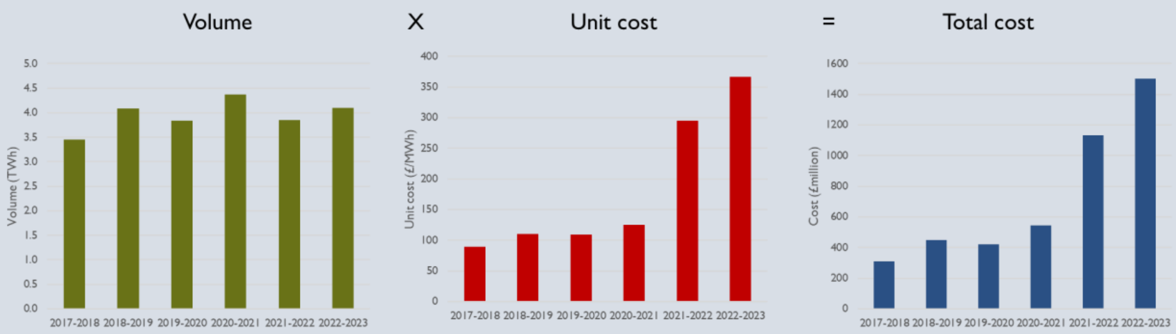


Figure ES2: The components of constraint costs in recent years. Volume on the left, average cost per MWh of constraint in the middle, and total costs on the right.

[2] https://www.nationalgrideso.com/data-portal/constraint-breakdown

Long-term constraint management markets

There is the potential to develop long-term CMM contracting, using competitive auctions or similar market-based approaches. Such contracts could be offered from a decade or more ahead down to a month ahead and could be used in situations where significant constraint volumes and costs are forecast. They would be particularly useful where, following uncertainty analysis, even the lower bound estimates of constraint costs are significant.

Such contracts could be structured to allow the FSO to procure availability based on a specified number of hours of service provision in a particular year. Close to real time, for example day ahead, the FSO would nominate the specific hours in which availability is required. Finally utilisation costs are only incurred if providers are ultimately dispatched within those availability windows.

This approach allows the FSO to lock in the costs for specific volumes, durations, and depths of constraint management at the time of contracting, whilst only committing to the availability component of the cost up front and if needed.

Short-term constraint management markets

Day-ahead and intra-day CMMs are also likely to be valuable. They would allow the FSO to contract with flexibility ahead of gate closure once there is a degree of certainty over the need for constraint management during particular settlement periods.

There are a number of models which could be used for short-term CMMs. One would involve procuring firm-response day-ahead or intraday, putting an obligation on a provider as soon as the contract is awarded. This would suit providers who need to schedule their activity with several hours notice or those that would be unable to make adjustments to their dispatch on timescales of less than an hour.

A second option is for the FSO to procure availability day-ahead (or intraday) with dispatch at or after gate closure. This would align with the structure of other ancillary services, such as the dynamic frequency response suite and the new balancing reserve service. It would also allow the FSO flexibility in the actions available to it as certainty in the scale and timing of constraints grows between day-ahead stage and delivery.

It is likely that both forms of short-term CMM – firm-response and availability / utilisation – would be of value in enabling the FSO to access provision from a wide range of potential providers, whilst balancing uncertainty, risk and cost.

Areas where greater understanding is needed

Three key challenges need better understanding in order to develop a constraint management portfolio: constraint forecasting; interactions between CMMs and other parts of the electricity market including the BM and wholesale market; and the integration of interconnectors into a constraint management portfolio.

Forecasting constraints

Constraint forecasting is challenging. For example, on planning timescales, in March 2019 NGESO forecasts the constraint element of Balancing System Use of System Charges for the year 2021/22 at £542million whilst the outturn cost was £1,071million[3]. On operational timescales, forecasting constraints during specific settlement periods means bringing together forecasts of wind and solar resources, other weather factors such as the impact of temperature on demand, network availability, interconnector flows, likely dispatch of schedulable generators and the impact of other system operability factors such as the need for voltage support and inertia.

Understanding our ability to forecast constraints and the uncertainty inherent in those forecasts, is central to efficient planning and operation of the power system.

Forecasting of annual volumes and costs with look ahead times of at least a decade is one of the main inputs to the NOA cost-benefit analysis. And forecasting of the volume and cost during specific settlement periods is used today to inform NGESO's existing energy trading and interconnector actions at day-ahead and intraday stage.

However, the process for conducting constraint forecasts and their accuracy is not well understood across the sector. This report recommends that NGESO should publish more information on constraint forecasts, including the volumes and costs used in each year of its NOA calculations (or, in future, similar calculations that will be carried out through the Centralised Strategic Network Plan (CSNP) process).

On operational timescales, there are concerns that publishing constraint forecasts will allow market participants to game opportunities across NGESO trading, wholesale market activities and the BM. However, without more information in the public domain, there is unlikely to be significant progress in improving forecasts or developing consensus across the sector on the most appropriate way to manage constraints.

The design of short-term markets should clearly take account of the level of confidence around constraint forecasts. For example, where there is significant uncertainty around constraint volumes at day-ahead stage, procurement of availability is likely to be a more appropriate method of managing consumer costs and risk than procurement of firm-response.

Market interactions

Interactions between wholesale markets, the BM and CMMs have the potential, if not identified and well managed, to lead to poor outcomes for consumers. The form of these interactions can range from legitimate trading strategies, through 'gaming', to leveraging of market power and illegal market manipulation.

We already see the opportunity for interaction between the wholesale market and BM, and we have developed regulatory tools to mitigate these issues to some degree. For example, the Constraint Management Licence Condition placed on generators provides a route to fine generators who are found to be unduly profiting from a transmission constraint.

We need to learn from this experience and CMM designs need to respond to the potential for each of these forms of interaction.

It is also important to remember that there is no perfect market design for managing constraints and that a 'good' market design will involve balancing the upsides and downsides of different options. Ultimately

[3] Note that these forecasts include a significantly wider range of actions within the definition of constraints than the values quoted in Figure ES1 above which only focus on thermal constraints. Forecast for 2020/21 made in March 2019: https://www.nationalgrideso.com/document/141946/download Outrun for 2020/21: https://www.nationalgrideso.com/document/284216/download

the solution chosen should be one that maximises value. That may mean accepting the potential for some negative interactions between CMMs and other components of the market design if the result is the ability to delivery significant reduction in overall constraint costs.

The preliminary explorations carried out in this report suggest that versions of constraint management markets which procure availability / utilisation rather than firm response are most promising for avoiding negative interactions between markets. It also suggests that markets designed around the explicit encouragement of stacking revenues between CMMs and wholesale markets could avoid some of the pitfalls that previous studies into gaming have identified. And that the use of long-term contracts which fix prices well in advance of delivery provide an additional step to decouple bidding strategies across markets.

This report has not had the scope to investigate market interactions in significant depth, rather these conclusions are presented in order to stimulate further analytical work focused on providing a clearer picture of the value and trade offs associated with using constraint management markets.

Interconnectors

Interconnectors represent a unique challenge. Due to GB's position outside the EU Internal Energy Market, there are a variety of trading arrangements across existing interconnectors including different gate closure times for capacity trading, varying combinations of day-ahead and intraday trading, and a division between interconnectors that are explicitly traded and those that are implicitly traded.

There are clear opportunities to integrate at least some of today's interconnectors into CMMs, this is most obvious for explicitly traded interconnectors with both day-ahead and intraday explicit auctions. At present this would mean the interconnectors between GB and France, Belgium and the Netherlands might be relatively easily integrated into CMMs.

Understanding the market operation of interconnectors, and the options available for developing new trading arrangements including CMMs is a highly specialist area where even many of those with significant GB electricity market experience struggle.

There would be value in the UK Government convening an expert-led working group to carry out a deep dive into the impact of interconnectors on future constraints. This group could be tasked with developing the 'best possible' set of future interconnector arrangements from a GB market perspective, that would support constraint management alongside of a bilateral national wholesale market. This is an important part of an evolutionary approach to market reform and we need a much clearer picture of what is possible.

This group should also consider how such a set of arrangements would align with plans, laid out in the Trade and Cooperation agreement between the UK and the EU, to return to a form of implicit trading known as Multi-Region Loose Volume Coupling.

Recommendations

1 **Forecast and publish estimated annual constraint volumes and costs for future years across a range of scenarios:** estimates made by NGESO of constraint volumes and costs, such as those produced as part of the existing Network Options Assessment, should be published for a range of future scenarios and for timescales out to 2050. This would allow the sector as a whole to develop a clearer understanding and more informed debate about the challenge of managing constraints. Work should also be undertaken to quantify, where possible, the uncertainty in future constraints to ensure an understanding of risk to consumers.

2 **Carry out analysis to understand constraint forecasting on operational timescales:** detailed analysis should be conducted into the accuracy with which constraint volumes and costs can be forecast for specific days, and settlement periods over look-ahead times of weeks, days and hours ahead. This analysis should consider the degree to which different factors drive uncertainty including forecasts of weather, demand, interconnector operation, and the operation of dispatchable power stations.

3 **Develop a constraint management portfolio:** this would form a core component of a market reform option based around evolving the current national bilaterally traded wholesale market. A constraint management portfolio should have a clear objective and overall architecture agreed up front, but it should be capable of being developed in agile and flexible way, for example formalising and integrating existing trading strategies and pathfinder projects first, before adding more complex aspects later.

An example of a fully developed portfolio is shown in Figure ES3.

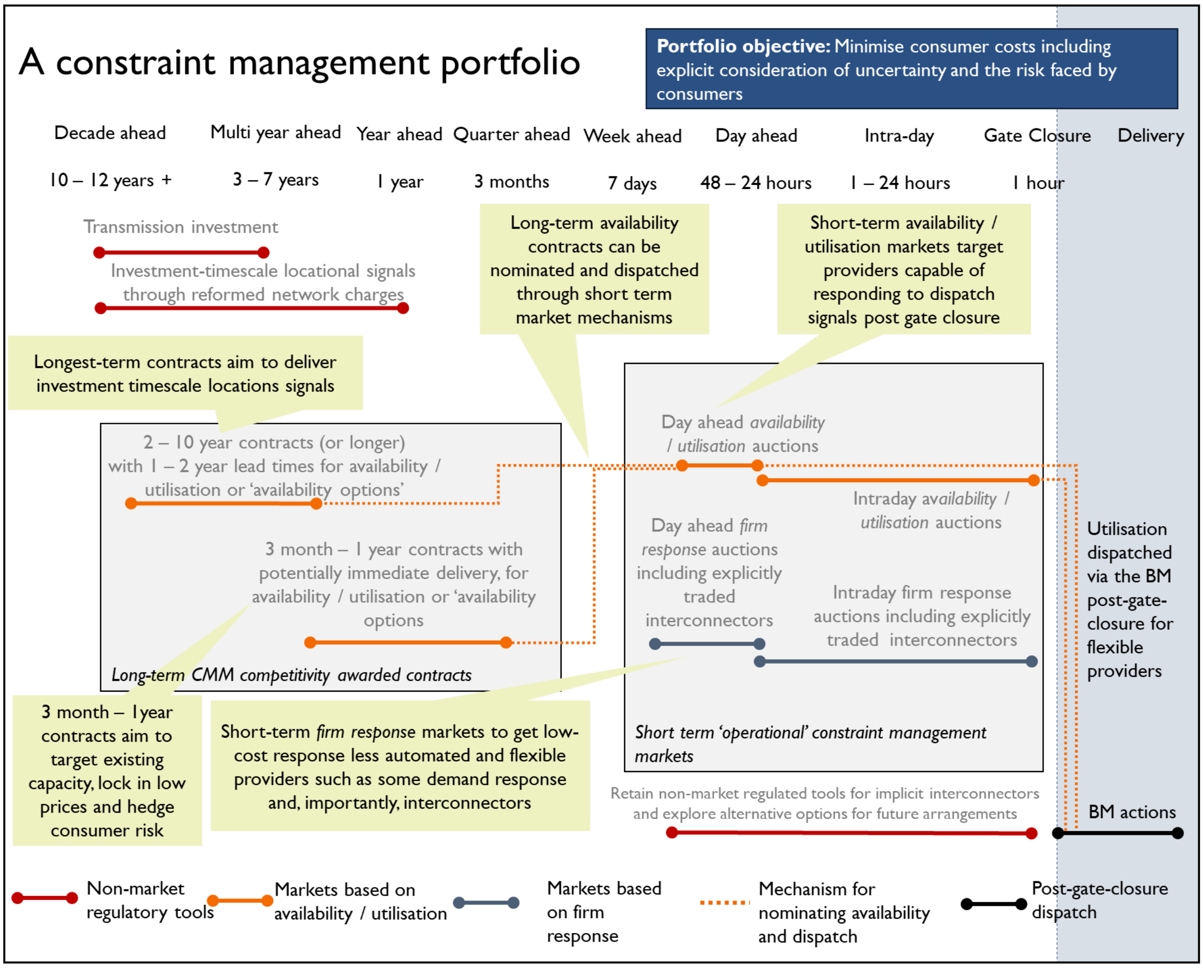


Figure ES3: An illustration of a constraint management portfolio which has the potential to effectively manage constraints, including uncertainty, through a mix of tools applied at different timescales from year ahead to post-gate closure.

4 **Define a clear constraint management objective:** A constraint management portfolio should have a clearly stated objective that is used across all timescales and tools. The objective should be based on maximising value and is likely to include a balance between minimising forecast consumer costs and managing consumer risk.

5 **Types of constraint management actions:** Within the portfolio, include options for both downward constraint management actions behind an export constraint, and upward constraint management actions in front of a constraint.

6 **Timeframes:** Structure the portfolio to include both long-term (e.g. 3 months to 12 years or longer) and short-term (day ahead and intraday) CMMs. There should be the potential for long term CMMs to contract flexibility for any timescale after the option for network investment has passed or for the lifetime of the asset.

7 **Long-term CMMs with competitively awarded contracts:** Explore the value of long-term contracts for availability to provide constraint management actions, awarded through competitive tender or auctions. These contracts could provide investment signals for new investment in flexibility or strategically situated demand. They can support delivery of good outcomes for consumers by acting as 'options' which lock in volumes and prices providing hedges against uncertainty.

8 **Short-term day-ahead and intraday CMMs:** Develop options for a set of short-term constraint management markets which, collectively, are accessible to the full range of potential providers: BM participants, individual domestic and business consumers potentially through aggregators, EV and heat pump fleet operators, non-BM embedded generation and flexibility, and interconnectors.

9 **Markets for availability / utilisation and markets for firm response:** Short-term constraint management market designs should prioritise, where possible, allowing the FSO to procure availability at day-ahead and intraday stage, with utilisation costs incurred closer to real time. However, complementary services, including purchase of 'firm response' ahead of gate closure may be required to allow some providers to offer constraint management who would not be in a position to participate in an availability / utilisation market design.

10 **Integrating explicitly traded interconnectors:** Develop specific options to integrate explicitly traded interconnectors into the short-term CMM designs developed in recommendations 7 and 8.

11 **Interconnector expert group:** Bring together a group of experts in interconnector trading with a mandate to develop the 'best possible' set of arrangements to allow interconnectors to participate in CMMs. This group should take account of the practicalities of trading arrangements including auction timing, explicit vs implicit trading, the impact on the connected markets, the direction of travel laid out in the EU target model and the EU-UK Trade and Cooperation Agreement, along with the pros and cons of diverging from that model in GB.

12 **Understanding market interactions:** Commission work to explore and understand the risks associated with interactions between constraint management markets, the wholesale energy market and the BM. This work should carry out analysis to compare any disbenefits that negative interactions such as gaming might create against the overall benefits that CMMs could deliver.

1 Introduction: constraints

Constraints arise because the wholesale electricity market does not take into account the location of generation and demand. When it comes to delivery, the market often creates a pattern of supply and demand which cannot be accommodated by the transmission network. The quantity of supply and demand that cannot be accommodated is called the constraint volume, and the cost of adjusting the market dispatch to fit within network limits is referred to as the constraint cost.

Following gate closure, an hour before delivery, National Grid ESO (NGESO) has the task of assessing the dispatch that the market has delivered and adjusting it to bring flows within network limits. Today, this is largely achieved through the Balancing Mechanism (BM).

The cost of constraints has been rising over recent years. In the financial year 2022-23 the total cost of constraints on the GB transmission network was £1.5bn[4]. Under current arrangements these costs are socialised across demand which equates to around £5.70 / MWh consumed. That adds about £15 / year to a typical domestic consumer's bill.

Constraint costs consist of two categories of cost: turn down costs behind an export constraint and turn up costs in front of an export constraint (Box 1 provides a description of terminology and conventions used in this report to describe constraints).

The majority of today's constraints are related to wind generation in the north of Britain, particularly Scotland, areas which have limited transmission capacity to carry power to demand centres further south. However, the pattern of congestion will change over time. By the 2030s, depending on our success in building new transmission capacity, we could expect to see constraints associated with offshore wind generation connected in East Anglia, solar generation in the south-west of England and interconnectors connected in several areas around the British coast.

1.1 Aim of this report

This report presents a discussion of constraint management options which could be implemented alongside a bilateral national wholesale market. As such it aims to add to the development of a credible evolutionary pathway for market reform to compare against the more revolutionary options based on locational marginal pricing also being considered. Recent work by Ofgem highlights the importance of developing a reformed national market model for the wholesale electricity market framework due to the "potential risk to investment and distributional impacts on consumers" [5] of locational wholesale pricing options.

The scope of work underpinning this report does not allow for detailed analysis of the options, challenges and solutions, rather it aims to identify some deficits of the current approach. This is largely limited to investment in new transmission capacity and operational redispatch in the BM. The report goes on to propose a number of potential improvements which could lead to a strategic constraint management portfolio consisting of a mix of market mechanisms and regulatory tools.

The report puts particular emphasis on the potential for Constraint Management Markets (CMMs) as part of such a portfolio. These appear to have the potential to improve current arrangements, allowing more options for intervention, over a greater variety of timescales, to minimise consumer costs and, importantly, manage consumer risk.

The ideas presented here should be treated as a starting point for further discussion rather than a firm proposal for exactly how a constraint management portfolio should be designed. The aim of the report is to support UK Government, Ofgem, NGESO and the wider electricity sector in developing the type of proposal that Ofgem suggests is important.

---

[4] https://www.nationalgrideso.com/data-portal/constraint-breakdown

[5] https://www.ofgem.gov.uk/publications/assessment-locational-wholesale-pricing-great-britain

Box 1: Constraint terminology and convention

Discussing constraint management needs careful use of terminology to ensure clear reference to actions on both sides of a constraint. This report uses a convention based on the perspective of an **exporting constraint** and **generation response.**

A constraint is, in effect, a line cutting the power system in two. An exporting convention (as opposed to an importing convention) means taking the perspective of the part of the power system which has an excess of generation over demand. A generation response convention (as opposed to a demand response convention) means taking actions to relieve that constraint – bring it back within limits – by changing the scheduled output of generators.

Using the export constraint and generation response conventions, resolving a constraint requires that the system operator takes the following two actions:

- **Turn down actions behind the constraint:** this could be to reduce generation, increase demand, reduce storage discharging, or increasing storage charging.
- **Turn up actions in front of the constraint:** this could be to increase generation, decrease demand, increase storage discharging, or decrease storage charging.

The terminology is summarised in Figure 1. The convention is somewhat arbitrary. Importantly, the choice to use a generation response convention does not imply that generation response is preferable to response from flexible demand or storage. In fact, many of the recommendations made in this report aim at facilitating greater contributions to constraint management from these sources of flexibility.

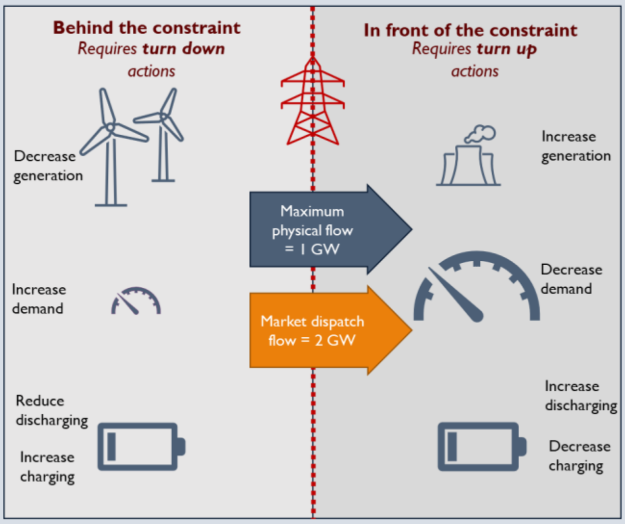


Figure 1: Export constraint, generation-orientated terminology used throughout this report

1.2 The cost of constraints today

The rise of constraint costs over the past decade has been driven by two factors: a growing mismatch between transmission capacity and generation capacity in the north of Britain, and more recently the significant increase in the price of gas-powered generation.

Growth in volume: mismatch between generation and transmission

The mismatch between generation and transmission comes largely from the significant growth of wind generation in Scotland paired with a sluggish expansion of the transmission network. In 2010, UK government changed the rules around connection to the transmission network from an 'invest and connect' approach – where generation had to wait until all network upgrades needed to make their connection economic had been completed – to a 'connect and manage' approach where connection was allowed earlier. The rationale was to remove a barrier to the development of renewables, one which was seen as a potential obstacle to delivery of the 2020 renewables targets. This view has been borne out in practice: the faster connections and greater confidence over connection timing that it provided renewable developers is seen as an important factor in the speed of renewable growth during the 2010s.

To manage the risk of unnecessary costly constraint that the move to connect and manage created, UK government identified the importance of delivering transmission upgrades quickly as possible. Their response to the original Connect and Manage consultation highlighted this, saying "[t]he ultimate solution to the problem of network constraints and connecting new generation is investment in the transmission network, and we are working closely with Ofgem to ensure that this is delivered in a timely and efficient manner."[6]

However, despite identifying this potential problem in 2010 the development of transmission has been slower than needed. This has been driven partly by lengthening timescales to gain planning consent. For example, a recent review by the National Infrastructure Commission suggested that typical consenting times for nationally significant infrastructure projects have grown from 2.6 years in 2012 to 4.2 years in 2021[7]. But a key bottleneck in transmission development has been the process to gain regulatory approval. Ofgem recently agreed with this, stating in late 2022 that the current regulatory framework "offer[s] heightened regulatory scrutiny and consequential consumer protection on a project-by-project basis" but "assessing large transmission projects under this framework is unlikely to afford the necessary pace to deliver the government's 2030 ambitions, and limits the scope for the required investment to be considered and delivered in a programmatic fashion."[8]

Despite these delays, some new transmission capacity has been commissioned over the past few years, including the HVDC western link between Scotland and North Wales and the Caithness Moray HVDC link in northern Scotland. This has meant that after rising through the first part of the 2010s, there has been relatively little growth in constraint volumes since 2018. For example, during the year 2017-18, volumes of thermal constraints were 3.5 TWh, whilst in 2022-23 it had only risen to 4.1TWh (see Figure 3 below)[9].

Growth in unit cost: rising prices

The majority of increase in the total cost of constraint since 2018 is related to the unit cost of constraints. The average cost of relieving constraints in the year 2018-19 was £109 / MWh, in 2022-23 it was £366 / MWh[10]. Of the two components of constraint costs – the cost of turn down behind a constraint, and the cost of turn up in front of a constraint – it is the latter which is responsible for the majority of the cost increase. Figure 2 illustrates the two elements that make up the constraint.

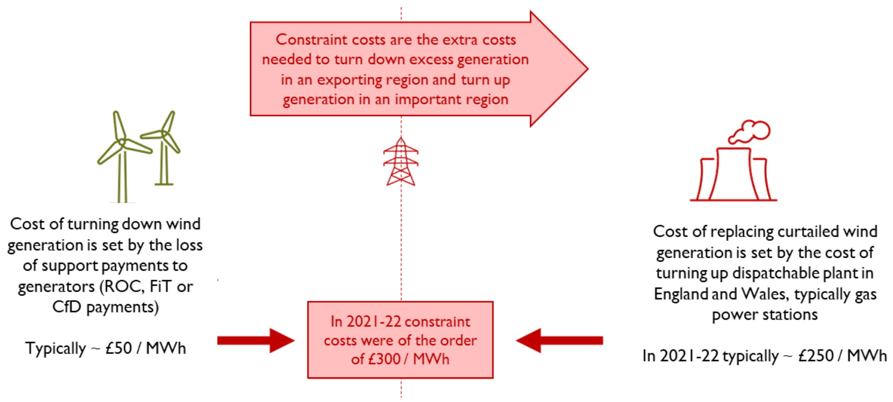


Figure 2: Illustration of the turn up and turn down replacement elements of constraint costs. Source: Regen[11] and NGESO[12]

[6] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/42979/251-govt-response-grid-access.pdf
[7] https://nic.org.uk/studies-reports/infrastructure-planning-system/delivering-net-zero-climate-resilience-growth/
[8] https://www.ofgem.gov.uk/sites/default/files/2022-12/ASTI%20decision%20doc%20-%20Final_Published.pdf
[9] https://www.nationalgrideso.com/data-portal/constraint-breakdown
[10] *ibid*
[11] Seven solutions to the rising cost of transmission network constraint management: https://www.regen.co.uk/wp-content/uploads/Regen-Insight-Managing-Constraint-Costs.pdf
[12] Constraint Breakdown – costs and volumes: https://www.nationalgrideso.com/data-portal/constraint-breakdown

Given that the majority of constraints are related to wind generation, the current situation for unit costs can be summarised as follows:

- **Turn down costs** are related to the loss of a generator's support payments due to the decrease in its generation and a small amount to reflect additional operation and maintenance costs. Typically, turn down costs are in the range £50 - £70 / MWh, reflecting the value of Renewable Obligation Certificates. As the prevalence of generators with Contracts for Difference (CfD) increase, this may change, but should not significantly exceed generators' strike prices[13].

- **Turn up costs** are related to the short run marginal costs of the replacement generation. This is usually a gas power station and therefore is directly related to the cost of the gas used to generate the replacement power. The increase in gas prices by a factor of eight during the recent price crisis therefore had a direct impact on the cost of constraint management[14].

In addition to the short run marginal impact on turn up costs there are a number of other factors that feed into the turn up side. In particular 'unit commitment' issues for schedulable power stations – which relate to how quickly a power station can turn on and off, the minimum amount of time they need to be on or off for, and the minimum stable output they can produce – means that NGESO can face additional costs beyond the pure short run marginal energy cost.

Figure 3 shows the components of thermal constraint costs over the past six years. It shows that volumes have only increased by a small amount, whilst the unit cost, the cost per MWh of constraint, was four times higher in 2022-23 than in 2017-18[15].

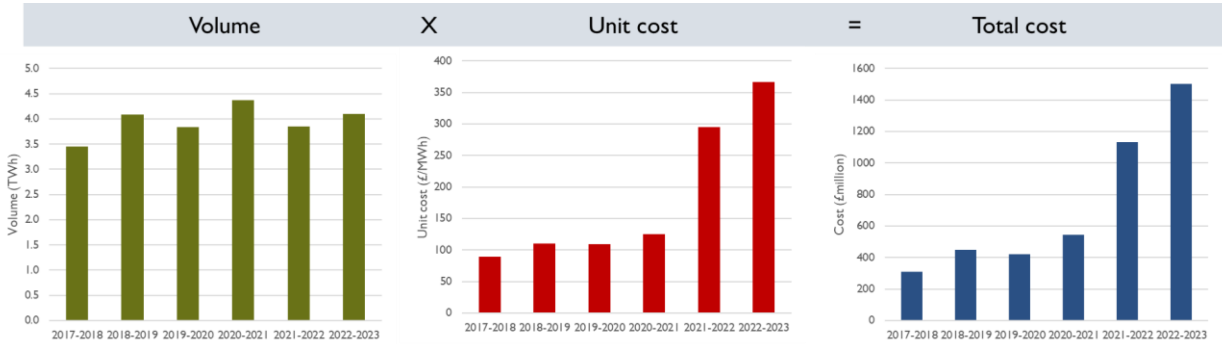


Figure 3: The components of constraint costs in recent years: volume on the left, average cost per MWh of congestion in the middle, and total constraint costs on the right.

1.3 The current approach to constraint management

Actions taken in today's market to manage constraints can be split into two groups: actions taken on planning and investment timescales and focused on investment in new transmission capacity; and actions taken on operational timescales and focused on re-dispatching the outcome of the wholesale market through the BM.

[13] A code modification, P462, may remove the cost of support payments from bids into the Balancing Mechanism, however they will remain a 'cost' of constraint because under the modification support payments would still be paid, but through an explicit mechanism, separate from the bid price. https://www.elexon.co.uk/documents/change/modifications/p451-p500/p462-initial-written-assessment/

[14] Ofgem Weekly average forward delivery contract data, averaged by quarter and comparison with Q2 2021 and Q3 2022. https://www.ofgem.gov.uk/energy-data-and-research/data-portal/wholesale-market-indicators

[15] Note that as of 31st October 2023, year-to-date volume for thermal constraints was 2.9 TWh compared with 2.1 TWh for the same point in the previous year, however with energy prices falling the total cost to date in 2023 was £567 million compared with £979 million for the same period in the previous year.

1.3.1 Investment timescales: planning the future transmission network

Effectively managing constraint costs, along with ensuring security of supply and minimising emissions, is one of the main reasons for investing in new transmission capacity. Constraints are fundamentally caused by limited network infrastructure and investing in new transmission capacity is an important way to keep constraints under control.

However, there are two important considerations. Firstly, the timescale for developing new transmission capacity is currently estimated to be around 12 years from the identification of need. Therefore decisions need to be made in the face of significant uncertainty around the future operating context. Secondly, investment in transmission is about balancing the reduction in constraint costs (and other benefits that transmission capacity brings) against the cost of that transmission infrastructure. Therefore optimal transmission investment will not eradicate constraints, but lead to the situation where investment in the next unit of transmission capacity is balanced against the benefits it will bring.

The reduction in constraints is one of the key benefits assessed when making decisions about transmission investment. For the past decade, NGESO has used a three-part cycle to estimate the need for new transmission capacity. Initially four detailed scenarios are developed through the Future Energy Scenarios (FES), scenarios include the location of generation, demand and flexibility. Next, for each scenario the transmission capacity required across major system boundaries is calculated and published in the Electricity Ten Year Statement (ETYS). Finally, the Network Options Assessment (NOA) uses a cost-benefit analysis (CBA) in which the costs of specific transmission investments are compared against the reduction in constraint costs they would facilitate for each scenario. Where the CBA shows that a particular transmission reinforcement is important to developing an efficient system, that investment is given a 'proceed' recommendation and is then passed to the Transmission Owners (TOs) to develop, gain regulatory approval, and deliver.

Uncertainty is incorporated in the NOA process through its use of the four FES scenarios. The impact of each transmission investment is assessed under each scenario, and the final decision on whether to proceed or hold is based on a least worst regret process[16]. The process is summarised in Figure 4[17].

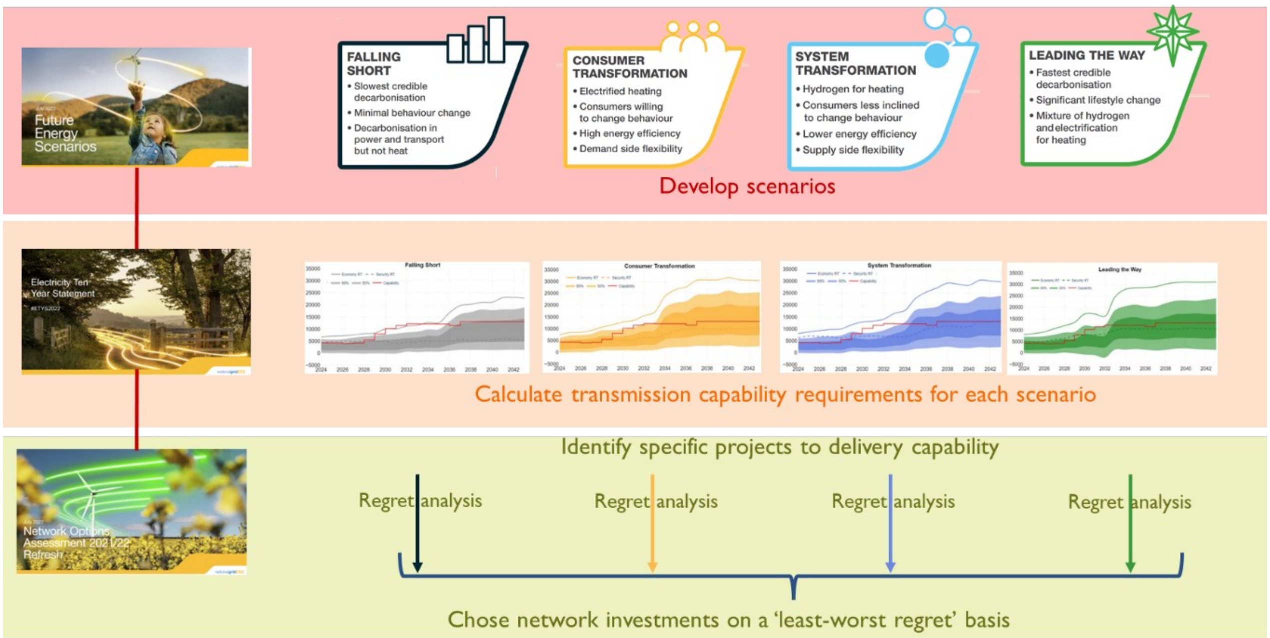


Figure 4: Summary of the transmission planning process used between 2015 and 2022. The process uses scenarios and 'regret analysis' to consider uncertainty in future constraint costs. The process could form a model for long term CMM contracting.

---

[16] A least worst regrets approach involves calculating the result of a CBA for an investment in each of the four scenarios then highlighting the 'worst regrets' scenario – i.e. the scenario where the decision to invest led to the worst outcome. Investment decision are made where the 'worst regret' is minimized.

[17] The details of the NOA methodology for 2023 are available here: https://www.nationalgrideso.com/document/285321/download

The annual cycle described above has operated each year from 2015 to 2022. In 2023 the focus of transmission planning moved towards a strategic approach and a new framework, the Centralised Strategic Network Plan (CSNP), is expected to operate from 2024 onwards. The CSNP will include many of the features of the previous approach including pathway and scenario development, and a CBA for network investment, but will do so against pathways more focused on delivering strategic outcomes[18].

Constraint cost forecasts

Figure 5 shows estimated annual constraint costs for the period 2017 to 2035 combining historic outturn and future forecasts. It shows that constraint costs in future are expected to be higher than those experienced before 2020. For the period 2025 – 2029 two estimates are shown based on the FES *2022 Leading the Way* and *System Transformation* scenarios.

Under *Leading the Way* constraint costs during the second half of the 2020s are expected to be significantly higher than recent years, whilst under *System Transformation* they will remain at similar levels to 2022.

The time series shows a significant fall in costs in 2030, reflecting the assumed delivery of new transmission capacity which has been planned through the Holistic Network Design process and given accelerated regulatory approval pathway through Ofgem's ASTI framework. Constraint costs rise again beyond 2030, but this rise reflects the fact that further new transmission capacity has yet to be planned and approved. If generation were to develop inline with the *Leading the Way* scenario it would be expected that further transmission upgrades would be commissioned, hence those constraint costs below those shown. In fact since the forecasts used in Figure 5 were created in 2022 significantly more capacity has now been identified in the early 2030s[19].

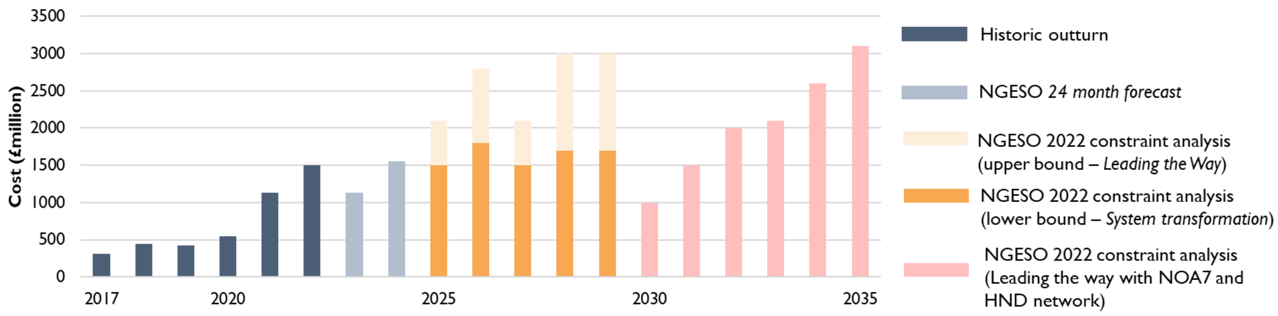


Figure 5: Constraint costs: historic outturn and future forecasts. Historic costs are from NGESO constraints breakdown[20]. Forecasts for 2023 – 2025 are based on the 24-month forecasts made by NGESO, forecasts from 2025 onwards are based on 2022 analysis carried out to support market reform debate. For the period 2025 – 2029 upper and lower bounds are shown based on the net-zero compliance scenarios from FES2022 and a network developed according to NOA7. From 2030 onwards only an upper bound (Leading the Way) is shown and the network is based on the NOA7 update which included the HND network. Source: NGESO[21]

There are, however, additional uncertainties and risks not factored into the analysis presented in Figure 5. These include:

- **Transmission delay:** the risk that transmission network capacity planned for delivery during the 2020s is delayed meaning the HND is not delivered on time. This would lead to higher constraint costs.

[18] For more details on the CSNP framework, see Ofgem's recent decision https://www.ofgem.gov.uk/publications/decision-framework-future-system-operators-centralised-strategic-network-plan
[19] See ETYS 2023, https://www.nationalgrideso.com/document/286591/download
[20] https://www.nationalgrideso.com/data-portal/constraint-breakdown.
[21] Historic data: https://www.nationalgrideso.com/data-portal/constraint-breakdown
Forecasts for 2023 and 2024: https://www.nationalgrideso.com/data-portal/24-months-ahead-constraint-cost-forecast
Forecasts from 2025 onwards: https://www.nationalgrideso.com/document/266576/download

- **Generation development:** the risk that low carbon generation isn't delivered inline with what is needed for a net zero electricity system. This would likely lead to lower constraint costs but would also represent a failure to deliver on decarbonisation.
- **Flexibility development:** a lack of flexibility could lead to an increase in constraint costs, whereas the faster development of (the right types) of flexibility could decrease those costs.
- **Electrification of heat and transport:** uncertainty around the rate at which heat and transport demand are electrified – slower electrification of demand would likely lead to higher constraint costs.
- **Locational uncertainty:** uncertainty in the location of generation, flexibility and demand.

The conclusion of existing analysis is that constraint costs in the late 2020s could be between £1.5bn and £3.0bn per year but that there is actually even greater uncertainty over the actual volume and cost of constraints due to uncertainty over the speed of the net zero transition, the timeliness in which new network capacity is built and the success in ensuring flexibility to minimise constraints.

The uncertainty over constraint can be highlighted by comparing recent forecasts and outturn. For example, on planning timescales, in March 2019 NGESO forecast the constraint element of Balancing System Use of System Charges for the year 2021/22 at £542 million[22] whilst the outturn cost was £1071million[23].

1.3.2 Operational timescales: managing constraints through the balancing mechanism

Once the timescale for network investment has passed, there are no formalised frameworks for taking constraint management actions until gate closure and the opening of the balancing mechanism (NGESO does undertake some energy trading and interconnector actions on an ad-hoc basis. These are described below).

The BM is the route for the majority of operational timescale-constraint management today. All large-scale generators participate in the BM as do a number of other assets including batteries, pumped storage and some large industrial consumers. Recent improvements in the way the BM operates, such as the introduction of 'virtual lead parties' have increased the opportunity for aggregators to facilitate participation for smaller flexibility providers, pooling response from distribution connected assets[24].

Figure 6 summarises the BM process. Initial Physical Notification (IPN) must be submitted on an asset-by-asset basis by all BM units by 11 AM the day before delivery. Whilst these give an indication of what the unit is expecting to do, they are not binding, and can be adjusted to reflect either intra-day wholesale trading, or a rearrangement of how a company's portfolio will meet its wholesale market commitment across different assets. All BM participants must submit a Final Physical Notification (FPN) ahead of gate closure, indicating what they expect to generate or consume on an asset-by-asset basis[25]. They can also choose to provide offers and bids to decrease or increase their generation / demand. These are prices that NGESO can pay or be paid to ask those assets to change their generation away from the submitted FPN to a new level following gate closure.

[22] https://www.nationalgrideso.com/document/141946/download
[23] https://www.nationalgrideso.com/document/284216/download
[24] https://www.elexon.co.uk/about/roles/virtual-lead-party/
[25] The process is slightly different for suppliers who will hold BM Units (BMUs) for each grid supply point at which they operate. They therefore submit Final Physical Notifications, bids and offers based on the aggregation of demand and distributed generation / flexibility for which they are responsible.

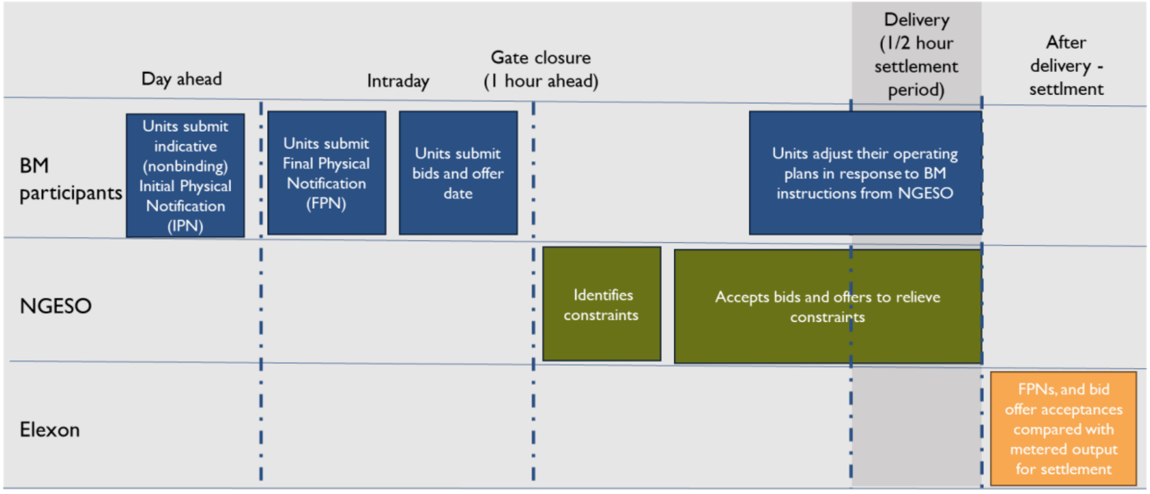


Figure 6: Summary of the BM process involving individual BM participants, NGESO and Elexon who settle the market following delivery.

The process of constraint management effectively involves accepting bids behind a constraint to reduce generation (or increase demand) and at the same time accepting an equal quantity of offers in-front of a constraint to replace that generation (or cover the additional demand)[26].

This description of the BM helps clarify why there is a cost to constraints. If the non-locational wholesale market is competitive and efficient, it will dispatch the cheapest set of generation available in the country to meet the national demand. However, constraints mean that output from some of the cheapest generators cannot be transported to consumers. The process described above is a process of replacing some cheap generation with more expensive generation elsewhere in the country. In addition to this 'theoretical' cost increase, the compression of this process into the last hour before delivery and the practical considerations which are needed in order for the control room to confidently maintain the safety and security of commitment can add additional expense.

1.3.3 Trading ahead of gate closure and interconnector management

In addition to waiting for the BM, NGESO can also carry out forward trades in the wholesale energy market to proactively reduce constraints at the point of gate closure. In the financial year 2022-23, around 14% of constraint costs, or 19% of volume, were due to forward trading[27].

Trading options available to NGESO include contracting with specific GB generators in particular locations relative to forecast constraints, and the ability to 'countertrade' with owners of capacity on some interconnectors. However, the ability to do this is limited to interconnectors whose capacity is 'explicitly traded' where market participants buy the interconnector capacity for given settlement periods separately from their trading of energy in each of the connected markets. By contrast the North Sea Link to Norway and the interconnectors between GB and the Island of Ireland are implicitly traded and counter-trading is not possible. On these interconnectors, non-market regulatory actions can be used to limit flows.

In summary, NGESO is able to use the following actions ahead of gate closure to manage constraints:

- **Energy trading**: means the process of NGESO themselves buying or selling energy in the wholesale market. They will trade with specific assets in strategic locations relative to constraints to reduce or relieve the forecast constraints before gate closure. Once a trade has been agreed it will be reflected in the asset's FPN at gate closure leaving less constraint volume to manage within the BM.

[26] Note that, at the same time the BM is carrying out a separate 'energy balancing' exercise as well as dispatching some ancillary services. The result is that, in practice, it may not be possible to identify precisely which actions correspond to constraint management and which to other activities, particularly in relation to the turn up actions in front of a constraint. In fact one action could be used to meet multiple system needs.
[27] Derived from Monthly Balancing Services Summary data for March 2023: https://www.nationalgrideso.com/data-portal/mbss

- **Interconnector management**: can be broken down into several subcategories. The options available vary by interconnector and depend on the timing and type of auction held (a more detailed summary of interconnector arrangements is given in Section 4.2):

  - **Counter trading:** National Grid contracts with parties who hold capacity on an explicitly traded interconnector, have nominated that capacity, and hold a position in the GB wholesale market. Countertrading involves paying those parties to reverse their position. For example, where an interconnector trader plans to import into GB and has sold that energy in the GB wholesale market, a countertrade allows them to agree a price with NGESO which pays them to buy back the energy in the GB market and set their import on the interconnector to zero. (A more comprehensive example is given later in Box 6.)

  - **Net transfer capacity (NTC) restrictions:** this is a non-market tool that allows NGESO to limit the usual interconnector capacity in return for a payment to the interconnector owner. NGESO aims to avoid using NTC restrictions except as a last resort[28], preferring to use market solutions. However NTCs can be the only available option in the case of implicitly traded interconnectors.

  - **System Operator to System Operator (SOSO) trade:** this is where NGESO adjusts flows by working directly with the System Operator on the far end of an interconnector rather than with either market participants or the interconnector owner. SOSO trades are only a small part of interconnector management for constraints today and, as with NTCs, NGESO prefers to avoid this as a non-market approach.

Whilst these options are available, and are used increasingly to help manage constraints, NGESO has not set out a clear approach describing where and when energy trading or other ahead of gate closure actions will be used in preference to the BM.

Energy trading and interconnector management does not operate within a formal, transparent, well-structured framework. As such it is difficult to understand how NGESO make decisions about when to use these tools and when to avoid using them.

One of the recommendations of this report is that the expansion and formalisation of these activities, in conjunction with a clear articulation of the objective which they are applied, would be likely to improve outcomes for consumers.

1.4 New approaches to constraint management

The 'business as usual' approaches, described above, to operational constraint management represent the primary ways in which constraints are managed. There have been a number of pathfinders and demonstration projects which show the potential for new constraint management tools to deliver benefit for consumers.

1.4.1 The Local Constraint Market: Scotland

In 2023 NGESO introduced a constraint management market for Scotland[29]. This is a day-ahead and intra-day market to purchase firm response for turn down actions behind the constraint. It procures from non-BM providers. As such it aims to increase the pool of providers able to provide turn down services in Scotland and reduces the reliance on BM wind curtailment.

---

28 https://www.nationalgrideso.com/sites/default/files/documents/Amended%20Methodology%20for%20GB%20Commercial%20Arrangements%20relating%20to%20Interconnector%20Capacity%20Calculation.pdf

29 https://www.nationalgrideso.com/industry-information/balancing-services/local-constraint-market

The market operates to manage constraints on both the B6 boundary (Scotland to England) and B4 boundary (between central and northern Scotland) and aims to accept actions when there is an expected 2% saving in comparison with waiting for the BM.

Participants in the LCM include portfolios of EV charging points; batteries; industrial, commercial and domestic demand response and non-BM onshore wind. Volumes are relatively small, reflecting the early stage of the market's development. The majority of volume has been delivered through onshore wind turn down. After allowing more expensive bids to be accepted in the proof-of-concept phase, since October 2023 NGESO have implemented the rule that only bids more cost-effective than the BM will be contracted through the LCM. Since then a number of low volume bids at £60 / MWh have been contracted[30].

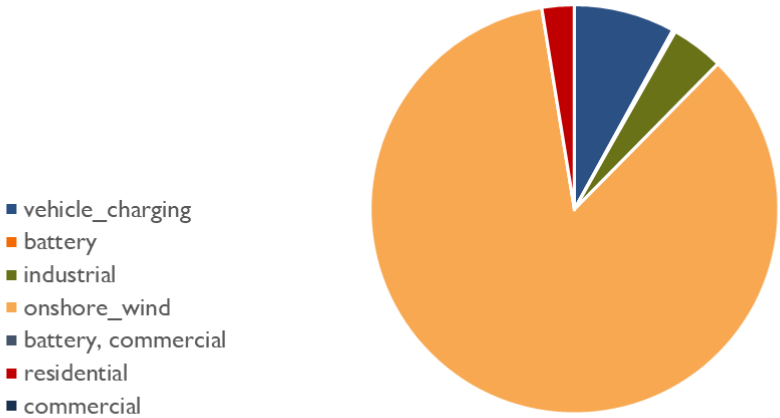


Figure 7: Breakdown of contracted volume through the LCM in Scotland between April and November 2023. Total volume contracted was 104 MWh, the majority was from non-BM onshore wind but there was also contracted volume from vehicle charging, residential and industrial demand response.

1.4.2 Constraint management pathfinder: Intertripping

Intertripping is the process of allowing power flows across the network to exceed their usual limits when there is generation capacity which will trip off within a very short time following a network fault. It has been used with some generators for many years, but there has been a recent push to increase its use.

Intertripping effectively increases the capacity of a boundary under normal operating conditions. Usually the power flows must respect redundancy: no limits must be breached even after a fault on one or more transmission lines. For example, although the B6 boundary between Scotland and England has around 12.2 GW of thermal transmission capacity crossing it, the secure boundary capability is 6.6 GW, reflecting the fact that the loss of any two circuits would mean that power flows were not exceeded.

Intertripping allows boundary flows to exceed 6.6 GW but in return some generators behind the constraint will be tripped in the event of a network fault and will need to stop generating. The Constraint Management pathfinder has been developing arrangements to reward providers willing to accept an intertrip. Terms involve payment for being 'available' to intertrip, known as arming, and payment for periods when tripped off. The potential to increase the use of intertrips to manage constraints across B6 is being explored through the Constraint Management pathfinder[31]. NGESO has also suggested using the approach for other boundaries, including the EC5 boundary exporting from East Anglia towards the Midlands and London[32].

Intertripping can effectively increase the boundary capability under normal operating conditions and could be a useful tool as part of a constraint management portfolio.

1.4.3 International experience: developing approaches for energy storage

One of the issues with the current GB approach is that there is a lack of opportunity to use batteries effectively in constraint management. An approach currently being explored on the island of Ireland is to develop options for the system operator to procure system services, including constraint management,

[30] Operational data available from: https://data.piclo.energy/
[31] https://www.nationalgrideso.com/document/185696/download
[32] https://www2.nationalgrideso.com/document/277921/download

from long-duration energy storage. A recent call for evidence notes that whilst short-duration batteries appear to have a strong route to market there is a missing money problem for long-duration storage created by a gap between the market revenues available and the value provided to the system.

One option that scores particularly highly in the consultation's multi-criteria assessment is a long-term system service contract with centralised control. The approach laid out involves offering contracts for either 15 years or for the lifetime of the asset, with the system operator sending dispatch instructions based on optimal operation from the system-wide perspective[33].

1.5 Managing constraints: the need for change

Constraints will be an important component of system cost for the next decade as we attempt to catch up on transmission build and deliver a well-planned network suitable for a net zero power system. However, even if we develop an optimal network in the 2030s, we should expect some level of constraints to remain. The balance point is found where the cost of investing in new network capacity is equal to the savings created from reduced curtailment.

It is important that we have access to a well-designed constraint management framework capable of dealing with the high level of constraints we are likely to experience over the next decade, and the enduring level of constraints beyond.

Today, our strategy of managing constraints through only (a) decade-ahead transmission investment decisions and (b) post-gate closure BM actions means interventions are forced to the two extremes of the timeline. There are also some important gaps in the articulation of the holistic approach to constraint management:

- **A lack of a clearly defined objective for constraint management that specifically deals with risks to consumers as well as expected costs:** NGESO's obligations and incentives are set through a combination of license conditions and price control outcomes. Overall NGESO is expected to facilitate an efficient, co-ordinated and economic electricity system[34]. Specifically relating to constraints: there is an incentive to reduce balancing costs (including constraints) against a specified metric[35], and guidance which highlights that NGESO should have an ambition to 'implement a comprehensive plan to proactively mitigate material increases to balancing costs'[36].

- **No actions available to NGESO between the lead time for transmission and operational timescales:** Once the time has passed for transmission reinforcement, there are currently no further tools available to respond to changing constraint forecasts until the day-ahead stage. Recent work by Ofgem puts timescales for transmission investment decisions at around 12 years[37], although the objective of recommendations made by the Electricity Networks Commissioner in 2023 is to halve the development time[38]. This is despite the fact that our knowledge of constraints will evolve significantly in the period between 12 years ahead and operational timescales. During this gap we will have increasing certainty over many of the factors which affect constraints: the capacity of transmission likely to be available in a given year, the generation and flexibility fleets including geographic distribution, and the size of demand and its degree of flexibility.

[33] See pg 36 - 40: https://www.eirgridgroup.com/site-files/library/EirGrid/LDES-Call-for-Evidence-EirGrid.pdf
[34] License condition C28: https://www.ofgem.gov.uk/sites/default/files/2023-03/Electricity%20Transmission%20Consolidated%20Standard%20Licence%20Conditions%20-%20Current.pdf
[35] See metric 1A: https://www.ofgem.gov.uk/sites/default/files/2023-03/ESORI%20Guidance%20Document%202023-2025.pdf
[36] https://www.ofgem.gov.uk/sites/default/files/2023-03/ESO%20Roles%20Guidance%202023-2025.pdf
[37] For example see consultations on the development of Centralised Strategic Network Plan such as https://www.ofgem.gov.uk/publications/centralised-strategic-network-plan-consultation-framework-identifying-and-assessing-transmission-investment-options
[38] https://www.gov.uk/government/publications/accelerating-electricity-transmission-network-deployment-electricity-network-commissioners-recommendations

During this gap there are a wide range of actions that could benefit consumers, including contracting with generation, flexibility and demand for availability to provide constraint management.

- **No formal frameworks or clear articulation of objective and strategy for pre gate closure actions:** whilst the level of energy trading and interconnector management has grown in recent years, the approach remains ad hoc. Several tools are used which can deliver value to consumers including energy trading with specific assets and interconnector counter-trading. However there is a lack of strategy and a coordinating framework. This is an important gap and one that can be filled by the development of day-ahead and intraday constraint management markets.
- **Overreliance on the final hour before delivery:** with at least 80% of balancing actions being dispatched through the BM, most constraint actions need to be taken with less than 90 minutes warning. Whilst many providers are capable of responding on these timescales some are not. In particular, less flexible generation, some forms of demand response and interconnectors. Although this can be partially mitigated through existing ahead of gate closure actions, the lack of a clear and well-socialised framework for doing so limits their availability.

- **Limited options for battery storage and other forms of flexibility to offer constraint management scenarios under current processes:** the way the BM is operated today means that battery storage is largely excluded from constraint management. This is partly a practical issue related to current control room processes. For example, for purposes of dispatching flexibility, the control room designates a maximum 15-minute firm response time from batteries, reflecting the fact that (a) many batteries in the market today are short duration e.g. 1 hour; and (b) the control room doesn't currently have access to State of Charge information.

  Many forms of flexibility, including energy storge, have operating constraints which need to be managed across multiple settlement periods, but current processes do not allow that. Intertemporal considerations for market participants are those where the ability to respond at one time depends on actions at another time. One example is energy storage charge management. The ability to discharge energy storage at one time relies on it having charged at a previous time. Therefore, if there is likely to be a period of constraints lasting several hours, the ability to relieve that using energy storage relies (a) on storage behind the constraint being discharged prior to the constraint becoming binding; (b) on storage in-front of the constraint being charged prior to the constraint becoming binding.

  Managing intertemporal constraints effectively means considering operation across many settlement periods together. For example, dispatching energy storage to consumers during a potential constraint several hours in the future requires that the storage is able to discharge ahead of the constraint becoming active.

2 A portfolio approach to constraint management

The need for change, articulated in section1.5, suggests that there is likely to be value in exploring a more strategic approach to constraint management. This will involve the clear articulation of the objective, explicit consideration of consumer risk, and a joined-up approach that runs through from transmission planning to real time and provides appropriate tools to a Future System Operator (FSO).

This report proposes that a 'portfolio approach' to constraint management should be explored, which uses both market and regulatory tools applied across a wide range of timescales. The bookends of the portfolio are already in place: over the longest timescale of around a decade or more this portfolio would include an evolution of existing processes for transmission planning and investment. At the shortest timescales it would continue to use an evolution of the BM but with significant improvements to processes.

Within these two bookends, there is likely to be value in developing both market frameworks and regulatory approaches to provide tools that can respond to our changing forecast of constraints in the period from decade-ahead to gate closure; in effect providing the 'books' between the bookends.

These market and regulatory frameworks would be designed to maximise and protect consumer value, considering both costs and risks faced by the consumer.

The role of markets in constraint management

Market-mechanisms are likely to deliver value where it is feasible to use them. They can leverage competitive pressure and, if well designed, encourage commercial providers to deliver the services that the system needs in a flexible way. The next chapter discusses the use of market-mechanisms, defined and described as constraint management markets (CMMs), and their potential role in overall constraint management. From a portfolio perspective there are two groups of CMM:

- **Long-term CMMs**: the use of competitive approaches to procuring long-term constraint management contracts over timescales of three months or more, including over timescales of a decade or more.
- **Short-term CMMs:** market frameworks which support scheduling of constraint management actions ahead of gate closure, formalising and expanding existing trading activity and focused on day-ahead and intra-day timescales.

In order to develop CMMs in the most appropriate way it will be important to properly articulate the objective. As noted above, the objective should be based around maximising consumer value and include appropriately weighted considerations of both overall consumer cost, and management of consumer risk. It is likely that a portfolio will involve different contract types, not just contracts for 'firm response' but also a variety of availability / utilisation contracts, and potentially other 'option' based contracts (these are described in more detail in Section 3.2.1).

Consumers can also benefit from a portfolio that manages the costs of both downward and upward actions. As discussed in Section 1.2 a large part of consumer costs and risk exposure is related to the turn up element of constraint management.

Box 2 and Figure 8 provide an illustration of the design principles and potential components that could make up a constraint management portfolio.

Box 2: illustrative objective and design principles for a constraint management portfolio

Objective: To maximise and protect consumer value, including minimising consumer costs including a specific requirement to manage consumer risk appropriately.

Design principles:

- Provide appropriate tools for managing constraints which can be applied on timescales ranging from decade-ahead to post gate closure BM actions.
- Provide the FSO with the capacity to respond to evolving constraint forecasts at each timescale.
- Develop options that support investment in low-carbon flexibility in appropriate locations where the probability of future constraints, and the likely cost of resolving them, are high.
- Develop constraint management options that cover both downward constraint management actions behind a constraint and upward constraint management actions in front of a constraint.
- Consider tools that allow the FSO to lock in availability and utilisation prices where doing so is in the consumer interest.
- Develop operational timescale tools which maximise consumer value.
- Develop tools that balance the potential downsides of interactions between wholesale and constraint management markets and the BM.
- Consider any new *pre gate closure* interventions against an appropriate counterfactual such as the cost and challenge of continuing to wait for *post gate closure* actions in an improved BM.
- Develop options that support response from all forms of flexibility including demand-side response, energy storage, flexible generation and interconnectors. This could be through sufficiently broad, technology-neutral contracting or through the development of different tools to suit different types of provider.
- Enable interconnector participation in the constraint management portfolio where feasible within broader GB and international electricity market trading arrangements.
- Focus on the use of contracts for ‘availability’ or option-based contracts where possible in order to minimise consumer risk.

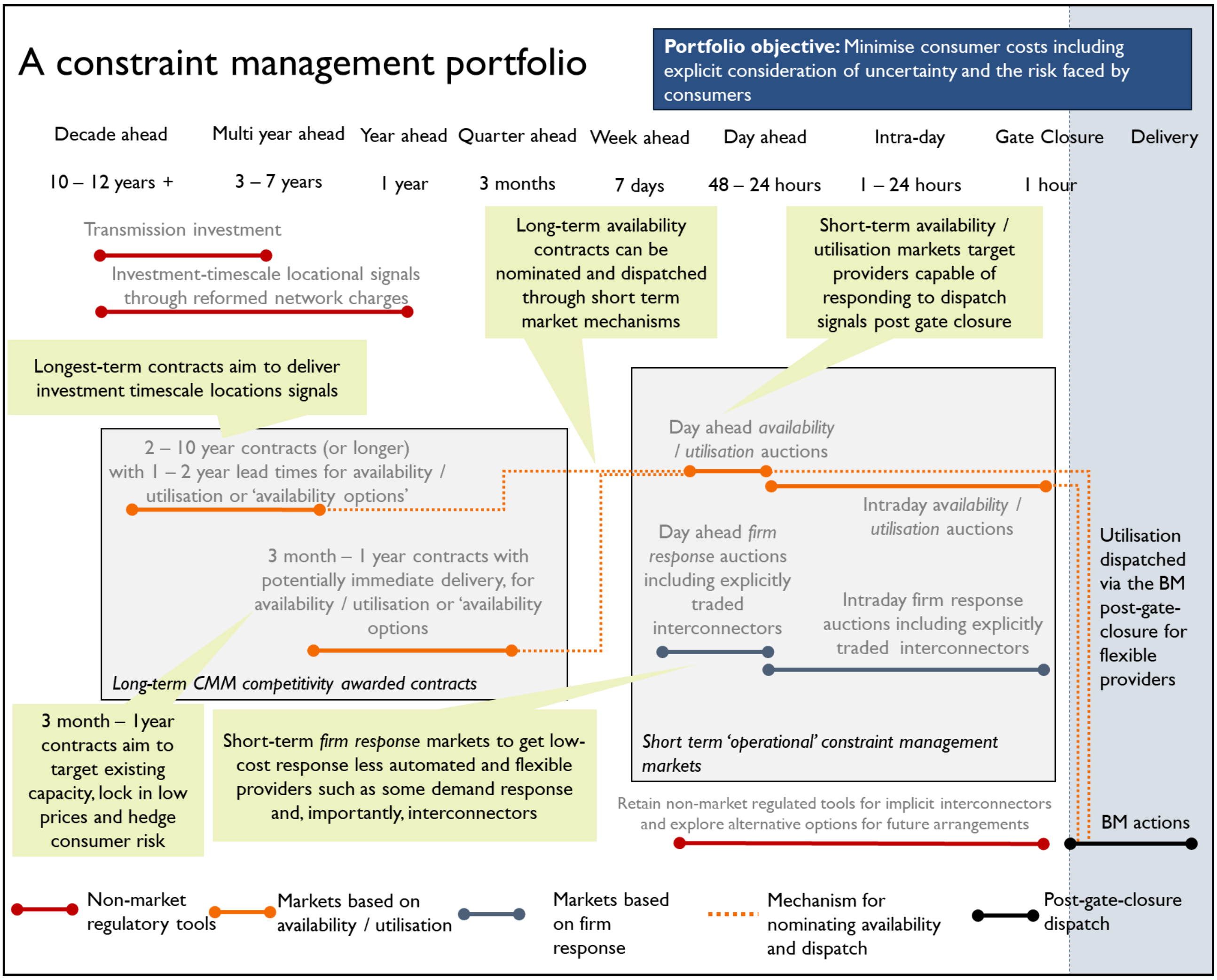


Figure 8: An illustration of a constraint management portfolio which can effectively manage constraints.

3 Constraint management markets

The previous section has set out a rationale for a broad portfolio-based approach to constraint management which includes regulatory approaches to transmission investment, the use of market mechanisms over different timescales, and the retention of a BM over the shortest timescale.

This section expands on the use of CMMs as part of such a constraint management portfolio. It specifically considers both long-term and short-term CMMs.

3.1 What are constraint management markets?

CMMs can take a range of forms. In this report they are defined as:

> Any market-based approach operating ahead of gate closure through which the FSO can buy or sell flexibility, or related products such as 'availability' or 'options' in order to relieve constraints on the transmission network. This includes both downward flexibility behind a constraint and upward flexibility in front of a constraint. They can include options for contracting for constraint management over several months or years (long-term CMMs) or running auctions for constraint management days, hours or minutes ahead of delivery (short-term CMMs).

In addition to the definition there are several characteristics that are likely to be present in any constraint management market design:

- A central buyer market with the FSO defining the quantity and characteristics required.
- Non-mandatory participation by market participants, following the approach used with ancillary service products.
- Open to a wide range of potential providers and not unduly restricted by size and type of provider (e.g. it should not require being a BM participant).

3.2 Long-term constraint management markets

For the purposes of this work, long-term CMMs are any mechanism with a lead-time of three months or more, that use competitive approaches to procure constraint management services. Here, the type of long-term CMMs discussed are based on long-term contracts, for example a contract to provide constraint management services over several months or years.

One advantage of using long-term contracting is that it allows the price of constraint management actions to be fixed in advance, hedging the risk of prices increasing in the future. There may be some concern over committing consumers' money to constraint costs far in advance, particularly if there is uncertainty within forecasts of future constraint volumes. However, this concern should be balanced against the fact that failing to take hedging actions can also place consumers at considerable risk and erode the ability to deliver consumer value.

This calculus of risk needs to be a central factor in designing long-term CMM markets. For example the structure of contracts can be designed in different ways to balance the risk held by the FSO (on behalf of the consumer) and constraint management providers.

The NOA process used for making transmission investment decisions based on constraint forecasts could provide a template for procurement decisions in long-term CMMs. Such an approach would use scenario forecasts along with regret analysis (or other approaches that incorporate uncertainty), to decide where and when there is consumer value in long-term contracting.

Long-term CMMs are unlikely to allow purchasing of response for specified settlement periods. Rather, they would most likely buy response to cover an extended period of a year, season or month, with nominations of the specific hours of delivery left until operational timescales. For example, contracts might buy a specific depth of response for a specified number of hours during the year. This would align with the nature of constraint forecasts a year or more in advance: it is possible to forecast expected value and probability distributions related to the total volume or the number of hours per year that constraints across a boundary will occur.

Figure 9 illustrates an example of a 'constraint-duration graph' for the B4 boundary between central belt and northern Scotland in 2030. The graph effectively takes a time-series of constraints across a year and reorders them so that the deepest constraint becomes hour 1, the second deepest constraint becomes hour 2 etc. It therefore shows the duration or number of hours in a year at which the constraint is at least that deep. The figure shows four scenarios, which might be considered as 'forecasts' of constraints, created from modelling different mixes of generation, demand, flexibility and transmission capacity[39]. The point at which the plot hits the vertical axis shows the maximum depth of constraints in the year and the point at which it hits the horizontal axis gives the number of constraint hours forecast.

For the least constrained scenario the graph shows that the 500 most constrained hours all have constraint depths of 2 GW of more, whilst for the most constrained scenario, the equivalent figure is nearly 8 GW. At 1000 hours duration, constraint depths are at least 1 GW and 6.5 GW respectively for the extreme scenarios.

Under a low-regrets approach, a long-term CMM might be used to procure contracts for availability of 2 GW behind the constraint for 500 hours during the year, and 1 GW of response for 1000 hours during the year. If the FSO felt able to put meaningful estimates on the probability of different scenarios it may be suitable to procure more.

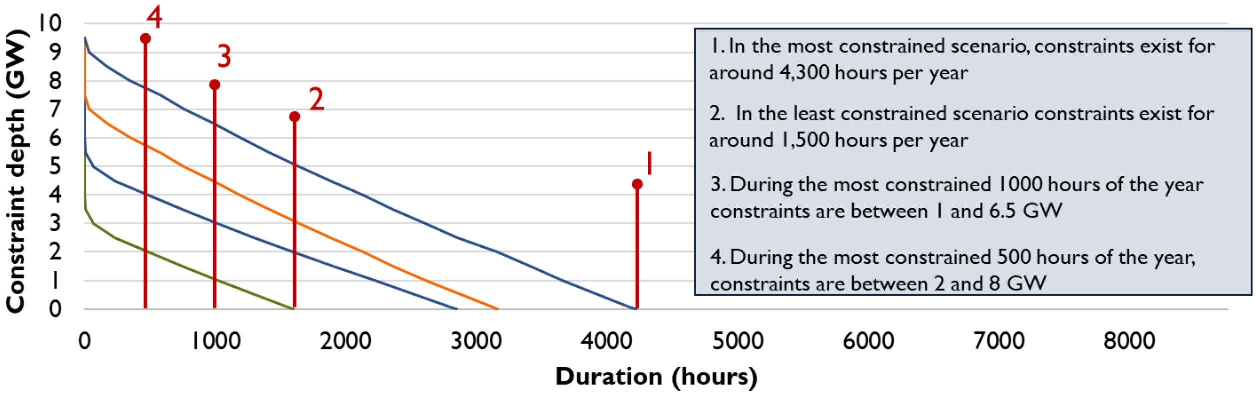


Figure 9: A duration curve for constraints showing the number of hours per year that constraints are equal or greater than a specific depth in four scenarios based on the B4 boundary in 2030. The point where the plots meet the vertical axis is the maximum constraint depth, the point where it meets the horizontal axis is the number of hours with constraints.[40]

In the examples discussed below, contracting for 1,000 hours per year is used to illustrate the point. The specific 1,000 hours for which the availability is used for can be deferred to close to real time: day-ahead, intra-day or even post gate closure, and dispatch of long-term assets can be integrated within arrangements for short-term CMM markets as discussed in Section 3.3.3.

3.2.1 Design choices for long-term CMM contracts

Starting from the basis of contracting for a specified number of hours per year, the following list describes three forms of contract that could be explored: firm response contracts, availability / utilisation contracts, and availability options. Each subsequent option gives the potential to commit a lower fraction of consumers' money upfront, giving a different balance between risk to consumers and value for, and therefor appetite from, providers:

- **Firm response contracting (one stage):** contracting for the firm provision of constraint management service for the specified number of hours with the price fixed upfront. A contract would be to deliver response during 1000 hours a year with a commitment by the FSO to pay 1000 hours worth of delivery. The contract would involve a process for dispatching the response on operational timescales. The FSO would be committing consumers' money upfront for the full cost of delivery.

[39] Note that the example is meant to be illustrative rather than robust quantitative outcomes, however a summary of the approach used to produce the data is given in annex A.

[40] See section 5 and annex 1 for details of the modelling used. However, whilst modelled on realistic values, plots should be treated as illustrative of the potential approach.

- **Availability / utilisation contracts (two stages):** contracting availability for a specified number of hours with two prices: (a) the availability price, and (b) the utilisation price. The contract would involve a process for nominating availability (e.g. day-ahead) and dispatch within nominated hours (e.g. at gate closure). The FSO would only be committing consumers' money upfront to the availability component of cost and incurring the utilisation payment only if the response is physically dispatched.

- **Availability option contracts (three stages):** the FSO takes out an option on availability for a specified number of hours per year and agrees three prices: (a) the option premium, (b) the availability payment, and (c) the utilisation payment. The contract would include a process of exercising the option or allowing it to lapse (e.g. on a daily basis, two days ahead); for nominating availability where options have been exercised (e.g. day-ahead where options have been exercised the previous day); and for dispatch within nominated availability hours (e.g. at gate closure). The FSO would only be committing consumers' money upfront to the option premium, and incurring the availability payment if and when the option is exercised, and incurs the utilisation payment if and when the response is physically dispatched.

The three models are summarised in Figure 10.

	Contract timescales (months to years ahead)	Two day ahead	Day ahead	Gate closure	Delivery	Year end round-up
Firm response contract	Contract for 1000 hours firm response FSO commits full utilisation costs		Dispatch response for specific settlement periods			FSO pays for full 1000 hours regardless of whether hours are dispatched or whether cheaper options were available in operational timescales
Availability / utilisation contract	Contract for 1000 hours availability / utilisation FSO commits availability costs FSO agrees utilisation price		Availability nominated for specific settlement periods	Utilisation dispatched at or after gate closure FSO commits utilisation costs		FSO pays full1000 hours availability but only pays utilisation for actual MWh dispatched
Options / availability / utilisation contract	Contract for option of 1000 hours availability / utilisation FSO commits option premium FSO agrees availability and utilisation prices	Option exercises for specific number of hours availability to be used in two days time FSO commits to utilisation costs	Availability nominated for specific settlement periods	Utilisation dispatched at or after gate closure FSO commits utilisation costs		FSO pays full1000 hours options premium, but only pays availability prices for number of hours nominated and utilisation prices for number of MWh dispatched

Figure 10: Summary of three long-term constraint management market options. The table shows actions and (in pink) the point at which consumer's money is committed for a firm response market, availability / utilisation contracts and an options based contract.

3.2.2 Identifying the potential value of constraint management markets

The value of long-term CMMs for consumers is related to two elements of the contract: firstly the potential to strike contracts with low prices, below the expected price that the consumer would face if the FSO waited either for short-term CMMs or the BM. The second element of value is in the certainty that the contract provides, and the reduction of consumer risk related to the fact that they are no longer exposed to potential future price fluctuations.

Behind the constraint

Long-term contracts behind an export constraint would aim to buy downward response. That is generation turn down or demand turn up. The counterfactual to compare long-term CMM contracts against would likely be the turn down of renewable generation within the BM. This cost would likely be related to either a unit's physical short run marginal costs (approximately zero) or its support payments (ROCs or CfDs)[41].

It is most likely that a long-term CMM would aim to buy demand increase behind a constraint rather than generation turn down as the latter may encourage new generation to locate in the region simply to try and win a turn down CMM contract, something that may be unhelpful to the system[42].

Focusing on demand, the FSO would be looking for constraint management providers willing to receive a payment of less than £50 / MWh in order to increase their demand, or would be willing to pay to consume. Depending on the timescales over which the contracts were offered this could include options which supported the development of new flexible demand and energy storage.

In some cases there is likely to be a limited pool of providers for constraint management turn down behind a constraint. For example, constraints on the B0 boundary, in the far north of Scotland, would limit providers to the area north of Inverness. Therefore an objective of long-term CMM contacts behind a constraint may be to increase the pool of providers available.

The value to providers of long-term CMM contracts is both the guaranteed revenue (either full revenue, availability, or option premium depending on the model) along with commitments for utilisation at fixed prices. Although under availability / utilisation contracts and availability option contracts utilisation is not certain, there is increased confidence to the provider that they will be used and price risk is removed as utilisation price can be locked in at the contracting stage.

In front of the constraint

The context in front of a constraint is different. Not only is the opposite response required – generation turn up or demand turn down – but in general there are significantly more providers available due to the larger geographic area.

In these locations, the FSO would aim to buy upward response to replace the turn down action behind the constraint. The counterfactual to long-term CMM contracts is the forecast cost of dispatching flexible gas generation (or in future, other forms of generation or flexibility) at the prevailing prices. As future prices are unknown, the counterfactual is uncertain, whilst the CMM contract provides certainty.

For example, if the forecast *expected* upward response price in the BM in a year's time were £100 / MWh, it may appear sensible for the FSO to limit its long-term CMM contracting to options which lock in prices below £100 / MWh. However, if a distribution of future costs is also available, for example suggesting that the range of possible future costs has a standard deviation of £20 / MWh it might be appropriate for the FSO to contract some upward response at a slightly higher price, say £110 / MWh to hedge the risk of higher prices in future. Knowledge of the probability distribution may also lead the FSO to limit the quantity it buys at £100 / MWh due to the potential that BM prices outturn lower than expected.

Box 3 discusses a useful hypothetical: whether a long-term CMM in 2019 could have hedged consumer risk related to constraint costs during the period of the 2022-23 price crisis. The discussion acknowledges the risk of 'justification by hindsight' given that it would have been very difficult to predict the specific crisis back in 2019. However, there is still value in exploring the situation from the perspective of learning lessons. In particular, what strategy should the FSO have in future given that price crises are likely in the future even if the timing and the specific characteristics are difficult to predict? Given the scale of the impact on consumers, it might be expected that the FSO should explicitly consider the risk of similar events in terms of protecting consumer value.

It can also be argued that the opposite is true: locking in prices in a long-term CMM due to the perceived risk of a future price crisis could limit the ability of consumers to benefit from an unexpected price fall. Recent

[41] Footnote 13 highlights that a CUSC code modification, P462 is under consideration to remove support mechanism costs from BM bids and offers. However, it is worth noting that regardless of a decision on P462 support mechanisms costs remain a cost to consumer regardless of whether they are recovered through BM bidding or through a side-payment as the current modification proposes.

[42] However, there may be circumstances where it might be valuable to allow generation turn-down to participate For example a one year ahead CMM contract for existing generation where a CMM contract would provide confidence to the FSO that the generator will not contribute to a constraint during the specified number of hours.

Box 3: could long-term CMMs have mitigated risks face by consumers which manifest as increased constraint costs during the recent price crisis?

In 2019-20 average constraint costs were £109 / MWh, in 2022-23 they were £370 / MWh[43]. This was largely due to increased turn up costs in front of the constraint caused by increased gas prices related to the war in Ukraine (see section 1.2, Figure 3).

Would there have been value in 2019 in NGESO offering long-term CMM contracts for upward response in-front of the constraint for delivery during 2022-23?

With hindsight it is clear that contracting in 2019, when forecasts of future prices were low, would have led to significant savings compared with the actual outturn of BM prices in 2022-23? However, *without hindsight* should NGESO have committed consumers' money at that stage given that, in common with others involved in forecasting energy markets, it was not, at the time, forecasting significant price rises?

The same question can be asked in another way: would it be sensible to consider now, in 2023, striking long-term CMM contracts to mitigate the risk of *future* price uncertainty? Should the FSO consider long-term CMM contracts to hedge the risk of a *future* price crisis?

2019: the prevailing view of relevant factors
In 2019 forecasts for the following few years showed only a relatively small increase in wholesale electricity prices. UK government's 2019 baseload wholesale electricity forecasts for 2022 and 2023 were £58 / MWh and £56 / MWh[44] (this can be compared with an outturn in 2022 of £212 / MWh[45]).

Forecasts of the development of the electricity system infrastructure were more accurate: in 2019 the FES scenarios laid out pathways with wind generation capacity between 24.2 GW (consumer evolution) and 29.5 GW (two degrees)[46] (outturn capacity in 2022 was 27 GW)[47].

Transmission network capacity has been relatively predictable in recent years. In 2019, the reported transmission capability between Scotland and England was around 5.5 GW compared with 6.6 GW today[48]. The next major upgrades remain several years away. Therefore, forecasts made in 2019 of transmission capacity over the coming few years would be likely to be relatively accurate.

The conclusion is that, in 2019, it is likely that NGESO would have been in a position to make reasonable forecasts of constraint *volumes* for 2022 based on its knowledge of system development and used the FES scenario framework to understand some of the risks. However, its *cost* forecasts – affected by the price crisis – would have been significantly lower than outturn.

What was the potential value for consumers in CMM contracts in 2019?
The increase in constraint costs from £109 / MWh in 2019-20 to £366 / MWh in 2022-23 means that the cost of turn up was around £300 / MWh. Any forward constraint management contract, signed in 2019, to cover the financial year 2022-23, and with turn up costs of less than £300 / MWh would have resulted in a benefit to consumers. For example a contract for 1 TWh of firm turn up response (e.g. 1 GW of capacity for 1000 hours) at a price of £150 / MWh would have resulted in a £150 million saving for the consumer, relative to waiting to use BM actions.

Would there have been an appetite from providers for long-term CMM contracts for upward response in 2019? Generators, for example, may have been nervous of setting a ceiling on the price of some of their output if they foresaw the potential of even higher prices if they waited. However,

[43] Values derive from *NGESO Constrain Breakdown Cost and Volume* https://www.nationalgrideso.com/data-portal/constraint-breakdown
[44] https://www.gov.uk/government/publications/updated-energy-and-emissions-projections-2019
[45] Averaged 2022 electricity day-ahead baseload contracts: https://www.ofgem.gov.uk/energy-data-and-research/data-portal/wholesale-market-indicators
[46] See 2019 Archive: https://www.nationalgrideso.com/future-energy/future-energy-scenarios/documents
[47] See 2022 historical value quoted in FES 2023: https://www.nationalgrideso.com/future-energy/future-energy-scenarios-fes
[48] See Boundary B6 in ETYS 2019: https://www.nationalgrideso.com/document/157451/download and in ETYS 2023: https://www.nationalgrideso.com/document/286591/download

Box 3 (continued)

those generators were also involved in selling power forward in the wholesale market and were willing to fix prices on some fraction of their expected output at values significantly lower than £150 / MWh. It seems likely that there would have been at least some interest in contracts that could have guaranteed generators significantly higher prices than they expected to receive through the wholesale market.

Would long-term CMM contracts at £150 / MWh have been in consumers' interests?

The answer depends on the view of risk taken at the time and on the form of contract offered. For example, if contracts were of the form of availability / utilisation the level of money committed in the contracts could be low. For example, if the overall turn up of £150 / MWh was split between £20 / MW / hour for availability and £130 / MWh for utilisation, only £20 million would have been committed upfront with the remaining £130 million incurred only if used. This form of contract would have allowed NGESO to default to alternative providers at lower costs had turn up prices available through the BM in real time outturned lower than £130 / MWh.

Remarks
This discussion doesn't aim to suggest that NGESO, or others, should have predicted the price crisis back in 2019. However, it is appropriate to ask whether the FSO should, in future, take account of the risk caused by uncertainty in future price-outturn and whether new tools such as long-term CMMs could be used to more actively manage these risks.

years also provide an example of this with the pandemic leading to reduced electricity demand and therefore lower wholesale prices. The experience of these two shocks indicates that the risk to consumers of high prices is significantly higher (due in part to the uncapped upper end of potential increases) than the potential benefit to consumers of lower than expected prices which are largely floored around zero.

3.2.3 Why would market participants be interested in participating in long-term CMMs?

Generators, consumers and flexibility providers would likely be interested in securing long-term contracts as it would provide a degree of revenue security and help them manage risk. In this respect, long-term CMMs would be like any other market contract: a tool for providing confidence over future revenue or cost streams. Long-term CMM contracts could be important in securing debt or equity to finance initial construction, and as such they could provide a useful investment-timescale incentive to locate in particular areas.

Therefore long-term CMMs can provide both a locational investment-timescale signal and ensure that the FSO has access to locational operational-timescale signals.

As an example, a hydrogen electrolyser looking to connect to the network could face significant risks related to the price they would have to pay to consume electricity in the future. A long-term CMM could provide:

- **via a firm response contract:** a guarantee of a certain volume of energy at a fixed price, potentially lower than the electrolyser could expect to achieve on the wholesale market. The CMM price would be lower than the wholesale price because of the importance of purchasing constraint management turn up in a specific location, rather than purchasing simply from the national market; and it would also reflect the value the electrolyser provides in guaranteeing the flexibility to respond to dispatch instructions on operational timescales, guaranteeing to the FSO that the consumption would be targeted and firm months or years in advance
- **via an availability / utilisation contract:** an initial base-revenue from guaranteed availability payments, plus the probability of utilisation at a price which meant significantly lower electricity costs. This form of contract includes a risk that utilisation is not called for and therefore the only value stream for the electrolyser is the availability contract and this would be a decision that the electrolyser developer would need to build in when choosing how to bid.

3.2.4 Next steps and recommendations for exploring long-term constraint management markets and contracting

This section has discussed the potential for long-term CMM contracts to provide value to GB consumers by providing opportunities to lock in low costs for future constraint management actions where confidence in the need for those actions is high, and to manage consumer risk associated with price uncertainty. It has also highlighted some of the contract designs available to the FSO that could allow different distributions of costs, benefits and risks between potential constraint management providers and consumers.

This report does not have the scope to carry out quantitative analysis of the potential value that long-term CMM contracts could provide. **Therefore the key recommendation of this section is to explore the ideas discussed further and carry out quantitative analysis of the potential value they could create.**

3.3 Short-term constraint management markets

Short-term constraint management is about supporting operational efficiency on timescales within which there is an ability to meaningfully forecast operating conditions in individual settlement periods. This is likely to be within the period day-ahead to gate closure (although it could potentially include slightly longer timescales). The forms of short-term CMM discussed here exclude the BM although some designs may use the BM as the dispatch mechanism for utilisation associated with availability / utilisation markets.

3.3.1 Design choices for short-term constraint management markets

As with long-term contracts, there are a range of models that can be used. The design of short-term CMMs can be broken down into a number of parameters:

- **Objective:** The objective should be focused on maximising and protecting consumer value and include consideration of cost minimisation and risk management. It should be closely tied to the objective of the overall constraint management portfolio. Risk could be introduced explicitly through the objective function via probability estimates, by incorporating regret costs within the objective, or via a more qualitative approach which would agree to buy actions on a 'low-regrets' basis.

  The FSO could also choose whether its objective was to attempt to fully resolve forecast constraints in using a CMM or to relieve a proportion of the constraint whilst choosing to leave some constraint volume to be managed through the BM.

- **Products and contract types:** similar to long-term CCM contracts types explored above, short term CMMs can trade firm response, availability / utilisation or a form of availability 'option'. For example if running a day-ahead auction to procure constraint management, the auction could procure firm response, requiring the provider to dispatch to the specified level no matter what else occurs between the auction and real time. Alternatively, inline with the direction of travel for other ancillary services, an availability / utilisation contract could be used. This would require the provider to guarantee availability to dispatch at a specified level but that the dispatch instruction will only be sent closer to real time, for example through the BM, post gate closure.
- **Obligations on providers in the wholesale market:** CMM contracts will likely be held by providers also active in the wholesale market. The rules around whether CMM and wholesale contracts can be stacked will influence how providers act, including how they bid into a CMM auction.

  Important rules include whether a provider is allowed to sell or buy the same capacity in both markets, and whether there is any specific obligation in the CMM availability / utilisation contracts if a provider is not given a CMM dispatch signal during a settlement period nominated for availability. This is explored further in Box 4.
- **Procurement timescales:** This report primarily considers procurement timescales of day-ahead and intraday for short-term CMMs. There may be value in exploring periods longer than day-ahead under particular circumstances, such as at times where transmission capabilities will be significantly reduced due to maintenance outages. However, given that many of the major factors with constraints on operational timescales are due to weather-related effects, it is likely that only during the period of 48 hours before delivery or closer does confidence in forecasts begin to grow (note that 'day-ahead' timescales can be as long as 40 hours ahead if, for example, gate closure on a day-ahead auction is at 8 AM for delivery during all hours of the following day).
- **Dispatch timescales:** under availability / utilization markets, dispatch timescales need to be set to suit the provider and allow them sufficient time to physically respond. For many asset types dispatch can be very close to delivery. For example batteries, many flexible generators and many forms of demand response can be dispatched within minutes or seconds. However, some classes of demand response may require longer, either due to a lack of automation on the flexible demand itself, or due to interactions with the provider's other activities such as industrial processes. There will also be the need to factor in unit-commitment considerations such as warming time for thermal generators which may need to start from cold.

Box 4: Stacking wholesale and CMM, and CMM obligations when not dispatched

Participants in short-term the CMM are likely to be active in the wholesale market. As part of the CMM design there is a choice whether participants can stack revenues from the two markets.

In this context stacking is defined as selling the same capacity in two markets. For example a 1 MW battery could (a) sell 1 MWh of power to the wholesale market during a particular hour on the day-ahead power exchange and (b) offer to provide 1 MW of turn-up constraint management in the CMM for the same hour. If successful in both, the battery 'stacks' the revenue across the two markets.

This might be seen as allowing providers to sell the same output twice. However, if both markets are competitive, the ability to participate in both and stack revenues across the two would affect the prices that participants would bid and offer in both markets and, ultimately, the clearing price of both.

In the example given, if the day-ahead power exchange cleared first and was then followed by the CMM market running an hour later, and if the battery won a wholesale contract, it may bid a very low price to provide turn up response in the day-ahead CMM. For example, if the battery operator is highly confident that it intends to honour its wholesale contract, it may bid close to zero in the CMM and take any additional revenue that was available. However, if the battery operator believed there was a reasonable chance it may wish to sell back its wholesale obligation during intra-day wholesale trading, then its CMM bid may be higher, reflecting the opportunity cost that a CMM contract create, namely the removal of the option to buy back its wholesale position.

The CMM bid would effectively be the price that the FSO has to pay at day ahead stage to buy certainty around the battery's operation and remove the risk (from the FSO perspective) that the battery reverse its position in intra-day trading.

The alternative to stacking is that a battery could hold either a wholesale position or a CMM contract for a particular hour, but not both. It would likely use significantly different bidding strategies and would have the potential to trade off the options in both markets, potentially to the detriment of consumers. This is the essence of 'inc-dec trading' discussed in section 4.1.

A second important interaction, where stacking is allowed, is between wholesale market positions and CMM availability / utilisation positions. Under a firm-response CMM market, providers could meet their CMM obligation by taking the matching wholesale market position. However, with an availability / utilisation CMM market, a provider would likely need to take a wholesale position before a final CMM dispatch signal is sent (or not).

This could be managed by two rules: (1) making explicit that a CMM provider has no obligation to operate at a particular level if a CMM dispatch signal is not sent; and (2) using the BM to adjust a unit's FPN to deliver a CMM dispatch instruction if it has taken a different position in the wholesale market.

Figure 11 shows the revenue flows for the eight combinations of scenarios for a battery that has won a CMM availability / utilisation auction for 1 MW of generation (in front of a constraint). It involves five revenue streams:

- **wholesale market revenue;**
- **CMM availability payments** which are assumed to be paid unless the battery is sent a CMM dispatch instruction but fails to deliver;
- **CMM utilisation payments** which are assumed to be paid if a CMM dispatch signal is sent and the unit dispatches at the correct level;
- **a CMM 'non-delivery' penalty** if sent a CMM dispatched signal but final dispatch does not match what is required; and

Box 4 (Continued)

- **imbalance payments** where a unit's FPN (reflecting its wholesale position), adjusted to reflect any CMM dispatch, differs from its metered output.

Under this set up there is the potential for a provider to earn wholesale market revenue and CMM availability and utilisation payments at the same time.

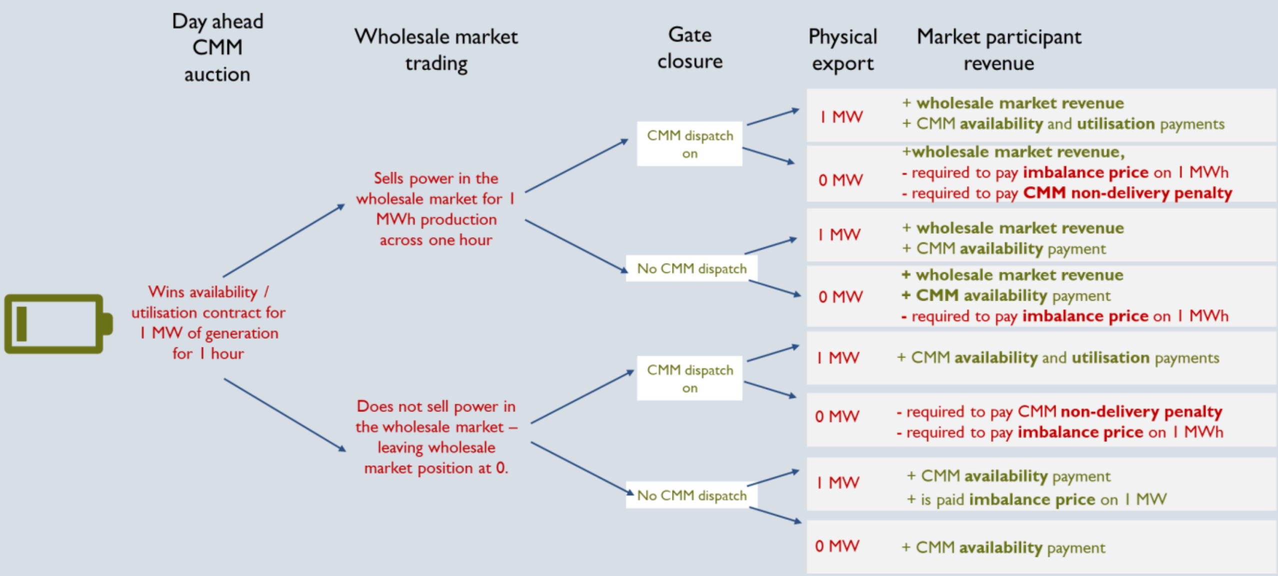


Figure 11: Potential outcomes for an illustrative situation assuming that CMM and wholesale revenues can be stacked. It represents the situation where the obligation on the provider is only to dispatch at a specified level if they receive a CMM dispatch signal and does not include an obligation to dispatch at any specific position if no CMM dispatch signal is received.

- **Counterparty:** there will need to be a range of counterparties that can contract with the FSO in short-term CMMs. For larger generators and flexibility providers who are active BM participants themselves, it is likely these parties are best placed to be the CMM counterparty. For small-scale demand response and small-scale embedded generation and flexibility it is unlikely that the FSO will want to contract directly with each individual provider. As with other flexibility services, suppliers or aggregators are the obvious counterparty options. For interconnectors (discussed in more detail in section 4.2), the most likely counterparty will be the interconnector trader who owns capacity and who themselves are active in the GB wholesale market.
- **Trading blocks:** constraint management could be traded separately for each settlement period or for blocks of settlement periods such as Electricity Forward Agreement (EFA) blocks. If trading were for individual settlement periods there are opportunities for linked-bids across multiple settlement periods. If trading were for blocks of settlement periods (for example for all settlement periods between 00:00 and 04:00) and the product traded was availability / utilisation it would allow the FSO a degree of flexibility over the exact time of dispatch. For example, the contract could be for 1 MW of availability throughout the period 00:00 – 04:00 and a utilisation price for up to 4 MWh.

These options are summarised in Table 1.

Table 1: Summary of design options for short-term CMMs

SO Objective	Product	Procurement timescale	Dispatch timescale	Counterparty	Trading periods
▪ Cost minimisation ▪ Risk minimisation ▪ No / low regret actions only ▪ Relieve constraint in the CMM or resolve constraints	▪ Firm Flexibility ▪ Availability and utilisation ▪ 'Option'	▪ Within week ▪ Day-ahead ▪ Intraday	▪ Ahead of gate closure ▪ At / after gate closure	▪ BM unit operators ▪ Consumers and embedded asset owners / operators ▪ Suppliers ▪ Aggregators ▪ Interconnector traders	▪ Individual independent settlement periods ▪ Individual but linked settlement periods ▪ Blocked settlement periods

3.3.2 Examples of short-term CMM models

This subsection describes two possible short-term CMM models in more detail.

Model 1: Day-ahead market for availability and utilisation

This model would align with other ancillary services in the current portfolio, in particular the new balancing reserve service[49] and could be integrated with the development of NGESO's 'Single Market Platform'[50].

Under this model, the FSO would run a daily day-ahead auction which closed after the wholesale power exchanges and day-ahead interconnector auctions. This would procure both availability for upward response in front of a constraint and availability for downward response behind a constraint. Dispatch would be post gate closure and could use the BM in much the same way that some response services are dispatched using the BM today.

The day-ahead auction would buy for all settlement periods for the following day, meaning that it has a look-ahead time of approximately 14 to 38 hours (assuming the auction cleared late morning for the following day). Bidding would allow linked settlement periods, helping providers reflect any intertemporal considerations such as State of Charge management for energy storage.

Contracts would be awarded on a pay-as-clear, least-expected-cost basis, with expected costs calculated as the sum of availability costs and a probability-weighted utilisation cost. Once contracted, providers would be required to be 'available' and if 'dispatched' would be required to maintain a specified output level for the duration of the settlement period. Availability payments would be awarded unless a provider was dispatched and failed to respond, and utilisation payments would be awarded if the provider was dispatched and delivered the specified output.

Interaction with the wholesale market

CMM contracts could be stacked with wholesale trading and if CMM utilisation was not dispatched there would be no obligation on providers regarding their generation or consumption.

Locations

The model is appropriate for procuring both turn down actions behind a constraint and turn up actions in front.

Behind a constraint, the FSO would be looking for options in which the expected cost of the service (availability payment plus probability weighted utilisation payment) was less than the cost actions that it would be able to take in the BM. For example, in today's market context this is likely to be the turn down of wind generation, with prices set by the combination of a close-to-zero physical marginal cost, support mechanism

[49] https://www.nationalgrideso.com/industry-information/balancing-services/reserve-services/balancing-reserve
[50] https://www.nationalgrideso.com/industry-information/balancing-services/single-markets-platform

payments, and small adjustments to reflect operational inefficiencies and increased maintenance implications. In future this might change to reflect changes to arrangements for recovering support payments[51].

In-front of the constraint the counterfactual BM action today is likely to be redispatch of flexible closed cycle gas turbine power stations that are already online with headroom available, or peaking plant capable of starting at short notice. This is related to the short-run marginal cost of gas generation, plus the premium that providers add to their costs in the BM today.

Additions and options

In addition to the main day-ahead auction a number of intra-day auctions could be held to refine the position as confidence in the volume of forecast constraints grows. This would be particularly useful for settlement periods later in the day where, at day-ahead stage, there was a significant look-ahead time.

Another design decision would be to vary the market-closure time form the CMM relative to power exchange and interconnector day ahead markets. This would reverse market participants' knowledge of other markets.

Interaction with long-term CMMs

A day-ahead availability / utilisation market can also provide the mechanism to nominate availability and dispatch capacity procured in long-term auctions (see section 3.3.3).

Potential providers

This approach is likely to be suitable for most large flexible market participants and the ability to use the BM as the dispatch mechanism means it is likely to be particularly suitable for these assets. The model could also work for smaller embedded providers providing working through suppliers or aggregators. For example, fleets of EV chargers or flexible electric heating could be aggregated into a Virtual BMU for dispatch. The supplier / aggregator would be rewarded or charged for delivery or failure to deliver. Arrangements for how revenue flows through to the ultimate provider would be left to contracts between provider and aggregator.

This model is unlikely to be suitable for interconnectors given that they cannot be dispatched post gate closure. Even if dispatch were made an hour ahead of gate closure, we currently lack mechanisms by which interconnectors can be reliably dispatched that close to real time.

It is unlikely that the model would lead to cheaper turn down dispatch of wind generation itself which would continue to lose its support mechanism if it did not generate.

Model 2: Ahead of gate closure firm response contracting

A second approach is to contract for firm-response ahead of gate closure rather than availability / utilisation. This could involve participants who need to have clarity around their operating position earlier and particularly where short-notice deviations from planned operation may not be possible. From the FSO and consumer perspective, this approach may be appropriate where it gives access to additional, cheaper providers that would be unable to participate in availability / utilisation markets, or where the FSO has high levels of confidence in its forecast of constraints.

This model is similar to the one being used in the current 'Local Constraint Market' for Scotland (see section 1.4.1).

An auction would be held, either day-ahead or intraday and would contract for firm response. This means that if awarded a contract, a participant immediately has an obligation to dispatch at the specified level during the delivery period. The auction would be a pay-as-clear, lowest-price-wins auction.

Locations

The model would be suitable for procuring turn down behind a constraint and turn up in front of a constraint. Behind the constraint the FSO would be looking for opportunities that were cheaper than wind turn down in the BM, in front they would be looking for opportunities that were cheaper than turn up costs from flexible generation in the BM, currently mainly gas power stations.

[51] As of January 2024, there is a live BSC code modification proposal to remove support mechanism payments from balancing mechanism bids and offers and pay them separately https://www.elexon.co.uk/mod-proposal/p462/

Additions and options

The timing of auctions could be varied with respect to other day-ahead and intraday markets, this would include both the power-exchanges for GB wholesale trading, day-ahead markets for ancillary services, and trading windows for interconnectors.

There may be greater risk of interactions between a firm response market, the wholesale market and the BM in comparison with the availability / utilisation market described above. This could include opportunity for gaming that is to the detriment of consumers. Therefore there may be a greater need to consider mitigations of this risk such as those discussed in Section 4.1.

Potential Providers

The firm response model lends itself to technologies that are less responsive. This includes some forms of demand response such as those linked to industrial processes which need to be scheduled. It also includes responses which are not easily able to be automated. Whilst there is an expectation that the roll-out of smart metering and smart energy management systems will increase the fraction of demand that will be automated in the future, in the short- to medium-term there is likely to be the need to provide opportunities for response that is manually dispatched.

Firm response markets may also be needed in order to integrate interconnectors into constraint management frameworks. As discussed in Section 4.2, there is an opportunity to procure constraint management from explicitly traded interconnectors (currently those between GB and France, Belgium and Netherlands) via owners of interconnector capacity. However, this response will need to be scheduled through the day-ahead and intra-day explicit auctions that run on these interconnectors and cannot be left to be dispatched post gate closure.

3.3.3 Linking long-term and short-term constraint management markets

The long-term CMMs described in Section 3.2 procure responses for a specified number of hours a year through contracts signed months or years ahead of delivery; but these contracts can leave the definition of *which* hours until close to real time. Short-term CMM structures provide a way to nominate the availability and finally dispatch capacity procured in long-term contracts.

The first short-term CMM model described above – day ahead and intraday availability / utilisation market – is the most obvious way to integrate the two. Capacity with existing long-term CMM availability / utilisation contracts are entered into the day-ahead auction. As the availability payment has already been sunk, it should be ignored in the short-term market – and long-term providers given a zero price for availability. The utilisation price for long-term capacity would remain as defined in the long-term contract.

By setting the availability price for providers with existing long-term CMM contracts to zero it will be more likely that they will be dispatched compared with new capacity that bids with a non-zero availability price. However, if utilisation prices for new capacity are significantly lower, long-term capacity can still be displaced in the short-term market. In effect, the short-term market still allows the most efficient trade-off to be made between dispatching providers with long-term contracts and those without.

3.3.4 Next steps and recommendations for short-term markets

The two market options discussed above appear to merit further consideration. Within a constraint-management portfolio it is likely that there is the potential for both market structures to operate, purchasing different products from different providers. In particular, it is likely that firm response markets will be needed to ensure that the FSO can access non-automated demand response and interconnectors. However, full reliance on firm response procurement may not deliver the best route to manage consumer risk, given the level of uncertainty remaining within some of the key parameters such as weather forecasts.

3.4 Managing multiple constraints: a constraint management market for the whole of GB

In order to develop a picture of different constraint management models, the discussion so far has focused on single constraint boundaries. However, the reality is that there are multiple constraint boundaries in GB today with the potential for more to develop in future. Figure 12 shows the main boundaries for which NGESO currently identifies future transmission needs.

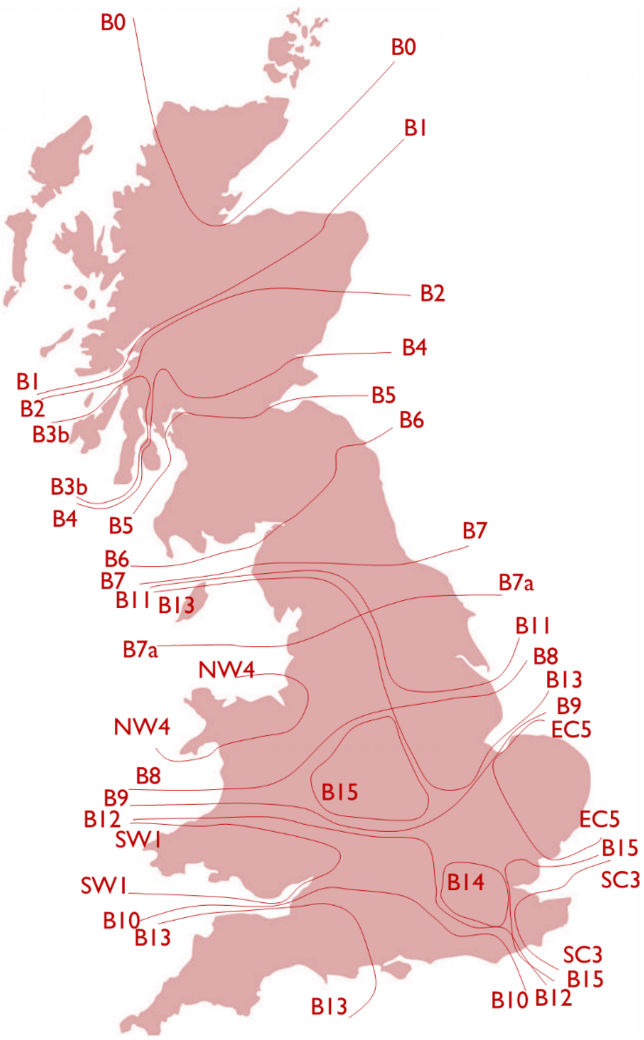


Figure 12: The main GB transmission system boundaries assessed by NGESO as part of its transmission planning[52].

Although much discussion about constraints focuses on the boundary between Scotland and England, constraints occur today across many of the major system boundaries in GB, and the number of boundaries with significant constraints is likely to increase over the next decade.

In north Britain there are a number of important boundaries starting in the very far north of Scotland just north of Inverness, this boundary is known as B0. There follow a number of constrained boundaries in series moving down through Scotland and northern England to B8 which cuts the country from mid Wales to the Humber Estuary north of Birmingham and B9 which goes south from mid Wales to The Wash south of Birmingham.

There are also several radial boundaries feeding in from other directions: from south Wales into the Midlands (boundary SW1), from south-west England towards the midlands and the central south coast (B13), from Kent up towards London and the central south coast (SC3) and from East Anglia in towards the midlands and London (EC5).

Managing serries constraints in north Britain

By the 2030s the boundaries from B0 to B8 will, under certain scenarios and conditions, experience constraints all at the same time due to power flowing south from renewable generation and interconnection in the north to load centres in the south. Because they form a linear stack, the location of response within the stack is important.

Turn down actions located north of B0 paired with turn up actions located south of B8 would relieve constraints on all boundaries and at least some sets of actions from these two geographic areas would be

[52] Note a few boundaries have been left out for readability and where they are only slight difference from a boundary shown. For example there are a number of additional boundaries which take slightly different lines through Kent and the area to the south and east of London.

needed if all boundaries were constrained. B0 turn down is therefore likely to be more valuable to the FSO than equivalent actions further south. However, it is unlikely that using all available turn down actions north of B0 will be sufficient to relieve constraints on all the relevant boundaries; it is likely that turn down actions south of B0 will be needed. Similar logic shows that turn up actions north of B8 will also be part of the lowest cost solution.

Managing radial constraints feeding into southern England

In addition to the stack of northern constraints, the radial constraints feeding into the load centres of the midlands and southern England from different directions form a different set of dependencies. These are boundaries where the region behind each constraint is completely geographically separate, but the region in front of all constraints is shared.

Managing parallel constraints will mean that completely separate sets of assets are available which can provide turn down actions for each individual constraint, whilst a single pool of potential providers for turn up actions is shared.

A framework for all constraints

The different models of CMM discussed above, both long- and short-term, can be developed into a framework for managing all of GB's constraints. The examples used so far refer to one constraint boundary. However, the logic can be applied to multiple boundaries at the same time. This will mean that when procuring CMM actions the FSO needs to consider which constraints an action can help mitigate and should value each action according to the aggregate benefit it brings across all constraints. For example, as noted above, an asset located north of B0 may provide significantly more value in a long-term CMM than an asset between B5 and B6 due to the greater number of periods of potential constraint that it could support. There may also be greater certainty over its need: two scenarios may show significantly different levels of constraint on the B6 boundary but similar levels of constraint across the combination of all boundaries between B0 and B6.

For long-term CMM contracting, an important consideration in offering contracts will be the strategic location of the asset. For example, long-term contracts for turn down with assets north of B0 are likely to provide greater value and greater flexibility to the FSO on dispatch timescales than an asset near the Scottish border.

In a similar way, contracts (either short- or long-term) with turn up assets in areas such as London and the central south coast and areas to the south of Birmingham may be more valuable than similar contracts further north in England. Assets further south would be able to provide the turn up requirements associated with different radial sets of constraints.

The approach to procuring across all constraints is best expressed as a costs minimisation: the objective is to minimise the re-dispatch cost in order to bring flows within *all* boundary limits.

4 Challenges for constraint management markets

There are a number of challenges which need to be explored and addressed in order to take CMMs forward. This section explores three particular challenges: the interaction between markets including the risk of gaming; forecasting constraints; and integrating interconnectors into CMMs.

4.1 Market interactions

One of the challenges with electricity market design is the overlap of multiple markets. Whilst any one market can be designed to deliver a particular outcome, the fact that multiple markets are required and that each market has a different set of objectives means that care needs to be taken to consider how those markets interact. In effect: is it possible to design a set of markets that are self-consistent with incentives in each working together rather than creating tensions?

The challenge of understanding market interactions is seen in many places across the electricity system: capacity markets interact with wholesale energy markets, as do ancillary service markets. Some argue, for example, that the use of a capacity market blunts the incentives driving efficient investment in peaking generators and unnecessarily subsidises capacity that the wholesale energy market would have naturally delivered. Others argue that the scale of these effects can be minimised and, if well managed, are a price worth paying for the confidence that capacity markets provide in terms of security of supply.

The main interactions considered in this report are between a CMM and:

- Wholesale energy markets
- Balancing mechanisms / markets.

There are also different ways in which markets can interact and it is worth making a number of important distinctions.

Firstly, there are general trading strategies across multiple market where an individual company aims to maximise their returns across all markets, within the rules of both and which support the two markets in delivering good outcomes for consumers.

Secondly, there are trading activities which might be referred to as gaming. The definition of the term gaming, when applied to electricity markets tends to be vague. For the purposes of this discussion, the term is used to mean trading strategies which remain within the rules of the two markets but, whilst maximising returns for a particular market participant, tend to deliver poor outcomes for consumers.

Thirdly there are situations where the design of two interacting markets accentuates the risk of market power issues where an individual market participant can influence market prices by their own actions and can specifically trade in one market in a particular way that affects prices or opportunities in the second.

Finally there are designs which could be more prone to opportunities for illegal activity such as collusion across multiple markets.

This section focuses on the risk of gaming or other strategies which might be legitimate but may lead to poor outcomes for the FSO and consumers. However, the discussions below can also identify ways to ensure that situations of market power and opportunities for illegal activity are also identified and mitigated.

4.1.1 Interactions today: wholesale market and balancing mechanism

The GB wholesale market interacts with constraint management today through the BM. In the early 2010s there was concern that BM participants behind an export constraint could exploit the constraint by offering high prices to turn down generation, and this would be a particular issue where the number of providers available to the ESO was small.

In 2012 Ofgem introduced the Transmission Constraint License Condition (TCLC) which prohibits generators behind a constraint from deriving an excessive benefit during periods of transmission constraints[53]. Guidance on the TCLC states that BM participant bids will be compared with estimates of avoided short run marginal costs (SRMC) adjusted to account for additional maintenance and ramping costs as well as any opportunity

[53] https://www.ofgem.gov.uk/sites/default/files/docs/2016/05/consultation_on_the_future_of_the_transmission_constraint_licence_condition_1.pdf

costs which could include the cost of lost support payments (e.g. ROCs or CfD uplift payments)[54]. There have been a number of recent cases of Ofgem finding market participants in breach of the TCLC[55].

The TCLC responds to several of the types of market interaction discussed above. Firstly, before the TCLC it wasn't necessary clear that the submission of *offers*[56] above SRMC would be in breach of the BM rules[57]. The TCLC also provides a route to fine market participants found to have illegitimately used market power.

In effect the TCLC imposes a quasi-regulated price on BM bids in relation to a transmission constraint. With the growth of storage capacity and other non-traditional providers of flexibility, it is likely that the definition of 'excessive benefit' under the TCLC may need to be revisited. For example, the opportunity cost faced by energy storage which bids or offers within the BM is a complex calculation that depends on its owner's overall revenue strategy and estimates of potential prices in other markets in subsequent settlement periods.

The TCLC doesn't completely manage the risk to consumer value created by interactions between the wholesale market and BM. For example, there is the risk that market participants dispatch themselves through the wholesale market at a contracted price that differs from their SRMC simply in order to be able to bid or offer in the BM at their SRMC to return to their non-dispatched position. This is challenging to identify as there is no obligation to publish wholesale contract prices which may be kept private between contracting parties. This is a form of increment-decrement trading which could also occur between a CMM and the wholesale market.

4.1.2 Increment decrement (inc-dec) trading

An inc-dec trading strategy is most easily described from a generator's perspective, but can also apply to demand and energy storage with the appropriate change of terminology. It involves a generator behind a constraint contracting in the wholesale market to *increase* its output above the level that would be economically efficient, this is achieved by contracting at prices below their SRMC. Later, the generator bids into the constraint management market to *decrease* their output and relieve a constraint at a price that covers any losses in the original transaction and more. The result is that the unit is not dispatched on but receives a net revenue for doing so.

At the same time generators in front of a constraint avoid contracting in the wholesale market, decreasing their output relative to what it would have been had they contracted at their SRMC. They then offer into the constraint management market to increase their output at a higher price than they would have received in the wholesale market.

Inc-dec trading clearly leads to negative outcomes for consumers: it creates a tendency for market participants to operate in the two markets – wholesale and constraint – in ways that artificially exaggerate network constraints in order to profit from them. Box 5 presents an illustrative example of inc-dec trading[58].

Two important questions need to be answered regarding inc-dec trading: what is the likely scale of consumer detriment that it will create? And is the scale of detriment smaller or greater than the wider benefits that the introduction of a constraint management market will create? These are empirical questions and ones that need to be explored quantitatively in the process of developing the potential for CMMs in more detail.

There are some important characteristics of inc-dec trading strategies:

- They do not require market power nor do they involve any activity that is usually defined as illegal

[54] https://www.ofgem.gov.uk/sites/default/files/docs/2017/05/2017_tclc_guidance.pdf
[55] https://www.ofgem.gov.uk/publications/ofgem-closes-its-compliance-engagement-drax-pumped-storage-limited-relation-breach-transmission-constraint-licence-condition-tclc
https://www.ofgem.gov.uk/publications/investigation-ep-shb-limiteds-compliance-tclc
https://www.ofgem.gov.uk/publications/notice-penalty-sse-generation-limited
[56] Note that term 'bid' is used in the BM to refer to actions that reduce generation or increase demand (i.e. a bid to buy power) and the term 'offer' is used to refer to actions that increase generation or decrease demand (i.e. an offer to sell power)
[57] This was particularly the case because *bids* away from SRMC do occur and are allowed in other circumstances in the BM, namely in periods of very tight margin where there is an energy imbalance. In this situations bids above SRMC are generally considered to be a legitimate way for peaking plant to cover their sunk capital costs. The TCLC clarified the situations regarding offers behind transmission constraints and made it clear that the rules did not allow market participants to make excessive returns
[58] https://www.econstor.eu/handle/10419/222925

- As such it may be reasonable to define inc-dec trading as a legitimate strategy. Even if accepted as legitimate it doesn't mean that measures shouldn't be taken to discourage it.
- The ability to implement an inc-dec trading strategy does not require that the two markets operate at the same time. For example it is possible that inc-dec strategies may be being used by some participants within the current GB market to arbitrage between the wholesale market (where the submission of unit Final Physical Notification must come before gate closure and the BM which dispatches constraint management post gate closure).
- However, it does become easier to apply inc-dec trading when two markets are accessible at the same time. For example, in a situation with a day-ahead CMM auction for firm response combined with simultaneous trading in the wholesale market the revenue from the two markets cannot be stacked[59]. there would be very little risk to a market participant who offers to sell energy in the wholesale market and bids to buy back energy in the constraint management market.
- Inc-dec trading cannot be fully mitigated against by pay-as-bid market structures, or by simple rules such as requiring equal bids in both markets (note that this last strategy isn't even possible within the GB bilaterally traded wholesale market where contract prices can be private between parties).

The conclusion of a number of studies is that inc-dec trading leads to detriment to consumers[60]. These studies also conclude that the solution is a move to locational pricing in the wholesale market. However, they do not provide commentary on what the best approach would be to manage internal constraints within a national non-locational market.

The literature suggests both that inc-dec style trading strategies may already be used within the BM in Britain today[61]. Therefore the pertinent question is not whether inc-dec trading would occur within a CMM market *per se,* but how the introduction of CMMs will change the current situation. Would it lead to better or worse outcomes for consumers in comparison to the current approach, and where CMMs deliver other benefits, how does any change in consumer impact associated with inc-dec trading compare to these wider benefits?

Exploring inc-dec trading further

Inc-dec trading is clearly a risk for market designs which layer a non-locational wholesale market on top of a locational CMM. However, there is reason to believe that the downside to consumers identified in the model studied in the existing literature can be avoided by using different market designs.

It is also important to reflect that all market designs will have imperfections. The aim of the process of designing electricity market design isn't to deliver a *perfect* market but to identify and acknowledge the opportunities and drawbacks that each market option brings with it and to select the one that works best in the context under consideration.

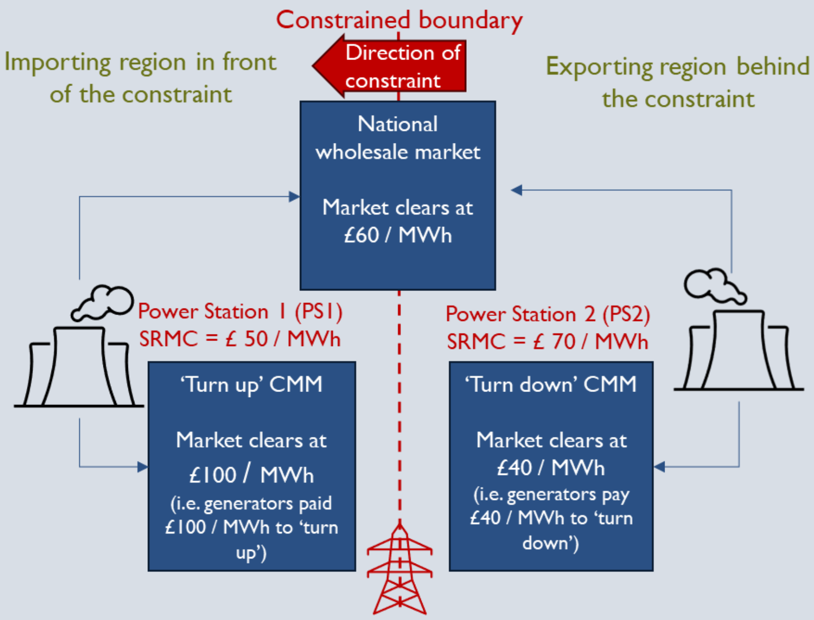
Inc-dec trading has only been explored in frameworks where stacking wholesale and constraint markets isn't allowed: the literature considers situations where participants can sell their energy in either the wholesale market or the CMM but not both. Therefore calculations about the optimal strategy include the assumption that revenue streams cannot be combined. For example in front of a constraint, in order to get any revenue through the constraint market, it is imperative that a market participants is not bought in the wholesale market. This is what leads to the need for market participants to place out-of-merit bids in one market in order to leave themselves in a position to benefit from the other.

However, this report suggests stacking revenues across wholesale and constraint markets should be explored (see Box 4 for more details). This would allow market participants to combine revenues from both markets to cover their short run marginal cost and make a profit. Under this arrangement participants acting in both markets would have an advantage over participants acting in one as they would be able to bid lower in both markets. The ability to mitigate the risks of inc-dec trading, and similar gaming and market trading strategies should be explored further under the assumption of stacking.

[59] For a description of a firm response CMM see section 3.3.2, Model 2, page 37. For a discussion of revenue stacking between wholescale and CMM markets see Box 4, in section 3.3.1.
[60] See, https://www.econstor.eu/handle/10419/222925, http://datasets.opentestset.com/datasets/Enron_files/full/gilbertsmith-d/stoft-intra-zonal.pdf, and https://www.sciencedirect.com/science/article/pii/S2211467X19300203?via%3Dihub
[61] See page 7 https://www.econstor.eu/handle/10419/222925

Box 5: Increment – Decrement (inc-dec) trading strategies

The illustration below uses the day-ahead power exchange (pay-as-clear) as an example of trading in the wholesale market and assumes that the two CMM markets are pay-as-clear markets for firm response. Whilst power-exchanges of this form operate in GB it is important to remember that the wider GB wholesale market does not follow this model. This is discussed further in the text.



Wholesale trading (Day-ahead power exchange)

- PS1 would normally bid into the day-ahead power exchange at its SRMC, £50 / MWh, willing to accept any clearing price at or above this level.
- But foresees the potential to get a more lucrative contract in the 'turn up' CMM which it forecasts would be worth £100 / MWh.
- So PS1 bids into the power exchange at £100 / MWh (its forecast opportunity cost).
- Day-ahead exchange clears at £60 / MWh with PS failing to get dispatched on (but would have been dispatched on if it had bid its SRMC).

- PS2 would normally bid into the day-ahead power exchange at its SRMC, £70 / MWh, willing to accept any clearing price at or above this level.
- But foresees the potential to gain a low price CMM contract which would require then to pay £40 / MWh to 'turn down'.
- So instead, PS2 bids at only £40 / MWh, on the expectation that it won't generate because it will get turned down later (at no more than £40 /MWh) in the CMM.
- This artificially exacerbates the constraint.
- Day-ahead exchange clears at £60 / MWh with power station dispatched on but wouldn't have been had it bid at its SRMC.

'Turn-up' CMM

- In the 'turn-up' CMM, PS2 offers at its SRMC to 'turn up': £50 / MWh
- 'Turn-up' auction clears at £100 / MWh.

'Turn-down' CMM

- In the 'turn-down' CMM, PS2 offers at its SRMC to 'turn down': £70 / MWh
- Note that the 'turn-down' CMM is a negative auction with the lowest price setting the clearing price.
- 'Turn-down' auction clears at £40 / MWh.
- PS2 has to pay £40 / MWh to reduce its output

Net result

- Wholesale revenue = 0
- Turn-up CMM revenue = £100 / MWh
- Physical position = dispatched to generate
- However, PS1 could have been dispatched in the day-ahead power exchange had it bid its SRMC

- Wholesale revenue = + £60 / MWh
- Turn-down CMM revenue = - £40 / MWh
- Net result = + £20 / MWh
- Physical position = not dispatched to generate

Box 5 (Continued)

The result is that generators increase their economic surplus at the expense of consumers for the same outcome relative to the most economic dispatch that both the wholesale and constraint market could have produced.

A number of studies[62] have highlighted that this practice has been seen in real-world markets including GB, does not require market power, and may be consider a legitimate trading strategy given that it does not require any anti-competitive behavior (as usually defined). One study also highlights that simple mitigation strategies such as pay-as-bid or a requirement to bid the same in both markets does not remove the opportunity for inc-dec trading.

In addition to the simple example given, it is also important to remember there may be additional opportunities for companies with a portfolio of assets to develop more sophisticated trading strategies.

Current studies assume all parties have access to both markets: One of the assumptions in existing studies into inc-dec trading is that all market participants have access to national wholesale and locational constraint markets. In GB, at present, this is not true. Rather all market participants have access to the national wholesale market (smaller participants access this via their supplier), but only a smaller set of market participants have access to the BM. Specifically, the following groups are excluded:

- ***Small-scale embedded flexibility,*** whilst theoretically able to access the BM via aggregators using virtual BMUs, are in practice largely excluded. The Scottish Local Constraint Market, offering a firm response CMM, offers an alternative but is, as yet, only at an early stage of development.
- ***Some battery storage*** can access the BM, however practical considerations related to control room processes mean that today they are largely excluded from constraint management.
- ***Interconnectors*** are unable to participate in the BM due to a lack of balancing-timescale market coupling arrangements.

One of the objectives of developing CMM markets is to increase the pool of potential providers for constraint management, putting downward pressure on prices.

To identify the overall impact on consumers, work is required to quantify the potential impact of gaming under today's system, the decrease in costs associated with a wider pool of providers under future CMM markets, and the change on gaming / market interaction impacts created by the introduction of a CMM.

The existing literature focuses on markets that buy firm response not utilisation / availability: There is the potential that consumer detriment associated with inc-dec trading could be reduced by a focus on availability / utilisation rather than firm response.

Under availability / utilisation arrangements participants only have certainty over the availability element of the payment; utilisation payments remain uncertain until dispatch. Whilst there may be some situations where providers may be reasonably confident of dispatch, for example if the forecast of constraints has relatively high confidence, the additional uncertainty is likely to affect bidding strategies.

The confidence of dispatch can also be affected by the quantity of availability that the FSO buys in a short-term CMM, it could choose to overbuy if availability prices are low, in order to have greater flexibility in dispatch or to account for assets which fail to deliver. This would also have the effect of reducing market participants' confidence of dispatch.

Locking in prices in long-term CMMs may mitigate gaming risk in day-ahead and intraday markets: existing studies assume that decisions on bidding strategies are made at day ahead and intraday timescales. However, the long term CMM contracts discussed in this report provide a mechanism to fix constraint market prices for some providers months or year ahead of delivery (see section 3.2). This significantly reduces the ability of providers to game between markets on operational timescales.

Whilst it is beyond the scope of this report to explore these points quantitatively, it will be important that both existing GB arrangements and specific future CMM models are assessed in terms of the risk of gaming.

[62] See footnote 60

4.1.3 Other forms of gaming and interaction

There are other ways in which constraint management can be gamed either in isolation or in conjunction with other markets.

One example identified through stakeholder feedback and relevant to current arrangements is the ability of energy storage to 'churn' the BM against its wholesale position. This strategy involves energy storage dispatching itself to generate behind a constraint in the wholesale market. At the same time it submits offers to decrease its output back to zero at a lower price than in the wholesale, assuming its BM offer is accepted it receives wholesale market revenue for full output and pays back a smaller amount (or is paid) through the BM to return its output to zero[63].

It can then repeat the process in the next period given that its state of charge remains unchanged. If the FSO is required to accept the lowest cost BM actions, it may be effectively required to perpetuate this situation throughout the constraint.

Such a situation is clearly not in the consumer interest and should be avoided. It is a situation which introducing CMMs could remove. It would not be possible to churn the constraint and wholesale markets in this way where CMMs buy services day-ahead for all settlement periods the following day, either on a firm response or availability / utilisation basis. This is because the CMM contracts would effectively set a schedule either explicitly (firm response) or implicitly (availability) for energy storage across multiple settlement periods together.

4.1.4 The day-ahead bottle neck

There are a growing number of electricity markets which have frameworks based around day-ahead timescales. These include power-exchanges which form part of the wholesale market, ancillary service markets which are increasingly moving to day-ahead procurement and interconnector trading arrangements. This means that the period between around 8 AM and late morning is becoming increasingly crowded and fitting additional day-ahead markets into the period is challenging. Existing timings are summarised in Table 2.

The order in which different markets close and then publish results is important because it affects the knowledge that market participants bring to each subsequent market. For example, if a CMM day-ahead framework involved gate closure at 07:00, then day-ahead positions or day-ahead interconnector schedules would be unknown to CMM market participants and the FSO. If CMM results were published by 07:30, these would be known when bidding in day-ahead power exchanges, ancillary service markets and day-ahead interconnector trades between GB and France, Belgium and Netherlands.

However, if the day-ahead CMM were scheduled with gate closure at 10:15, the process would be reversed with bids to the CMM being made with full knowledge of power exchange and interconnector results. In addition there are separate timetables for the wider European Single Day Ahead coupling auction (market closes at 11:00) and for NGESO's day-ahead auctions for the new suite of dynamic frequency response services (market closes 2PM).

In developing more detailed proposals for short-term CMMs, specific consideration needs to be given to the impact that different running orders for day-ahead markets would create.

[63] This strategy appears to be a credible way in which energy storage could game the wholesale / BM markets. However, as noted elsewhere in the report battery storage appears not to be used for constraint management in the BM today. Hence batteries are unlikely to be doing this today.

Table 2: Summary for timings of the main day-ahead auctions which a short-term CMM may interact with.

Market mechanism	Gate closure	Results published	Comment
Britned DA auction	07:50	08:10	Separate explicit auctions for interconnector capacity
Nemo link DA auction	08:00	08:30	Separate explicit auctions for interconnector capacity
IFA and IFA2 DA auction	08:40	09:00	Separate explicit auctions for interconnector capacity
EPEX day-ahead wholesale auction	09:20	09:30	UK government has proposed re-coupling the two day-ahead wholesale auctions[64]
NP day-ahead wholesale auction	09:50	10:00	UK government has proposed re-coupling the two day-ahead wholesale auctions
NSL capacity allocation	09:50	10:00	Implicitly allocated as part of the NP day-ahead wholesale auction.
European hourly day ahead coupled auction	11:00	11:42	

4.2 Interconnectors

Interconnectors present a unique challenge for constraint management, both in terms of the constraints they create or exacerbate, and integrating them into constraint management processes. The core underlying difference between interconnectors and other CMM participants is that any action in GB must be matched by an equal and opposite action in the connected market at the same time. The trading arrangements described below, and the options for interconnector involvement in CMMs reflect different ways of dealing with this fundamental point.

Physically, interconnectors are network assets, but networks that connect two different markets. However, each national market sees the interconnector as a producer or consumer of electricity. The European target model for electricity markets aims to use interconnectors to integrate different markets both physically and in terms of financial trading, however the UK's exit from the EU and the Internal Energy Market complicates the picture for GB interconnectors.

Engagement with stakeholders suggests that there is no fundamental reason why interconnectors cannot be incorporated into CMMs. However, there are a variety of practical considerations which will be challenging to overcome. And although the UK is outside the Internal Market at the moment, it will be important to consider whether any CMM design involving interconnectors would be compatible with a return to the internal market in future.

Although a detailed review of interconnector trading arrangements is beyond the scope of this work, it is valuable to provide an overview and introduce some of the key concepts that need to be considered.

4.2.1 Interconnector trading – an overview

Interconnectors are developed as large private sector infrastructure projects. Investment is raised in anticipation of revenue that comes primarily from price arbitrage between markets: they facilitate the flow of electricity from low price markets to high price markets, with the difference in those prices representing revenue from interconnector capacity. In addition, many GB interconnectors benefit from a cap-and-floor regulatory mechanism which shares some of the investment risks between investors and GB consumers.

Once built, interconnector owners sell the capacity of the interconnector to traders twice: once in each direction: a 1 GW interconnector between GB and France will sell up to 1 GW of capacity from France to GB

[64] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1179971/recoupling-gb-auctions-for-cross-border-trade-with-eu-consultation-government-response.pdf

and 1 GW of capacity from GB to France. Initial sales of interconnector capacity may be made a year or more in advance. These forward sales are, in effect, options which need to be exercised (a process called nomination) ahead of day-ahead trading.

At day-ahead and intraday stage interconnector flows are calculated through a series of auctions which involve the interconnector capacity itself and the trading of energy in each market. These processes can involve **explicit trading** or **implicit trading** of interconnector capacity**.**

It is at this stage that significant differences emerge between different interconnectors operating in GB today, and between today's arrangements and potential future arrangements.

4.2.2 Explicit and implicit trading

Explicit trading means that a trader must do three things:

- Purchase and nominate interconnector capacity from one market to another;
- Purchase energy in the 'from' market equal to the interconnector capacity they hold; and
- Sell energy in the 'to' market equal to the interconnector capacity.

If the trader fails to match their purchases, sales and interconnector nominations they will end up in an imbalance position in one or both markets.

Under explicit trading arrangements, the overall physical flow on the interconnector is equal to the sum of all nominations by traders who hold interconnector capacity. Currently, the interconnectors between GB and France, Belgium and the Netherlands are explicitly traded.

Implicit trading means that flows are set by a 'market coupling algorithm' linked to the relevant day-ahead power exchanges operating in the two connected energy markets. The algorithm determines the optimal interconnector flow to maximise social welfare[65] and schedules the physical flow of the interconnector to match. The interconnector still auctions interconnector capacity; however, those who buy that capacity no longer need to nominate the interconnector capacity or trade energy, that is automated by the algorithm. Holders of interconnector capacity receive revenue calculated from the price difference between the two markets with an adjustment for losses.

Within the EU, interconnector flows are increasingly implicit, with market coupling at day-ahead stage provided by an algorithm called Euphemia[66], and market coupling at intraday stage provided by a continuous process called XBID[67]. Day-ahead trading between GB and the EU also used to be largely implicit, but following Brexit implicit trading is currently only used on the North Sea Link between GB and Norway (day-ahead implicit trading only) and between GB and Ireland (intraday implicit trading only).

Implicit is widely accepted to deliver more efficient energy trading than explicit trading. Typical values quoted are around a 5% improvement in trading efficiency[68]. The scope of this efficiency is limited to better aligning interconnector flows with the real-time price differentials between markets, and doesn't include efficiency in terms of operating the system within either of the two markets. For example, it doesn't measure how efficient the interconnector is in terms of balancing requirements or managing constraints within a market.

Given the greater efficiency of implicit trading, the 2021 Trade and Cooperation Agreement between the UK and the EU, required that a form of implicit trading called "Multi-region loose volume coupling" is re-established between GB and the EU[69] in which volumes are implicitly allocated with prices calculated in a subsequent step[70].

---

[65] https://www.nordpoolgroup.com/globalassets/download-center/single-day-ahead-coupling/euphemia-public-description.pdf
[66] https://www.epexspot.com/sites/default/files/2020-02/Euphemia_Public%20Description_Single%20Price%20Coupling%20Algorithm_190410.pdf
[67] https://www.amprion.net/Energy-Market/Congestion-Management/Multi-Regional-Coupling-(MRC)-and-Cross-Border-Intraday-(XBID)/Content-Page.html
[68] For example see https://www.ofgem.gov.uk/sites/default/files/docs/2019/10/value_of_international_electricity_trading.pdf
[69] See annex 29 https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/982648/TS_8.2021_UK_EU_EAEC_Trade_and_Cooperation_Agreement.pdf
[70] See https://consultations.entsoe.eu/markets/cost-benefit-analysis-of-multi-region-loose-volume/supporting_documents/MRLVC_CBA_summary_report_April_2021_final_publication.pdf

Table 3 shows the day-ahead and intra-day trading arrangements on the existing GB interconnectors.

Table 3: Day-ahead and intraday auction arrangements for current GB interconnectors

Market connected	Interconnector	Day ahead	Intra-day
Norway	North Sea Link	Implicit	None
Irish Single Electricity Market	Moyle	None	Implicit
	East-West Link		
Netherlands	Britned	Explicit	Explicit
Belgium	Nemo	Explicit	Explicit
France	IFA	Explicit	Explicit
	IFA2		
	Elec Link		

4.2.3 Opportunities for interconnectors in CMMs

The most obvious way to involve interconnectors within constraint management markets is through a formalisation of the countertrading approach used by NGESO today (see section1.3.3) on explicitly traded interconnectors. This approach could see interconnector traders participating directly in either day-ahead or intraday constraint management markets, however it is likely to be limited (under present arrangements) to trading for firm response.

To illustrate the approach an example of intraday CMM trading is shown in Box 6. The box illustrates both how interconnector traders could participate in an appropriately designed CMM, but also highlights some of the limitations.

In terms of limitations, firstly the process would not be possible (at least in this simple form) for implicitly traded interconnectors. This is because there are no individual market participants trading energy in the GB market related to the flows on these interconnectors, rather the centralized algorithm uses information from power exchanges in both markets to set the flows and buy and sell the correct quantities of energy.

Secondly, a CMM based on availability / utilisation with utilisation dispatched close to real time would be difficult to implement for interconnectors under current arrangements. The reason is that availability products defer the final dispatch until gate closure, whilst at present there are no mechanisms to deliver balancing timescale adjustments over interconnectors, including the requisite rebalancing of the connected market. Within the EU there are arrangements to develop mechanisms to share balancing services, for example through the Trans European Replacement Reserve Exchange (TERRE)[71]. However, GB is outside such approaches.

There are options to integrate interconnectors with explicitly traded intraday auctions into CMM markets and this report recommends that these are developed as a priority and would only require a formalization of existing countertrading by NGESO. This would be possible over the existing interconnectors to France, Netherlands and Belgium.

[71] https://www.entsoe.eu/network_codes/eb/terre/

Box 6: Example of intraday CMM trading

The following example shows how a CMM market can be used to integrate existing ad-hoc counter-trading into a firm response CMM across an explicitly traded interconnector:

- At day-ahead stage the interconnector flow is fully dispatched importing into GB and is expected to exacerbate a constraint. There is high confidence that the constraint will materialise and that following intra-day trading, the interconnector will remain dispatched for full import.
- An intra-day CMM auction for firm response turn down is opened to traders on the interconnector who hold nominated capacity for import.
- Interconnector traders place offers to reduce their nominated capacity for import. This will incur costs for the traders associated with rebalancing their position in both GB and the connected market. Their CMM bid reflects these costs.
- Any interconnector trader who receives a CMM contract is obliged to adjust their position accordingly via the nomination process and neutralise that capacity, reducing the flow into GB.

There are situations where trading of this type may not provide full confidence of the aggregate response on the interconnector. For example, if the day-ahead action led to a partial flow over the interconnector – for example 1000 MW into GB on the IFA which has a capacity of 2000 MW. This means that the owners of some interconnector capacity have chosen to nominate at day-ahead stage whilst others have not yet nominated.

If a CMM were to effectively neutralise some of the interconnector capacity that had been nominated day-ahead, it is possible that owners of some of the additional interconnector capacity that had not been nominated day-ahead could still choose to do so in later intra-day trading.

Such situations could be avoided by limiting the ability for further trading to override CMM actions. In the example given, if a CMM reduced the day-ahead scheduled flow by 500MW, the overall capacity of the interconnector could be reduced by an equivalent amount.

4.2.4 Opportunities for implicitly traded interconnectors

Under the current structure of implicit trading used by some GB interconnectors, and favoured by the European Target Model, there would appear to be little opportunity to influence interconnector flows through CMMs.

Theoretically, however, it would be possible to incorporate CMMs, developing an algorithm which optimises across both the wholesale markets (represented by day-ahead or intraday trading on power exchanges) and constraint management markets. Doing so would mean introducing the CMM price alongside the wholesale price within the market coupling algorithm.

This would represent a significant innovation compared with current algorithms, and a change to the direction of travel compared with the EU target model.

4.2.5 Next steps for interconnectors and constraint management

As noted above, implicit interconnector trading is generally considered to be more efficient than explicit in terms of wholesale market trading, and this often drives the desire to return GB interconnectors to implicit trading in future.

However, retaining explicit trading may deliver greater overall benefits for GB. For example, if explicit trading is more compatible with effectively managing constraints within the GB system, and if the benefits of doing so outweigh the disbenefit of less efficient wholesale market trading, it would be of net benefit to GB to retain explicit arrangements. The costs and benefits of implicit versus explicit trading arrangements, including the potential to better manage constraints, should be explored directly with quantitative analysis.

The future of interconnector trading needs to be explored to identify the 'best possible' arrangements for incorporating interconnectors into constraint management. This should consider options for implicit and explicit trading. Where options diverge from the EU target model, this should not necessary be a reason for ignoring them although the risks of choosing to diverge from the EU model should be considered as part of the process. It will also be important to be realistic in considering how practical those arrangements will be and how compatible they will be with developing market arrangements on the far end of interconnectors.

5 Case study: managing constraints in Scotland and north England

The north of Britain provides a good case study to highlight the context for CMMs. Recent estimates by LCP suggest that nearly 90% of wind curtailment over the period 2020-21 was in Scotland[72]. Plans to build new transmission capacity to link northern generation with southern demand centres represent both a major constraint management intervention which will, in time, reduce constraint costs significantly; and a source of consumer risk stemming from potential delays in delivery. Figure 13 shows the key transmission boundaries between the north of Scotland and northern England that have the potential to contribute significant constraint costs, along with graphs showing the current capability of each boundary and the expected evolution of that capability over the coming 12 years.

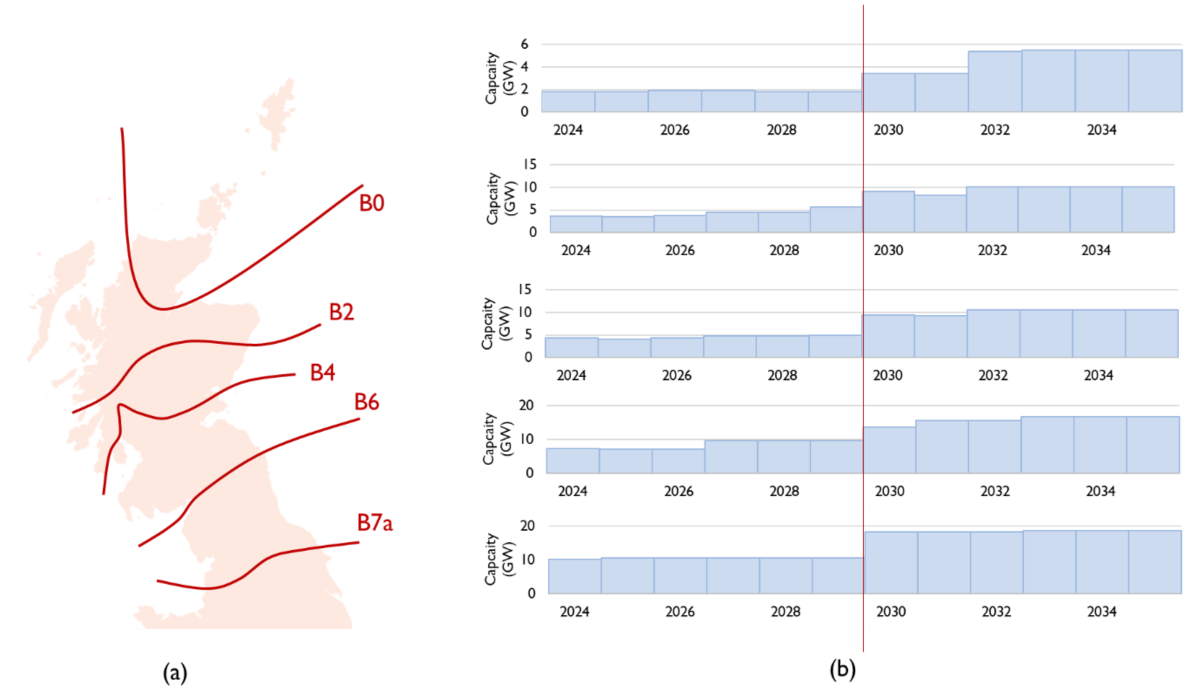


Figure 13: Some of the important strategic boundaries that will require upgrade over the coming decade together with the planned increase in transfer for each over the next 20 years[73].

Currently, much of the public debate focuses on the boundary between Scotland and England, known as B6, but there are also constraints on boundaries further north (B0, B2 and B4) and further south including B7a. Over time the relative capacities of each boundary will change. For example between today and 2029, B6 is expected to increase its capacity from 6.6 GW to 9.6 GW whilst B7a is expected to stay at its current capacity of 10.3 GW until 2030 when a number of transmission reinforcement projects mean it is planned to increase to 18.4 GW. In 2029, the two boundaries have almost the same capacity despite the fact that there is significant additional generation between the two (including nuclear power stations and the North Sea Link interconnector).

This means that in the late 2020s constraints on boundary B7a may become significantly more important relative to B6. At the same time there is expected to be growth in constraints across boundary B4, which separates central celt and northern Scotland.

The result is a shifting set of constrained boundaries which NGESO are attempting to reinforce in a sensible and coordinated way. However, as with any set of major interacting infrastructure projects, it is likely that some may deliver on time or early, whilst others will end up being delayed. There is also risk of common-cause delays across all transmission upgrade projects related to supply chains, planning, consenting and other factors. These uncertainties can be foreseen today and, if not considered explicitly, will result in a uncontrolled risk falling on consumers through highly uncertain constraint costs.

[72] https://www.drax.com/wp-content/uploads/2022/06/Drax-LCP-Renewable-curtailment-report-1.pdf

[73] Boundary capabilities from ETYS 2023: https://www.nationalgrideso.com/document/286591/download

5.1 Long-term contracting: constraints on the B4 boundary

Today B4 experiences constraints due to the large quantity of wind generation connected in northern Scotland and relatively small levels of demand. That situation is expected to be magnified over the coming decade and management of B4 constraints may be an example of where consumers would derive value from long-term contracts.

Major upgrades for the B4 boundary expected in the next seven years include the development of two off-shore HVDC links between Peterhead and northern England, due for commissioning in 2030, along with some reinforcement of the existing onshore transmission network. Figure 1 illustrates the scale of mismatch between wind availability and demand north of B4 in 2030 under the FES2023 System Transformation scenario. It shows that generation will exceed demand for almost the entire year. The grey line shows the ability to use generation output thorough either local demand or transmission export to the rest of GB (assuming the current planned boundary capability in 2030). A summary of how the modelling used to produce the time series and the analysis in Table 4 is given in Appendix A.

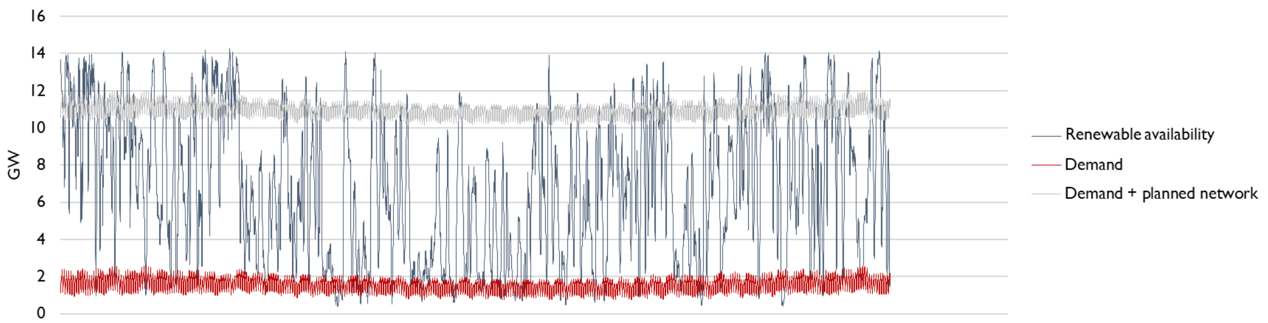


Figure 14: Illustrative profiles showing a timeseries of renewable generation available and demand north of B4 across the full year of 2030 based on FES2023 Leading the Way *scenario.*

Timeseries of this form can be used to carry out scenario analysis of future constraint levels and can be used to estimate annual constraint volume and the number of constrained hours. Table 4 shows congestion volume, hours of congestion and average depth for B4 in 2030 for a number of scenarios based on generation and demand backgrounds from FES 2023 and with B4 capability ranging from its current level up to its planned 2030 level, if all reinforcement projects are delivered on time.

The scenario shown in Figure 14, which uses the *Leading the Way* generation background would mean 10.3 TWh of congestion over 3161 hours (around 36% of the year) with congestion having an average depth of 3.2 GW. The extremes of the 10 scenarios modelled are highlighted in red (greatest constraints) and green (lowest constraints).

*Table 4: Illustrative scenario analysis of constraints on B4 in 2030 giving a range of constraint levels which might be experienced under net-zero compliant scenarios. Variables include scenario (*Leading the Way *– highest constraints, and* System Transformation *– lowest constraints) are shown along with a range of B4 transfer capabilities ranging from the current capability to the planned capability[74].*

Leading the Way			
Boundary capability (GW)	Congestion (TWh)	Hours	Average depth (GW)
4.4	32.7	5851	5.6
5.4	27.1	5273	5.1
6.4	22.1	4744	4.7
7.4	17.6	4225	4.2
9.4	10.3	3161	3.2

System Transformation			
Boundary capability (GW)	Congestion (TWh)	Hours	Average depth (GW)
4.4	18.4	4928	3.7
5.4	13.8	4245	3.3
6.4	10.0	3537	2.8
7.4	6.8	2848	2.4
9.4	2.3	1599	1.5

[74] Source: NGESO FES2023, NGESO ETYS2023, TEL analysis

Long-term contracting for 1 MW of downward constraint management availability for 1600 hours a year would be used in all the scenarios in Table 4and could reduce constraints by 1.6 GWh. Assuming a typical cost of £50 / MWh for delivering that service through the BM, if long-term contracts could be agreed for 1 MW / 1600 hour provision at less than £80k it has the potential to deliver value for consumers.

Electrolysers are one important group of potential future constraint management providers. Electrolyser capital costs for mature chemistries are currently in the range £400k - £1200k per MW[75]. At these prices, it would be likely that with the correct design of long-term CMM contracts, electrolysers would be willing to sign contracts at less than £80k / MWh per year. Through availability payments the electrolyser could access a fixed future revenue stream, and fixing utilisation payments would help the electrolyser manage price risk. The result could be a substantial incentive for electrolysers to locate in northern Scotland.

5.2 2030: a risky year for constraints

During 2022 and 2023 the level of ambition for transmission upgrade has increased. This has been seen through the development of the Holistic Network Design and the 2022 NOA Refresh. These plans have now fed through to estimated boundary capabilities published by NGESO in summer 2023 in the Electricity Ten-year Statement. These plans could see constraint costs fall in the early 2030s as the capability for transporting power across boundaries increases. Many of the transmission reinforcement projects are due for commissioning in 2030, reflecting the fact that the HND was developed in part to deliver the UK government 2030 targets. The year 2030 sees major step-changes in capability across several important boundaries[76]:

- B4: A step change of 4.5 GW from 4.9 to 9.4 GW
- B6: A step change of 4 GW from 9.6 to 13.6 GW
- B7a: A step change of 7.8 GW from 10.6 to 18.4 GW
- EC5 (Export from East Anglia): A step change of 5.3 GW from 5.9 to 11.3 GW

Given the complexity of managing the delivery of multiple transmission projects across these boundaries for delivery in the same year, there is risk that at least some projects will be delivered late. Delays to transmission upgrades would cause additional constraint costs and this would impact directly on consumer bills. It would therefore seem sensible for the ESO / FSO to take mitigation action against the risk of relay.

The illustration presented above for B4 in Table 4 provides an example of analysis which would be beneficial across all boundaries. It would help quantify constraint management volumes and costs under different scenarios and hence quantify consumer risk. By carrying out such analysis, the FSO would be in a position to take rational long-term actions to manage constraints.

5.3 Short-term markets: the characteristics of curtailment

When designing short-term CCMs it is also important to consider the characteristics of the constraints that they will be expected to manage. Figure 15(a) shows the timeseries of constraints across B4 in 2030 in the *Leading the Way* scenario in which current plans for network reinforcement are delivered on time. This corresponds to the renewable and demand timeseries show in Figure 14. At an aggregate-level constraints correspond to 10.3 TWh across 3161 hours of the year.

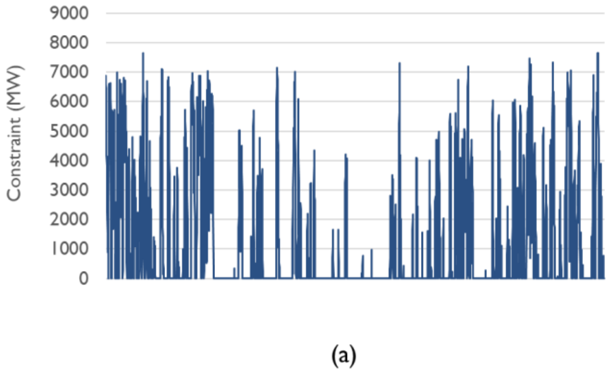
In terms of planning frameworks to manage constraints it is also important to consider how these hours are distributed through the year. The timeseries in Figure 15 (a) breaks down into 98 individual constraint events. These have a median duration of 10 hours: 50% of the events are of 10 hours or less and 50% are longer. The distribution of constraint event lengths is show in Figure 15 (b).

This highlights the importance of developing CMMs that support the right type of technology. Today's battery storage fleet consists largely of batteries with durations of well under three hours. This is because the majority of value for batteries in today's system comes from frequency response, a service which only requires durations of a few tens of minutes or an hour at most. However, when short-duration batteries began to bid

[75] Based on a USD to GBP exchange rate of $0.8 to the pound and capital costs for full Alkaine and PEM system quoted here: https://www.oxfordenergy.org/wpcms/wp-content/uploads/2022/01/Cost-competitive-green-hydrogen-how-to-lower-the-cost-of-electrolysers-EL47.pdf

[76] Note that these boundary capacities are taken from NGESOs 2023 ETYS publication https://www.nationalgrideso.com/research-and-publications/electricity-ten-year-statement-etys

for capacity market contracts, it became clear that batteries with durations of an hour or less would not be able to fully support supply during a stress event lasting four hours. In response storage has been de-rating in capacity factor auctions according to its duration, among other factors. This provides an incentive for longer-duration storage than that provided purely by the ancillary service markets. Similar points need to be considered for CMMs. In a similar way, the B4 example shows that flexibility with durations of ten hours or more will be required to manage much of the volume of constraint across this boundary. CMMs will need to be designed to incentivise long-duration storage and be capable of dispatching flexibility over curtailment events that will last for a significant part of a day, or longer.



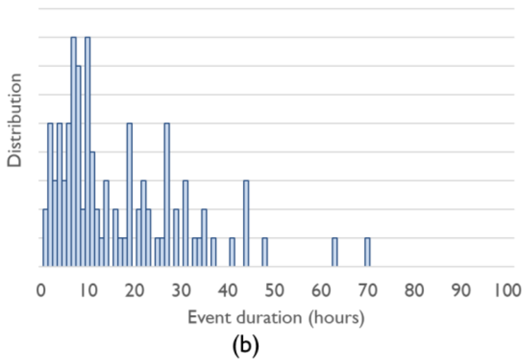


*Figure 15: Illustrative constraint characteristics for B4 in 2030 (*Leading the Way, *full HND network). (a) shows the timeseries of curtailment and corresponds to Figure 14. (b) shows the distribution of lengths for individual curtailment events.*

6 Recommendations and next steps

The report has explored the need for improved tools for managing constraint markets. It identifies that a portfolio of tools spanning all timescales within the electricity system is likely to deliver value for consumers by minimising costs and, importantly, actively managing consumer risk. In particular the report focuses on the potential to use commercial mechanisms through both long- and short-term constraint management markets combining these, where necessary, with regulatory tools. These could form the books within the existing bookends of transmission investment and the balancing mechanism, the two main tools used today to manage constraints.

The advantage of a portfolio approach, developed within the existing bilaterally-traded national wholesale market, is that it could be developed iteratively with 'quick wins' identified and introduced relatively quickly and more complex components developed later. The foundation would be a clear statement of intent together with guiding principles and an outline of the principles defining the overall objective.

As such the development of a constraint management portfolio could follow the approach taken by the development of ancillary services markets. Over the past few years these have developed from relatively static processes largely designed to award contracts to dispatchable, synchronous power stations, through bespoke but static contracts focused on particular technologies (such as the 2014 Enhanced Frequency Response Contracts) through to the implementation of dynamic day-ahead contracting through the Dynamic Containment / Dynamic Moderation / Dynamic Regulation services and the expected implementation of day-ahead contracting for Balancing Reserve.

Ancillary service arrangements therefore provide an example of a portfolio of evolving and developing tools which will stabilise as the approach becomes more mature with each element of the portfolio driven by a common overall objective.

Whilst this report points towards a potential solution it doesn't have the scope to conduct a detailed quantitative analysis of the potential costs and benefits that each of the proposed components of a constraint management portfolio could deliver.

As such the following recommendations represent a set of high-level starting points for more detailed analysis. As Ofgem highlight in their recent review of Locational Marginal Pricing (LMP), it is important that in addition to LMP we develop more concrete, realistic options which could be taken forward whilst retaining a national market. The constraint management portfolio and constraint management market options presented here should be explored, as a matter of urgency, as part of such an approach.

1. **Forecast and publish estimated annual constraint volumes and costs for future years across a range of scenarios:** estimates made by NGESO of constraint volumes and costs, such as those produced as part of the existing Network Options Assessment, should be published for a range of future scenarios and for timescales out to 2050. This would allow the sector as a whole to develop a clearer understanding and more informed debate about the challenge of managing constraints. Work should also be undertaken to quantify, where possible, the uncertainty in future constraints to ensure an understanding of risk to consumers.
2. **Carry out analysis to understand constraint forecasting on operational timescales:** detailed analysis should be conducted into the accuracy with which constraint volumes and costs can be forecast for specific days, and settlement periods over look-ahead times of weeks, days and hours ahead. This analysis should consider the degree to which different factors drive uncertainty including forecasts of weather, demand, interconnector operation, and the operation of dispatchable power stations.
3. **Develop a constraint management portfolio:** this would form a core component of a market reform option based around evolving the current national bilaterally traded wholesale market. A constraint management portfolio should have a clear objective and overall architecture agreed up front, but it should be capable of being developed in agile and flexible way, for example formalising and

integrating existing trading strategies and pathfinder projects first, before adding more complex aspects later.

4 **Define a clear constraint management objective:** A constraint management portfolio should have a clearly stated objective that is used across all timescales and tools. The objective should be based on maximising value and is likely to include a balance between minimising forecast consumer costs and managing consumer risk**.**

5 **Types of constraint management actions:** Within the portfolio, include options for both downward constraint management actions behind an export constraint, and upward constraint management actions in front of a constraint.

6 **Timeframes:** Structure the portfolio to include both long-term (e.g. 3 months to 12 years or longer) and short-term (day ahead and intraday) CMMs. There should be the potential for long term CMMs to contract flexibility for any timescale after the option for network investment has passed or for the lifetime of the asset**.**

7 **Long-term CMMs with competitively awarded contracts:** Explore the value of long-term contracts for availability to provide constraint management actions, awarded through competitive tender or auctions. These contracts could provide investment signals for new investment in flexibility or strategically situated demand. They can support delivery of good outcomes for consumers by acting as 'options' which lock in volumes and prices providing hedges against uncertainty**.**

8 **Short-term day-ahead and intraday CMMs:** Develop options for a set of short-term constraint management markets which, collectively, are accessible to the full range of potential providers: BM participants, individual domestic and business consumers potentially through aggregators, EV and heat pump fleet operators, non-BM embedded generation and flexibility, and interconnectors.

9 **Markets for availability / utilisation and markets for firm response:** Short-term constraint management market designs should prioritise, where possible, allowing the FSO to procure availability at day-ahead and intraday stage, with utilisation costs incurred closer to real time. However, complementary services, including purchase of 'firm response' ahead of gate closure may be required to allow some providers to offer constraint management who would not be in a position to participate in an availability / utilisation market design.

10 **Integrating explicitly traded interconnectors:** Develop specific options to integrate explicitly traded interconnectors into the short-term CMM designs developed in recommendations 7 and 8.

11 **Interconnector expert group:** Bring together a group of experts in interconnector trading with a mandate to develop the 'best possible' set of arrangements to allow interconnectors to participate in CMMs. This group should take account of the practicalities of trading arrangements including auction timing, explicit vs implicit trading, the impact on the connected markets, the direction of travel laid out in the EU target model and the EU-UK Trade and Cooperation Agreement, along with the pros and cons of diverging from that model in GB.

12 **Understanding market interactions:** Commission work to explore and understand the risks associated with interactions between constraint management markets, the wholesale energy market and the BM. This work should carry out analysis to compare any disbenefits that negative interactions such as gaming might create against the overall benefits that CMMs could deliver.

7 Annex A

The analysis presented in section 5 is for illustrative purposes only. This annex provides a brief description of the approach taken to the analysis.

- The case study presents a one-year time series of demand and renewable generation availability for the north of Scotland, the region north of boundary B4 which separates the Scottish Hydro Electricity Transmission (SHET) and SP Transmission (SPT) regions.
- The generation availability timeseries represents an illustrative 2030 north of Scotland renewables fleet including offshore wind, onshore wind and solar.
- **Generation capacities** for each technology type are taken from the *Leading the Way* and System Transformation scenarios from FES 2022[77], which include a breakdown of generation capacity by the network to which it is connected. As boundary B4 separates the SHET and SPT transmission network and the SHEPD and SPD distribution network this allows both distribution and transmission connected capacities to be allocated specifically to the north of B4 region.
- **Renewable availability timeseries** are created by a using meteorological reanalysis dataset created by the University of Reading[78]. This initially creates timeseries of meteorological parameters (wind speed at specific heights and solar irradiance) as an hourly timeseries for historic years. It then processes these to give estimated wind and solar capacity factor timeseries on an hourly basis.
- In the case study, the capacity factor timeseries is multiplied by installed capacity to give the hourly renewable availability timeseries. The case study uses historical weather data for the year 2007 applied to the 2030 generation capacity for the scenarios described above.
- **Demand timeseries:** FES 2022 scenarios are used to estimate the annual energy demand in the north of Scotland region for three categories of demand: electric vehicle charging, electrified heat demand (specifically for new heat pumps), and traditional electricity demand.
- For each demand category a methodology is used to convert annual energy demand into an hourly time series:
  - For EV the timeseries is based on work by Element Energy for an NGESO NIA project which developed an EV vehicle charging behavior profile[79].
  - For heat demand a model is used based on correlation with GB gas demand and a north-of-Scotland hourly timeseries of demand from the meteorological reanalysis data 2007. This model has been developed by Shanay Skellern, Callum MacIver and Keith Bell, University of Strathclyde.
  - For traditional electricity demand historic profiles are normalised and scaled.
- Total demand timeseries are calculated by scaling the profile for each component by the relevant annual demand.
- **The transmission capacity** quoted is taken from the 2023 Electricity Ten Year Statement and reflects the secure transfer capability of the B4 boundary[80].

[77] https://www.nationalgrideso.com/future-energy/future-energy-scenarios/documents
[78] https://researchdata.reading.ac.uk/321/
[79] https://smarter.energynetworks.org/projects/nia_ngso0021/
[80] Data available in excel format (note link downloads file directly: https://www.nationalgrideso.com/document/286691/download)