nationalgrid

Duration-Limited Storage De-Rating Factor Assessment – Final Report

EMR Modelling

December 2017

Duration-Limited Storage De-Rating Factor Assessment

November 2017

Disclaimer

This guidance document has been prepared by National Grid Electricity Transmission plc (NGET) and is provided voluntarily and without charge. Whilst NGET has taken all reasonable care in preparing this document, no representation or warranty either expressed or implied is made as to the accuracy or completeness of the information that it contains and parties using information within the document should make their own enquiries as to its accuracy and suitability for the purpose for which they use it. Neither NGET nor any other companies in the National Grid plc group, nor any directors or employees of any such company shall be liable for any error or misstatement or opinion on which the recipient of this document relies or seeks to rely other than fraudulent misstatement or fraudulent misrepresentation and does not accept any responsibility for any use which is made of the information or the document or (to the extent permitted by law) for any damages or losses incurred.

Copyright

No part of this document may be reproduced in in any material form (including photocopying and restoring in any medium or electronic means and whether or not transiently or incidentally) without the written permission of National Grid.

Copyright National Grid 2017, all rights reserved.

Contact

Any enquiries regarding this publication should be sent to National Grid at
emrmodelling@nationalgrid.com

Table of Contents

Executive Summary

The EMR Delivery Body, within the National Grid System Operator (SO), launched an industry consultation[1],[2] in August on a proposed methodology for updating the de-rating factors applied to duration-limited storage in the GB Capacity Market (CM). This was driven by the BEIS CM policy consultation launched in July[3], which suggested that the limits of new forms of short duration storage entering the market may need to be reflected in the CM auctions. Some of this new storage may be designed to have maximum durations as short as 30-minutes based on requirements in ancillary services, whereas modelling work conducted as part of this study suggests that CM adequacy stress events, if they were to occur, could last ~ 2 hours duration on average when the system is at the CM target reliability level of 3 hours LOLE per year, the GB reliability standard (see Section 6, page 26 for indicative histograms of GB system stress event durations).

While to date there has been one single de-rating factor for all storage based on the historical technical availability of pumped hydro at times of peak demand (96.11%), in future there is proposed to be a range of de-rating factors for storage sub-class durations ranging from 30-minutes up to around 4 hours. Moving to this alternative approach ensures that there is a transparent and fair means to account for short-duration storage contributions to security of supply and thus facilitate it's entry in to the CM, while at the same time ensuring that consumers pay an appropriate amount for the total CM capacity necessary to meet the GB Reliability Standard.

Our new proposed storage de-rating factor methodology presented in this report uses an Equivalent Firm Capacity metric (EFC), and has been designed to account for any such duration-limits when calculating the contribution to security of supply. An EFC is a very useful metric to normalise the security of supply contribution of non-conventional adequacy resources in the CM. It is defined essentially as "*for an given penetration of that resource, what is the amount of perfectly reliable infinite duration firm capacity it can displace while maintaining the exact same reliability level*".

Our industry consultation in August touched on a wide range of technical modelling attributes of duration limited storage in the GB market. We received a number of useful comments and feedback from industry which helped to inform the numerical modelling we have conducted since then. This report presents further detail on the methodology we have used, as well as the final de-rating factors that are applicable for the upcoming 2018/19 T-1 and 2021/22 T-4 CM auctions to be held in early 2018. We have also conducted a large amount of sensitivity analysis on the modelling parameters and assumptions of interest vis a vis their impact on these final de-rating factors.

From a CM governance process point of view, the introduction of the new storage de-rating factors based on our new methodology is a matter for the Secretary of State to decide. In that respect, we note the Government Response published today[4] which indicates that these de-rating factors will in fact be used in the upcoming 2018 CM auctions and thereafter.

---

[1] National Grid EMR Consultation on Duration-Limited Storage De-Rating - Methodology
https://www.emrdeliverybody.com/_layouts/15/listform.aspx?PageType=4&ListId=%7B612853B7%2D6D9A%2D4B23%2D9686%2D6AB632015C31%7D&ID=128&ContentTypeID=0x010400626754A76E41C74FA81B4D17EBF15511

[2] National Grid Consultation on Duration-Limited Storage De-Rating - Response
https://www.emrdeliverybody.com/_layouts/15/listform.aspx?PageType=4&ListId=%7B612853B7%2D6D9A%2D4B23%2D9686%2D6AB632015C31%7D&ID=134&ContentTypeID=0x010400626754A76E41C74FA81B4D17EBF15511

BEIS CM Policy Consultation
[3] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/631842/CM_consultation_-_detailed_proposals-template.pdf

[4] Government Response to consultation:
https://www.gov.uk/government/consultations/capacity-market-consultation-improving-the-framework-detailed-proposals

The Table E1 here summarises the applicable de-rating factors for duration-limited storage in the 2018/19 T-1 and 2021/22 T-4 CM auctions. These results are derived with studies conducted on an updated version of the 2017 EMR 5 year Base Case, with the most recent 'best view' of future storage capacity and duration penetration, and GB peak demand projections.

Table E1 – CM De-Rating Factors Proposed for Duration-Limited Storage Class in the 2018/19 T-1 and the 2021/22 T-4 Auctions

Final De-Ratings Per Duration in Hours	"2018/19"	"2021/22"
Storage Duration: 0.5h	21.34%	17.89%
Storage Duration: 1h	40.41%	36.44%
Storage Duration: 1.5h	55.95%	52.28%
Storage Duration: 2h	68.05%	64.79%
Storage Duration: 2.5h	77.27%	75.47%
Storage Duration: 3h	82.63%	82.03%
Storage Duration: 3.5h	85.74%	85.74%
Storage Duration: 4h +	96.11%	96.11%

Note that consistent with the Government Response, the table above presents the de-rating factors up to 4 hours maximum sub-class duration, with that duration and above receiving a de-rating factor consistent with the average technical availability of pumped hydro.

Based on the feedback we received from our storage de-rating methodology consultation with industry, we have also conducted a large amount of sensitivity analysis in our numerical modelling studies (see Section 4 of the report for details) to better understand the key assumptions and parameters influencing the EFC-based results. Sections 2 and 5 of this report outline these modelling assumptions and subsequent findings in detail, and indicate that:

- The capacity MW amount and constituent duration of the storage fleet assumed in the modelling Base Case has a material impact on the EFC results derived thereafter. For a greater penetration of duration-limited resources in the study case, then the marginally lower the outturn EFC results will be, as stress events will tend to be longer and the incremental contribution of short duration resources to security of supply will begin to saturate. For the 2018/19 and 2021/22 CM delivery year de-rating factors listed in the table above, these are based on an assumption of ~ 2.74GW of existing pumped hydro penetration in each case, as well as ~ 400MW and ~ 1,000MW of battery storage penetration respectively in either year. These updates reflect the EMR Base Case 'best view' of such issues at present, but will be updated annually as we conduct the de-rating factor analysis for the Electricity Capacity Report in each year.
- The use of the 'incremental' EFC of a small storage unit added to the margin at the point which the CM is expected to deliver is a more sensible approach to base the storage de-rating factors upon than using the 'average' EFC of the entire storage fleet overall. This is in keeping with the economic principle of payment in a market being linked to the marginal contribution of supply to meeting demand at the point at which the market is expected to clear. Furthermore the 'incremental' approach allows direct disaggregation of the

contributions of different storage durations to security of supply – the overall average EFC is a mixture of influences from all storage durations combined. The table of the final de-rating factors above is based on the incremental EFC approach.

- The statistical risk metric upon which the EFC values are derived also has a strong impact. Our modelling results show that Expected Energy Unserved (EEU) has a superior performance for conducting an EFC assessment of a duration limited resource than does Loss of Load Expectation (LOLE). This is because EEU is not as sensitive to the operational strategy of the storage during a stress event, and further, it has a direct link to the customer's economic cost of unreliability in the system by virtue of the GB Value of Lost Load (VoLL) parameter
- The EFC results are only moderately influenced by the underlying adequacy margin (LOLE level) of the EMR 5-year Base Case, the mean time to repair parameters of conventional supply, and the size of the incremental storage unit applied to the margin in the EFC reliability assessment

Throughout this work we have consulted with BEIS and their Panel of Technical Experts (PTE), who have endorsed our methodology and the proposed de-rating factors to be used. Full detail of the PTE commentary on this work is contained in Appendix 4, but an indicative summary quote is as follows

"The Panel of Technical Experts (PTE), having reviewed National Grid's (NG) proposed Storage De-rating Methodology, is content with the proposed approach and notes that NG has undertaken a convincing piece of work. The analysis to determine de-rating factors for storage is very thorough and based on appropriate fundamental principles in the context of the present CM framework, leading to the proposed approach to derating storage being robust."

We also acknowledge the benefits of discussions with academic experts from The University of Edinburgh on a number of high level issues, notably in relation to the choice of risk metric and the definition and calculation of the storage EFC.

Though we have consulted widely on the methodology used this year, we have also noted a number of modelling improvements that can be applied in subsequent years as the amount of duration-limited storage builds out and we have more experience with, and operational data from, this new form of capacity. Some of these future modelling areas are listed in Section 8 of this report.

We look forward to engaging with industry and other stakeholders in due course as the storage de-rating factor modelling approach evolves.

1. Introduction

This report summarises the results of the National Grid SO (the EMR Delivery Body) investigation in to provision of an alternative de-rating factor methodology for duration limited storage in the GB Capacity Market (CM).

This work was motivated by Ofgem's response earlier in 2017 on industry proposals to change the de-rating factors for short duration storage[5], as well as the BEIS Panel of Technical Experts' (PTE) commentary[6] on the 2017 Electricity Capacity Report[7], both of which encouraged National Grid to consider the issue further with a summer research and development project.

Storage is of course not a new form of capacity on the GB system, with pumped hydro being an important element of the generation portfolio here for many decades. However, with its relatively long duration, then pumped hydro has historically been considered to be almost equivalent to conventional plant in security of supply contribution terms, assuming the upper reservoirs are maintained full to capacity at time of peak system demands.

In recent years, there is a great amount of interest in new forms of storage capacity in response to growing flexibility needs in a changing GB system. This has led to an increase in the amount of battery storage on the system, with ~ 500MW of such capacity winning up to 15 year contracts in the 2020/21 T-4 auction, as well as ~ 200MW winning 4-year Enhanced Frequency Response (EFR) ancillary service contracts in the recent National Grid tender for this fast acting grid stability service (note that a majority of the EFR contracted amount also has a CM contract).

There are concerns however that some of this new storage capacity may be of relatively short duration at full power output, with some industry sources suggesting that much of it may be only 30 minutes duration. Early analysis by National Grid suggested that GB system adequacy stress events if they were to occur, could last for longer than this, circa 2 hours on average – we have updated that analysis in this report and the findings are consistent (see Section 6). If left unchecked, this recent pattern could lead to over-rewarding of such capacity in the CM, as well as overall GB system reliability being endangered.

Figure 1 – Duration Limited Storage Contribution to Stress Events

We note that the policy intent seems to be not to create a barrier to short duration storage participating in the CM as it has noted wider system benefits in the evolving GB power system - the goal is simply to make sure that any contributions from storage in a narrower security of supply context are fairly rewarded in the CM. In this respect, BEIS launched a CM policy consultation

[5]Ofgem response on proposed CM rules changes:
https://www.ofgem.gov.uk/system/files/docs/2017/03/statutory_consultation_on_amendments_to_the_capacity_market_rules_2014_final_23032017_0.pdf

PTE commentary on the 2017 ECR:
[6] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/625885/PTE_Report_2017.pdf

The 2017 ECR:
[7]https://www.emrdeliverybody.com/_layouts/15/listform.aspx?PageType=4&ListId=%7B612853B7%2D6D9A%2D4B23%2D9686%2D6AB632015C31%7D&ID=116&ContentTypeID=0x010400626754A76E41C74FA81B4D17EBF15511

earlier this summer with a view to revisiting the de-rating factors applied to short duration storage in the CM[8].

National Grid in response subsequently broadened out the early research and development work and launched a storage de-rating factor assessment methodology consultation[9], proposing that new de-rating factors for storage should be based on the concept of an Equivalent Firm Capacity metric. We held two industry workshops on the 10th and 15th of August, and received a great deal of useful written and verbal feedback from participants on the proposed modelling approach that could be used. Our formal response to this consultation was published at the end of August[10], and we are thankful for the feedback received.

The second half of our development work in the Autumn 2017 has involved the numerical assessment task associated with the earlier methodology proposed during our consultation. We have now completed this analysis and this report presents the results of that work, the derating factors that are applicable to future CM auctions, as well as detailed sensitivity case studies which help to understand the influential parameters and modelling assumptions within the new proposed approach.

This report is structured in the following manner:

- Section 2 details the background to the choice of the EFC modelling methodology, as well as necessary assumptions for storage operation in the reliability simulation
- Section 3 outlines the details of the time sequential reliability simulation, and the tool used to conduct it
- Section 4 details the storage penetration level and duration assumptions for the various Base Case and sensitivity analyses that we have conducted as part of this overall study
- Section 5 presents the EFC results for the various study cases
- Section 6 gives an indicative view of the histogram of stress event durations when the GB system has a plant margin equating to 3 hours LOLE, the GB reliability standard and CM target reliability level
- Section 7 then gives an indicative guide to the de-rating factors that are applicable to the upcoming T-1 and T-4 CM auctions
- Section 8 presents a discussion of the broader outcomes as well as lists a number of areas for future work to improve the modelling accuracy and assumptions set
- The Appendices present a summary of the review comments from the BEIS independent Panel of Technical Experts, miscellaneous details on methodology commentary from academic experts at the University of Edinburgh, a discussion on solar PV modelling, as well as a review on conventional plant mean time to repair parameters.

As with any new modelling methodology for an emerging system resource, the de-rating factor assessment process for duration-limited storage will be refined with experience and knowledge gained in the coming years, and we welcome any feedback from all stakeholders on any improvements that can be made in the future.

BEIS CM Policy Consultation
[8] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/631842/CM_consultation_-_detailed_proposals-template.pdf

[9] National Grid EMR Consultation on Duration-Limited Storage De-Rating - Methodology
https://www.emrdeliverybody.com/_layouts/15/listform.aspx?PageType=4&ListId=%7B612853B7%2D6D9A%2D4B23%2D9686%2D6AB632015C31%7D&ID=128&ContentTypeID=0x010400626754A76E41C74FA81B4D17EBF15511

[10] National Grid Consultation on Duration-Limited Storage De-Rating - Response
https://www.emrdeliverybody.com/_layouts/15/listform.aspx?PageType=4&ListId=%7B612853B7%2D6D9A%2D4B23%2D9686%2D6AB632015C31%7D&ID=134&ContentTypeID=0x010400626754A76E41C74FA81B4D17EBF15511

2. Methodology

There are a number of modelling complexities related to the valuation of duration limited storage in a capacity market that need to be taken into account when determining what is a sensible approach to deriving a de-rating factor. These modelling aspects were the focus of our industry consultation earlier this summer, and are summarised under the following sub-sections. Associated commentary from the academic experts is included in Appendix 3.

How to value duration-limited storage on a consistent basis with other CM capacity sources?

The CM auctions are targeting a specific level of capacity to secure in relation to the underlying GB reliability standard (3 hours loss of load expectation (LOLE)). The finite duration limit of storage implies that its nameplate capacity alone cannot be compared directly to other infinite duration capacity sources on the system, whose contribution to security of supply is represented by their average availability at time of system peak demand. Furthermore, projects within the overall storage class can have a range of durations based on the size and capability of each, and thus a distinct de-rating factor would be required for each duration. Initially therefore, this introduces a challenge of consistency of treatment in the CM which has a fundamentally desirable principle of technology-neutrality at its core. What is required is a means to translate the security of supply contribution of duration-limited storage back in to a consistent basis with that of all other conventional capacity sources.

One possible approach could be to develop a simple rule based on e.g. the “maximum typical length of a CM stress event” on the GB system, and de-rate duration-limited storage accordingly as a fraction of this stress event length. For example, if the maximum typical stress event length on the GB system was 4 hours, then a 1-hour duration storage device might be arbitrarily allocated a 25% de-rating factor, a 2 hour device a 50% de-rating factor, etc and so on. However, as one will see later in the results Section 6, CM stress events if they were to occur can have a range of possible lengths on the GB system so there is no objective “typical” stress event length. Also, a major drawback of this very simple type of rule is that it would not have a scientific basis to explicitly link storage de-ratings (i.e. CM payment) to the resource’s contribution to risk reduction (as defined by commonly used statistical metrics such as LOLE or expected energy unserved (EEU)) which the GB CM is explicitly designed to target.

There is however a suitable template for this challenge in the existing treatment of variable and uncertain renewable sources such as wind power in the GB security of supply assessment. Similar to storage, the contribution of such sources to system reliability is also complicated in that their nameplate capacity is not a direct reflection of their adequacy contribution, as their power productions are subject to the vagaries of the weather. Even though wind power is not a CM auction participant at present, there is still necessarily an account of it using an Equivalent Firm Capacity (EFC) in the annual Electricity Capacity Report as the contribution from wind power reduces the amount of CM auction capacity that is required to meet the GB reliability standard.

An EFC is a very useful construct to normalise the security of supply contribution of non-conventional adequacy resources, and it is defined essentially as “for an additional penetration of that resource, what is the amount of perfectly reliable infinite duration firm capacity it can displace while maintaining the exact same risk level”.

Therefore, using this template in the CM, a sensible means to translate the contribution of storage in to the same framework as all other capacity sources would also be via a de-rating factor based on the EFC, allowing a fair and scientifically justified comparative basis between all sources once again. As each storage duration will have a different contribution to risk reduction, therefore each storage technology duration sub-class would also have a different EFC.

Incremental EFC for each duration sub-class is more appropriate in a CM de-rating than the average EFC of the entire storage fleet

One noted effect with the use of an EFC metric as applied to a technology class is that the incremental value of it on the system may change as the penetration of that resource grows. What this may imply is that the sum of individual EFCs of individual projects may be different from the total or average EFC of the entire fleet of that resource class.

This has been observed with wind power from previous experience, whereby the wind EFC, when expressed in percentage terms of the entire fleet capacity, has reduced over time as the amount of installed capacity on the system has increased. As will be seen in the later results Section 5, this effect also extends to duration limited storage. For storage, the effect is driven by the fact that as more duration limited resources are built on the system then the length of adequacy stress events that can be expected changes – thus the contribution of even more duration-limited resources will tend to saturate as both the overall storage fleet capacity and its constituent durations evolves in time.

As wind power is currently not a CM participant, the impact of its changing EFC over time is quite straightforward to handle. We use the overall total EFC of the entire wind fleet in any given CM target year to reduce the amount of capacity to secure from other CM participant sources in the auctions. If the overall wind fleet EFC changes over time, then this has no impacts beyond the changing the total residual CM capacity to secure in each year accordingly.

For a CM participant class such as storage the EFC treatment is necessarily more complex however, as there will be a distinction between the security of supply contribution from long-term CM contracts awarded in previous years to those that will be awarded in upcoming auctions. Furthermore, for storage itself, there will also be a distinction between the contributions of each storage duration sub-class to the overall average EFC of the entire fleet in any given year.

A sensible principle upon which to base the CM auction would be that the payment should be linked to the contribution of each resource to security of supply in that auction year. Therefore, as the legacy of build out of storage capacity in previous years impacts the overall EFC of the subsequent storage fleet, then the contribution of new capacity contracts to security of supply will be different, and that ought to be reflected in the de-rating factors derived.

The means to handle this effect is by defining an "incremental EFC" of each of the storage duration sub-class around the system reliability operating point at which the CM is targeting for the delivery year in question – i.e. 3 hours LOLE. This is in keeping with the general economic principle of payment made in a market being linked to marginal contribution of each supply source to the overall resource requirement at the point of which the market is expected to clear.

For clarity, an indication of the methodology carried out to calculate both the average EFC of the entire storage fleet and the incremental EFCs of each storage duration subclass, in this modelling exercise is as follows:

- ❑ Set up the EMR 5-year Base Case with a credible generation supply portfolio, for the given CM target horizon year, and with a specific baseline reliability level at 3 hours LOLE (the GB reliability standard and CM delivery year target)
- ❑ Remove <u>the entire</u> storage fleet and recalculate via simulation the amount of perfectly-firm, infinite-duration capacity to add back in to bring the system to the same reliability level (<u>using the most relevant risk metric</u>) as the Base Case had – this firm capacity is deemed the "**average**" EFC
- ❑ Add a storage resource <u>of suitably small capacity/energy limit</u> to the original Base Case (e.g. 100MW power and variations of 0.5 – 24 hours duration limit) and recalculate the improved risk level (<u>using the most relevant risk metric</u>) via reliability model simulation for each duration
- ❑ Assess the level of perfectly-firm, infinite-duration capacity, that when added to the same Base Case, would give the same change/reduction in risk for each of the

incremental storage unit MW/MWh combination cases above – these values are then defined as the '**incremental**' EFCs of each respective storage duration sub-class

Any differences between the sum of incremental EFCs and the total fleet average EFC may have some impacts on the CM auction treatment of storage. However, any distinctions between the overall expected average EFC of the entire storage fleet and the sum of incremental EFCs of successful participants in each T-4 CM auction can be accounted for with an update in the T-1 auction's target capacity level. Any such distinctions at the T-1 stage between the expected average EFC of the entire storage fleet and the out-turn post-auction value should be small as the amount of new build capacity in each T-1 auction is usually low.

Storage response to a CM stress event is uniform regardless of transmission or distribution connection status and/or Balancing Market (BM) participant level

The diagram below indicates the sequence of events, CM and system operational notices, that may occur on the GB system in the lead up to a CM stress event. Note that (in the absence of internal GB network transmission constraints) a CM stress event is broadly defined as the application of emergency customer load control actions such as voltage reduction for a period of more than 15 minutes with a CM notice having being issued at least 4 hours in advance by the EMR Delivery Body. A CM stress event is also only officially defined by ex-post analysis (i.e. it usually takes up to a few days of post-processing time to assess if a CM stress has indeed happened) and thus no market actor can know for sure when a CM stress event starts or finishes in real-time. For further more extensive detail on the definition of CM stress events, and the various market information sources and system operational notices that may be used as a guide to them, see here[11].

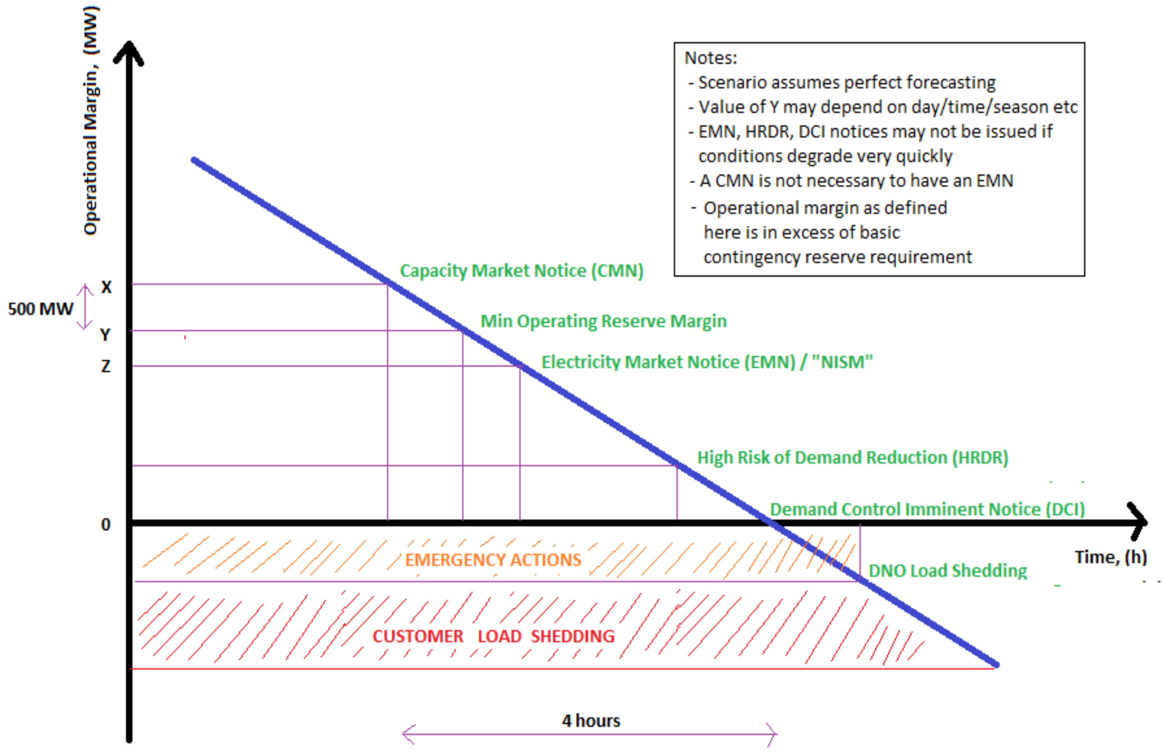

Figure 2 – Indicative GB System Margin and Operational Notices Leading up to Stress Events

[11]EMR Delivery Body 10th August Coordination Event - see circa slide #122 onwards
https://www.emrdeliverybody.com/_layouts/15/listform.aspx?PageType=4&ListId=%7B612853B7%2D6D9A%2D4B23%2D9686%2D6AB632015C31%7D&ID=127&ContentTypeID=0x010400626754A76E41C74FA81B4D17EBF15511

For a duration limited resource, it's triggering of the commencement of production should ideally coincide exactly with the beginning of a CM stress event if GB system reliability is to be maximised, and if such a resource is to avoid CM non-performance penalties.

Note that there may be a distinction between the likely behaviour of Balancing Market (BM) and non-BM-participant storage in this context of a system stress event. BM participant storage will be directly exposed to price escalations that would occur during a pending adequacy shortage and may be able to coordinate their power production to a shortage event accordingly. Non-BM storage may rely on other sources of information such as the issuance of warning notices from the system operator control centre. In the real system, then it is a commercial decision for each actor when to generate if they are to meet their CM obligations accordingly.

As there is no distinction between BM and non-BM storage in the CM technology class definitions, the modelling approach in this de-rating factor assessment work treats them consistently. However, as the experience with short-duration storage in the GB system grows in the future, the impact of any diversity of CM stress event response on actual GB risk levels will be important to monitor so that we have an accurate characterisation of the true reliability level as much as rewarding all CM parties fairly.

All resources can perfectly forecast CM stress events in the modelling environment

As explained above, the exact onset of a CM stress event is defined by ex-post analysis and no market actor can be sure exactly when a CM stress event begins in real-time.

In the real system then self-dispatch decisions on the day of a tight GB operational margin may be a challenging task, as it is the commercial remit of each duration-limited storage device when they decide to start and end power production. This of course is an issue for all CM resources if they decide to start production too late after a stress event begins, but may be additionally challenging for duration limited resources if they decide to start production too early and thus run out of energy by the time the event is later deemed to have started.

However, given that the purpose of this modelling assessment is to derive a reliability value for a duration-limited resource behaving in a manner consistent with its CM requirement, then a sensible starting premise in the modelling environment must be that the resource actually attempts to fulfil this requirement. Whether any resource behaves in a manner consistent with its CM requirement in the real system is an issue that would be covered by performance compliance assessment later on.

Therefore, a key assumption in the modelling works herein is that all system resources can perfectly forecast the starting point and duration of a CM stress event.

Storage is always charged to the level of its CM contract duration requirement at the onset of a stress event in the modelling environment

A further related modelling task in this de-rating factor assessment is what to assume as the starting charge level of the energy store of a duration-limited resource at the onset of a CM stress event.

On the real system, it is noted that storage may be performing multiple services across a number of different commercial contracts at any given time. For example, some of these services may be related to energy arbitrage in the BM, ancillary services provided to the SO, and/or triad avoidance with a supplier company for network charges reduction. As provision of these services earlier in a day may impact the storage charge state in the lead up to a later CM stress event, then in the real system they may have an impact on the ability of a duration limited resource to fulfil the CM obligation unless they are carefully managed.

The exact self-dispatch and storage change state management strategy is a commercial decision for each individual party on any given day of a tight GB adequacy margin. It is important to note though that there are performance penalties for any CM resource that does not fulfil its requirements during an actual stress event.

Analogous to the assumption above regarding perfect forecasting of stress events, then given that the purpose of this modelling assessment is to derive a reliability value for a duration-limited resource behaving in a manner consistent with its CM requirement, then again, a sensible starting premise in the modelling environment must be that the resource actually fulfils this requirement. Whether any resource actually behaves in a manner consistent with its CM requirement in the real system is an issue that would be covered by the ex-post performance compliance assessment later on.

Therefore, another key assumption in the modelling work herein in that storage is always charged to the level of its CM duration requirement at the onset of a stress event.

Storage response is immediate at the onset of a CM stress event

The response shape and co-ordination level of the fleet of storage resources at the onset of a GB adequacy stress event is another key modelling attribute of the de-rating factor assessment. As indicated by the stylised diagram below, there are a number of possible responses that duration-limited resources can provide, each of which will have a different impact on the depth and duration of the residual stress event capacity shortage (the simple case example here is a number of short-duration batteries but in principle the concepts apply to all short-duration storage resources of any type).

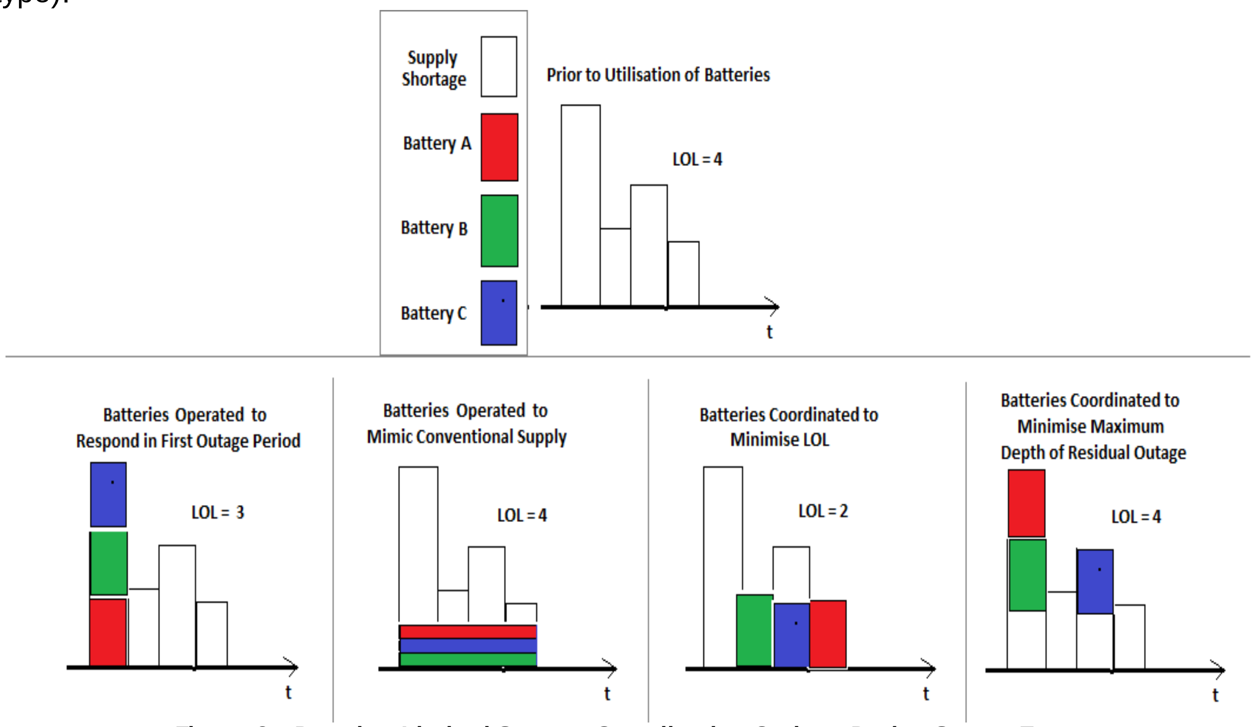

Figure 3 – Duration Limited Storage Coordination Options During Stress Events

These examples lend to a number of possible ways to model duration limited storage in the reliability assessment studies later in the report. The GB reliability simulation models were therefore designed to build in the 4 storage “Coordination Algorithms” proposed here:

- **Algorithm 1** – storage responds immediately at start of the event to fulfill as much as possible of the system shortage needs until it's energy store is exhausted
- **Algorithm 2** – storage tries to mimic conventional plant with a flat response during the event, with the exact power generation level in each event chosen so as to provide as much as possible of the stored energy to the system shortage
- **Algorithm 3** – storage coordinated so as to minimize number of residual loss of load hours
- **Algorithm 4** – storage coordinated so as to minimize the maximum depth of the outage in the worst period of each event

Note that these 4 Algorithms assume the least possible wastage of any duration-limited storage on the system. For example it is assumed that (i) the duration-limited storage is the last resource dispatched on the system, (ii) it's response is 'shaped' in the first periods of the stress event to reflect the residual load shortage shape (i.e. no overshoot of overall fleet storage response to a shallow shortage), and finally (iii) as much of possible of the stored energy is output as power during the stress event so that as little as possible stored energy remains after the stress event finishes.

In reality, it is difficult to know which strategy each individual storage operator will carry out during a real system stress event on the GB system – the exact shape and timing of their dispatch in once again their individual commercial decision. It was also noteworthy from our initial methodology consultation though that a majority of respondents felt that the immediate response in Algorithm 1 would be the most likely as stress events are anticipated to be more likely to be shorter in duration than extended. As discussed in the following sub-section though, and revealed in the results Section 5, then the materiality of this assumption is actually negligible provided we assume that EEU is a sensible risk metric upon which to base the EFC.

EEU is a more appropriate statistical risk metric to capture a storage EFC

As described above on Page 9, then the average and incremental EFCs of the duration-limited storage resources are defined with respect to the same reliability/risk-reduction contribution as an equivalent capacity of firm and infinite duration plant. An important question is then which statistical risk metric to use when carrying out the EFC assessment? There are two obvious candidates for the GB system CM – either the LOLE or the EEU.

One important observation from the 4 coordination Algorithms in the diagram above is that the number of loss of load (LOL) periods in the residual stress event shortage could be heavily impacted by the assumption of storage response strategy – accumulated over time therefore, the LOLE would be highest in Algorithm 4 and lowest in Algorithm 3. However, the overall energy unserved (EU) in each Algorithm is identical.

This indicates a possible weakness of the LOLE reliability metric <u>for this particular application</u> in that it is quite sensitive to the storage operational strategy assumed. The EEU however, is a direct representation of the customer economic damage on the system by means of the Value of Lost Load (VoLL) parameter (overall annual customer damage impact = EEU * VoLL). Therefore EEU as a risk metric may be both more independent of the storage response strategy and also more reflective of economic damage costs of the residual shortage.

For this reason, and in this particular application which aims to translate the duration-limited storage contribution to security of supply on to the same basis as the other CM participants, then the EEU risk metric will be used to define their EFC and thus de-rating factor. As discussed in the academic expert commentary in Appendix 3, the GB reliability standard of 3 hours per year LOLE is then once again able to be used for the overall CM target capacity to secure once all the supply resources are considered on a consistent basis.

Technical availability of storage can be applied as a linear scalar to the simulation based EFC in order to derive the overall derating factor

The above EFC assessment relates to characterising the impact of the duration-limited aspect of storage resources on security of supply. However, there is a further important aspect related to the technical availability of the storage resources that also needs to be accounted for – i.e. to reflect the fact that they can randomly suffer mechanical breakdown in the same manner as any other capacity resource in the system.

The assessment for the incremental EFC of the storage duration-category is carried out assuming the marginal storage unit added to the system is perfectly reliable. This perfectly-reliable duration-category EFC must then the linearly scaled by the technical availability of storage to produce the end result de-rating factor for that storage duration sub-class in the CM.

The average EFC of the entire storage fleet must be assessed with the possibility of random technical unavailability of the storage fleet components already built in to the reliability simulation, as these random failures could impact the shape and duration of overall system stress events themselves. The average EFC numbers presented in later results Section 5 thus need no further linear scaling in this manner.

An important question then relates to what should be the value of the linear technical availability scalar applied to the incremental EFC to get the end result de-rating factor. It is known that some modern battery storage projects are highly modular in design, implying that their overall plant reliability may be high. As the CM does not have any further distinction of the storage technology class by constituent technology (e.g. there is no distinction between battery storage, pumped storage, compressed or liquid air storage etc) then the same technical-availability must apply to all. As pumped storage is the only storage technology with a reasonable amount of historical technical availability performance data, and that furthermore, it is expected to be the mainstay of storage capacity in the CM for the short to medium term at least, then it was decided to use its technical availability as the scalar to apply in this case.

As indicated in this year's Electricity Capacity Report, the rolling 7-year average availability of pumper hydro at time of system peak demand is 96.11%. This parameter is used to scale the EFC values presented in the results Section 5 later on to arrive at the final de-rating factors in Section 7.

Storage recharging prior to/after electricity generation is generally an off-peak activity

Storage used as power output generation during a stress event necessarily requires a corresponding re-charge cycle beforehand and afterwards in order to fill up again for readiness of the next cycle of usage. An important question then relates to the timing and extent of this storage recharging, and its relationship with the overall GB system adequacy state on days of tight margin.

A basic assumption of this modelling framework is that the storage units are sufficiently charged at the start of each stress event, and that any recharging is carried out in an off-peak manner at times of much lower residual GB system demand. Furthermore, given that the recharging is carried out off-peak, then then impact of any round-trip efficiency of the storage will have no effect on the EFC parameter derived as a result.

Some respondents to our initial methodology consultation raised the risk that if there is a lot of storage capacity/volume on the system, that this recharging cycle could create new stress events themselves if it is not carried out sufficiently well in advance of the adequacy shortage, noting that there is only a 4-hour window between the CM notice warning and a potential stress event. While this is a possibility, it is an implicit assumption in this methodology that storage can forecast the onset of the stress events and will plan accordingly with sensible commercial decisions. As the level of storage on the GB system at present is relatively small compared to the overall CM capacity requirement, then this assumption may be allowable for now, but may be worthy of future investigation in later years as more operational data becomes available.

All storage resources have uniform CM de-rating factor regardless of ancillary service provision

It is noted that one of the important revenue sources for short-duration storage on the future GB system is the possibility to provide ancillary services. The provision of such services may impact the storage charge state both leading up to and during a CM adequacy stress event. For example storage resources providing Enhanced Frequency Response (EFR) will typically be half-charged so that they can respond symmetrically to over or under frequency disturbances accordingly. It is impossible to know what the system frequency level will be during a CM stress event and thus to forecast the power output pattern of such resources during an adequacy shortage.

However, the majority of ancillary services are included in the "List of Relevant Balancing Services"[12] such that any CM contracted quantity of capacity simultaneously holding such an ancillary service requirement is exempt from penalty payments during the CM stress event. The motivation for this is that important ancillary services such as frequency control provision must be maintained even during a CM stress event in order to safeguard the stability of the system. The overall requirement of capacity in the CM is thus reflective of this – with the annual Electricity Capacity Report adding the basic contingency reserve requirement to demand level when calculating the capacity to secure. Thus any CM capacity that is held on standby providing an ancillary service is still contributing to the overall system total capacity requirement to the same extent as that which is providing energy – for if this resource is not providing the relevant ancillary service then another capacity source on the system would have to be used instead.

A further issue is that the timeframe of CM auction contracts (which may be up to 15 years) and system ancillary service contracts (which may be as short as 1 month in future) are different, and therefore to factor any ancillary service contract attributes in to a CM de-rating factor assessment for any resource would be very difficult to forecast.

For both of these reasons, all storage resources have uniform CM de-rating factor in this work regardless of ancillary service provision.

[12] CM Rules, Schedule 4 – List of Relevant Balancing Services:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/629953/capacity-market-amendment-rules-2017.pdf

3. Time Sequential Storage Modelling and Software Tools

The standard reliability assessment methodology used to date on the GB system adequacy problem has been time collapsed convolution of the various conventional plant, wind power and demand probability distributions in order to assess LOLE and EEU. Time collapsed modelling breaks the inherent sequential nature of system characteristics from one settlement period to the next. This has been more than sufficient in the GB reliability monitoring and CM capacity to secure assessment tasks to date as the main GB reliability standard (LOLE) is a fundamentally time collapsed metric (i.e. there is no specific account of how loss of load events are grouped in time), and furthermore the penetration of significantly duration-limited resources (whose behaviour is dependent on successive past and future periods) on the system has been small.

For this storage EFC assessment task however, where the sequential charge and discharge of a storage device with respect to an aggregate GB system capacity shortage over consecutive time periods is absolutely necessary to represent, then an alternative modelling framework to time-collapsed is required. To that end, we have used the LCP Unserved Energy Model (UEM) which is a time sequential Monte Carlo simulation model of GB adequacy. This tool is a related sub-module to the time-collapsed Dynamic Dispatch Module (DDM) software used in the annual Electricity Capacity Report and thus preserves general consistency of GB system data and plant representation.

The UEM model uses 11 years of time-coincident hindcast historical wind and demand data, as well as a two-state (fully-available/fully-unavailable) representation of conventional plant technical availability and its mean-time-to repair (MTTR). Conventional plant availability is random Monte Carlo simulated on the basis of assuming a geometric distribution of plant failure outage duration, the mean of this distribution being the MTTR that is calculated from historical availability data. We have also included a time series representation of historical solar PV availability based on its growing importance to the GB system. More details of the plant MTTR data and the solar PV representation are outlined in the Appendix 1 and 2. Note that transmission system or network outages are not included in this modelling (only direct generation shortages) as is consistent with the overall focus of the CM.

When supplied with a relevant set of Base Case input data assumptions for projected demand and supply capacities, the UEM statistical simulation cycles through enough instances of possible annual GB system adequacy states at half-hourly resolution, based on simulated random availability of conventional plant, until such time as there is stochastic convergence of the relevant statistical reliability indices of interest. The UEM also has functionality to select a number of "capacity shifts" around the Base Case (i.e. the Base Case plus or minus a certain MW of firm capacity) so that the average EFC of the overall storage fleet and the incremental EFC of a marginal storage unit can be estimated via interpolation.

For this project, LCP have also included in the UEM a representation of the storage Coordination Algorithms 1-4 as described above in the previous Section. The modelling of storage in the UEM is also in adherence to all of the methodology principles outlined in Section 2 above. The full details of the UEM time sequential modelling framework, and the changes applied this year to specifically represent storage within it, can be found at the following resources [13],[14].

[13] LCP UEM Overview: https://insight.lcp.uk.com/acton/attachment/20628/f-060f/1/-/-/-/-/LCP%20Unserved%20Energy%20Model.pdf

[14] Storage Specific Changes to the UEM https://insight.lcp.uk.com/acton/attachment/20628/f-0610/1/-/-/-/-/LCP%20UEM%20storage%20proposed%20methodology.pdf

4. Base Case and Sensitivity Analysis Assumptions

The main focus of this project was to suggest incremental-EFC based de-rating factors for duration limited storage in the 2018/19 T-1 and the 2021/22 T-4 CM auction delivery years. However, a significant amount of sensitivity analysis was also prudent for some of the key modelling parameters of interest in order to gain a better understanding of their influence on the final EFC results.

The following studies and sensitivity analyses were carried out for storage EFC assessment using the 2017 EMR 5-year Base Case (see the 2017 Electricity Capacity Report for full details and general assumptions of this Base Case), with both the average EFC of the overall storage fleet and the incremental EFC of all storage durations ranging from 0.5-hour to 24-hours calculated. All cases below (unless otherwise stated) both include existing GB pumped hydro plus projected levels of battery storage plus a small amount of liquid air storage, and furthermore all were adjusted to 3 hours LOLE the GB reliability standard.

- **Case #1** - 2018/19 Base Case with "best view" of battery storage penetration (~ 400MW)
- **Case #2a** - 2021/22 Base Case with ~ 2000 MW of battery storage penetration
- **Case #2b** - 2021/22 Base Case with "best view" of ~ 1000 MW of battery storage
- **Case #2c** - 2021/22 Base Case with only committed or existing battery storage projects (~ 640 MW)
- **Case #3 –** Case #2a using LOLE as the EFC risk metric instead of EEU
- **Case #4 –** The 2021/22 Base Case with no pumped storage, batteries or other storage at all in the model
- **Cases #5, #6, #7, #8, #12** - The 2021/22 Base Case with storage capacities ranging from 1 to 5GW of 0.5 hr – 6 hr duration **only** (i.e. all other pumped hydro and battery storage removed)
- **Case #9** - Case #2a adjusted to 0.5, 1, 2, 3, 4, 5 hours LOLE as the incremental EFC assessment starting point
- **Case #10** - Case #2a with MTTR values of system supply resources either doubled or halved
- **Case #11** - Case #2a with the size of the incremental EFC unit ranging from 50 MW to 100 MW, 200 MW and 500MW

Note that as requested by market stakeholders during our industry consultation workshops, the EMR 5-year Base Case as used in this work has been updated with respect to the version which was applied earlier this year in the 2017 Electricity Capacity Report. More accurate information came to light regarding the level of existing storage on the distribution network at present, which was lower than initially thought. This combined with the ongoing policy reviews of embedded benefits and capacity market de-rating factors for duration limited storage, has led to a downwards revision in the EMR 5-year Base Case for the projected level of storage capacity expected to be on the system in the T-1 and T-4 CM delivery years.

The exact level of new-build storage capacity on the system as a result of these changing market and policy drivers is of course difficult to predict, but a single 'best view' of such capacity for the T-1 and T-4 year de-rating factor assessments is necessary for the purposes of this study, and is as projected by Cases **#1** and **#2b** above which are presented in detail in the following Table 1. Note that this updated 'best view' corresponds to an assumption of ~ 400MW of battery storage in total on the system for the 2018/19 T-1 assessment, as well as circa 1,000MW of battery storage in total on the system for the 2021/22 T-4 assessment.

As a result of these embedded storage revisions, the Average Cold Spell (ACS) peak demand forecasts have also been correspondingly updated since the 2017 ECR was published. The 2018/19 underlying demand ACS peak value is forecast to be 61.3 GW and the 2021/22 underlying demand ACS peak value is forecast to be 61.0 GW. We also then add the capacity requirement to cover the basic contingency reserve amounts of 900 MW and 1000 MW in the respective target years, so the total GB system underlying ACS peak demand level was thus assumed to be 62.2 GW and 62.0GW in the T-1 2018/19 and T-4 2021/22 delivery years respectively.

Note that existing GB pumped hydro capacities and durations are well known and are expected to be available to the system for the foreseeable future. However the exact assumption for existing and future battery storage capacity and duration capability are modelling assumptions here. The assumption for existing contracted battery storage capacity to either the CM or the EFR ancillary services markets is assumed to be 0.5 hours duration. Additional uncontracted battery storage is assumed to be 2 hours duration which is the standard assumption in the EMR Base Case. Pumped hydro and liquid air storage durations were estimated using various public domain data sources.

For clarity, the Table 1 below indicates the total amount of storage by duration category assumed in each of the Cases #1, #2a, #2b and #2c – the longer durations (>= 6 hours) correspond to pumped hydro capacity and the shorter ones (<= 2 hours) battery storage capacity. The small amount of liquid air storage capacity in the Base Case corresponds to the 3 hour duration category.

Table 1 – Storage Capacity by Duration Assumed in the Base Cases and Sensitivities

	2018/19 Case 1 (MW) – Best View	**2021/22 Case 2a (MW)**	**2021/22 Case 2b (MW) – Best View**	**2021/22 Case 2c (MW)**
Duration Category 22 hours	440	440	440	440
Duration Category 21 hours	300	300	300	300
Duration Category 6 hours	2004	2004	2004	2004
Duration Category 3 hours	5	5	5	5
Duration Category 2 hours	148	1327	419	60
Duration Category 0.5 hours	251	581	581	581
Total storage	3148	4657	3749	3390

5. Equivalent Firm Capacity Results

EFC Results Case #1 – 2018/19 Base Case with existing pumped hydro and "best view" of battery storage penetration (~ 400 MW)

The Table 2 below introduces the EFC results for the 2018/19 Case #1 study. Both the average EFC of the entire storage fleet and the incremental EFC of the different storage duration categories are presented for each of the 4 storage coordination Algorithms applied during the simulated stress events.

Table 2 – EFC Results for Storage Duration Classes for Case #1

	Algorithm 1	Algorithm 2	Algorithm 3	Algorithm 4
Whole Fleet Average EFC	87.3%	87.3%	87.3%	87.3%

Incremental EFCs	Algorithm 1	Algorithm 2	Algorithm 3	Algorithm 4
Duration: 0.5 hour	22.2%	22.2%	22.2%	22.2%
Duration: 1 hour	42.0%	41.9%	42.1%	42.0%
Duration: 1.5 hour	58.2%	58.0%	58.2%	58.2%
Duration: 2 hour	70.8%	70.6%	70.8%	70.8%
Duration: 2.5 hour	80.4%	80.2%	80.4%	80.4%
Duration: 3 hour	86.0%	85.8%	86.0%	86.0%
Duration: 3.5 hour	89.2%	89.1%	89.2%	89.2%
Duration: 4 hour	90.7%	90.5%	90.7%	90.7%
Duration: 4.5 hour	91.4%	91.3%	91.4%	91.4%
Duration: 5 hour	92.0%	91.8%	92.0%	92.0%
Duration: 5.5 hour	92.6%	92.2%	92.6%	92.6%
Duration: 6 hour	93.2%	92.6%	93.2%	93.2%
Duration: 6.5 hour	93.7%	93.3%	93.7%	93.7%
Duration: 7 hour	94.3%	93.9%	94.3%	94.3%
Duration: 7.5 hour	94.8%	94.5%	94.8%	94.8%
Duration: 8 hour	95.4%	95.1%	95.4%	95.4%

A number of observations are apparent from the above table of 2018/19 EFCs:

- It can be seen that the average EFC of the entire storage fleet at 87.3% is quite high given that the storage fleet in this case is dominated by the existing GB pumped hydro capacity which is of relatively long duration
- The incremental EFCs are seen to increase in a nonlinear fashion from the 0.5 hour duration category to the 8 hour category, with additional duration providing lesser additional reliability benefit beyond about 3 hours duration
- As expected, the EFC reaches a level consistent with the de-rating factor presently attributed to storage in the CM for the longer duration categories (durations 8.5h – 24h not indicated in this table for space reasons)
- The incremental and average EFCs across the 4 storage coordination Algorithms are almost identical for each duration. This underlines the value of using EEU as a stable statistical risk metric upon which to base the EFC study, one that is furthermore linked to the economic damage costs in each system shortfall event
- Note that the average EFC contains the effect of technical breakdown within the storage fleet, while the incremental EFCs are given for a perfectly reliable 100MW unit
- Note also that the average and the incremental EFCs are reported with respect to a Base Case adjusted to an LOLE of 3 hours, the GB reliability standard and expected CM reliability provision level

EFC Results Cases #2a, #2b, #2c and #4 – 2021/22 Base Case with existing pumped hydro and sensitivity case variations of battery storage penetration

Table 3 – EFC Results for Storage Duration Classes for Cases #2a, #2b, #2c, #4

	Case #2a(2GW)	Case #2b(1GW)	Case #2c(0.6GW)	Case #4(0GW)
Whole Fleet Average EFC	76.1%	80.3%	81.6%	0%
Incremental EFCs	**Algorithm 1**	**Algorithm 1**	**Algorithm 1**	**Algorithm 1**
Duration: 0.5 hour	14.9%	18.6%	21.1%	26.4%
Duration: 1 hour	29.9%	37.9%	41.7%	48.6%
Duration: 1.5 hour	43.2%	54.4%	59.4%	66.2%
Duration: 2 hour	54.6%	67.4%	72.2%	79.6%
Duration: 2.5 hour	65.8%	78.5%	81.4%	88.9%
Duration: 3 hour	74.5%	85.4%	86.9%	94.3%
Duration: 3.5 hour	79.6%	89.2%	90.1%	97.2%
Duration: 4 hour	82.3%	90.8%	91.3%	98.2%
Duration: 4.5 hour	83.8%	91.8%	92.0%	98.7%
Duration: 5 hour	84.9%	92.4%	92.5%	98.9%
Duration: 5.5 hour	85.9%	92.9%	92.9%	99.0%
Duration: 6 hour	86.8%	93.4%	93.3%	99.1%
Duration: 6.5 hour	87.7%	93.8%	93.7%	99.2%
Duration: 7 hour	88.6%	94.3%	94.1%	99.3%
Duration: 7.5 hour	89.5%	94.8%	94.5%	99.3%
Duration: 8 hour	90.3%	95.3%	94.9%	99.4%

The above Table 3 indicates the EFCs (for the storage coordination Algorithm 1 applied in each column) in the 2021/22 case ranging from 2GW (case 2a) to 1GW (case 2b), 640MW (case 2c) and 0 MW (case 4) of battery storage penetration.

It can be seen in the columns from left to right in each of the storage duration categories above that the installed penetration level of duration limited storage in the Base Case can have material impact on the resultant incremental and average EFCs. As indicated previously, this is because different levels of duration-limited storage will influence the shape and duration of stress events on the system (replacing the same capacity of firm plant with duration-limited storage will tend to cause stress events to be marginally longer), implying that the value of additional penetration of that same duration limited resource thereafter will tend to saturate.

EFC Results Cases #5, #6, #7, #8 and #12 - 2021/22 Base Case with ranging from 1 to 5GW of 0.5 - 6 hour duration storage only (i.e. all other pumped hydro and battery storage removed)

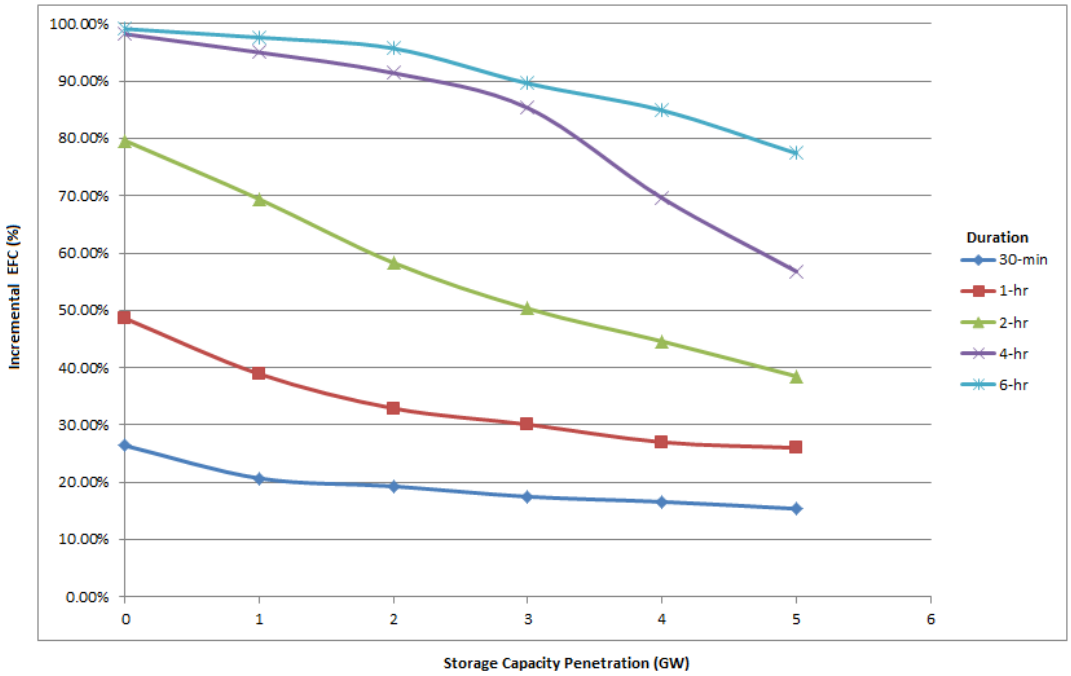

Figure 4 – Indicative Evolution of Storage EFCs as Penetration Level Increases

The Figure 4 above would indicate that the trends established in the EFC Tables 2 and 3 above persist with additional penetration of each duration limited resource – i.e. more capacity of a given duration will tend to have lower incremental security of supply contribution and thus a lower de-rating factor in the CM. Note that the starting point in this graph is a hypothetical case of no storage capacity at all on the system of any sort, and note also that this graph depicts the incremental EFC only for a given penetration of that duration.

Note that there are 30 different data points on this graph, that correspond to 6 penetration levels (in ranges of 0 – 5 GWs of capacity), of 5 different storage duration bands. There are necessarily 30 different assessment studies that underlie this graph – and thus a different/self-contained set of 48 (0.5-24 hour) duration storage EFCs (i.e. akin to the entire Table 2 or 3) for each Base Case that underpins each single point on this graph. We have only presented in this graph <u>one incremental EFC for each Base Case</u> that underlies each of the 30 different studies – but for each point presented here there are 47 other EFC values that are directly related. Comparisons of the EFCs across the 30 different points presented in this graph thus need to be carefully understood – one can only directly compare the 48 EFC values that correspond to within the same Base Case.

For example the bottom right data point on the graph shows the incremental EFC for 0.5 hours duration storage (<u>assuming there is already 5GW of 0.5 hours storage on the system</u>) as ~ 15%. However, the incremental EFC for 6 hours duration in this specific base case is ~ actually 98% (i.e. not visible on the graph), and not the ~ 77% data point in the top right hand corner data point on the graph. The ~77% figure for the top right hand corner data point in this graph corresponds to the incremental EFC of 6 hour duration storage, <u>if you already have 5GW of 6hour storage in the system</u> – and not if you have 5GW of 0.5 hour duration.

This indicates the flexibility of the EFC metric to respond to the evolution or ‘path dependency’ in the build out of storage in the GB system. What we can see therefore is that the EFC metric for the full range of storage durations remains sensitive to the changing needs of the system. For example in the blue line in the graph above, we can see that the EFC of the 30 minute storage penetration degrades as its own level of penetration grows. However, the relative value of long-duration storage, when there is the same expansion in the level of short-duration of storage, stays resilient at almost 100% EFC. Thus, all other things being equal, the CM derating factor would in this case tend to progressively value less any short duration storage if that is what substantially builds out first, while preserving the incentive for longer duration storage in the later years. It is an important point that underlines the value of an EFC metric to remain sensitive to the relative needs of the system.

It is difficult to project exactly how the de-ratings for storage will change in future years with additional capacity build out, as the likely system in future will have some mixture of capacities at different constituent durations – this is just an indicative rough sensitivity anlaysis. The graph might suggest though that the more duration limited storage on the system in future, then de-rating factors may change. Accordingly, then as per Table 1 we have used our 'best view' of the storage penetration in the 2018/19 and 2021/22 Base Cases when deriving the final T-1 and T-4 de-rating factors respectively.

EFC Results Case #3 2021/22 Case 2a using LOLE as the EFC risk metric instead of EEU

Table 4 – EFC Results for Storage Duration Classes for Case #3

	Algorithm 1	Algorithm 2	Algorithm 3	Algorithm 4
Whole Fleet Average EFC	82.1%	81.9%	93.5%	72.2%

Incremental EFCs	Algorithm 1	Algorithm 2	Algorithm 3	Algorithm 4
Duration: 0.5 hour	37.6%	25.4%	60.4%	15.0%
Duration: 1 hour	47.0%	42.7%	88.1%	22.9%
Duration: 1.5 hour	59.3%	56.5%	93.0%	36.5%
Duration: 2 hour	70.8%	66.1%	94.7%	47.2%
Duration: 2.5 hour	78.1%	78.9%	96.1%	58.4%
Duration: 3 hour	85.0%	84.1%	97.0%	67.9%
Duration: 3.5 hour	90.9%	86.6%	97.6%	74.0%
Duration: 4 hour	94.4%	88.0%	98.8%	77.6%
Duration: 4.5 hour	95.5%	89.4%	98.8%	78.9%
Duration: 5 hour	96.2%	90.0%	98.8%	79.2%
Duration: 5.5 hour	96.4%	90.9%	98.8%	79.2%
Duration: 6 hour	96.6%	91.7%	99.1%	79.7%
Duration: 6.5 hour	96.9%	92.6%	99.4%	80.4%
Duration: 7 hour	96.9%	93.3%	99.8%	80.9%
Duration: 7.5 hour	96.9%	93.8%	99.8%	81.2%
Duration: 8 hour	97.2%	94.6%	100.1%	82.9%

The Table 4 above indicates the sensitivity of the EFCs to the assumptions for storage response during the CM stress event when the risk metric used for the EFC assessment is LOLE.

- Contrary to the situation where EEU is used as the risk metric, this table indicates quite different EFCs for the 4 different storage coordination Algorithms (remember Figure 3 earlier) under an LOLE based EFC assessment
- The EFCs are most sensitive in the Algorithm 3 case, as this is the one in which LOL is being targeted to be minimised in each shortfall event
- It also shows that the EFC of the storage resources might be higher if LOLE were used – however this could be misleading as there is no specific link to the economic damage costs of the supply shortfalls for this metric in this application
- This table thus further underlines the unsuitability of LOLE as a risk metric for the storage EFC assessment as discussed in the methodology Section 2 above, and justifies the EMR modelling methodology consultation position that EEU is a far more stable and suitable metric for this particular purpose

EFC Results Case #9 - Case #2a adjusted to 0.5, 1, 2, 3, 4, 5 hours LOLE as the starting point for the EFC assessment

Table 5 – EFC Results for Storage Duration Classes for Case #9

	LOLE 0.5h	LOLE 1h	LOLE 2h	LOLE 3h	LOLE 4h	LOLE 5h
EFCs	**Algorithm 1**	**Algorithm 1**	**Algorithm 1**	**Algorithm 1**	**Algorithm 1**	**Algorithm 1**
Whole Fleet Average	78.5%	77.9%	77.0%	76.1%	75.2%	74.3%

Incremental EFCs	Algorithm 1	Algorithm 1	Algorithm 1	Algorithm 1	Algorithm 1	Algorithm 1
Storage Duration: 0.5	17.5%	16.7%	16.1%	15.5%	14.8%	14.3%
Storage Duration: 1	34.1%	33.5%	31.5%	30.3%	29.0%	28.0%
Storage Duration: 1.5	48.1%	48.4%	45.2%	43.3%	41.7%	40.2%
Storage Duration: 2	60.5%	61.7%	57.4%	54.9%	53.5%	51.5%
Storage Duration: 2.5	72.7%	74.3%	69.4%	66.2%	65.3%	63.0%
Storage Duration: 3	81.3%	83.1%	78.6%	74.7%	74.4%	72.3%
Storage Duration: 3.5	85.7%	88.0%	84.2%	79.8%	79.9%	78.1%
Storage Duration: 4	87.4%	90.0%	86.7%	82.6%	82.9%	81.3%

The above table indicates that the EFC values are only partially sensitive to the LOLE operating point of the underlying Base Case. Note that in these cases EEU is still used as the EFC risk metric, just that the marginal EFC unit addition is simply starting at an alternative LOLE set point - as the LOLE values decrease, this indicates that the GB system is becoming more reliable (i.e. higher de-rated margin).

It can be seen that as the LOLE reduces, then the incremental EFCs increase slightly (interestingly, it is generally the opposite effect observed for wind power EFCs based on experience to date – wind power EFC is higher when the LOLE is higher and the margin lower). This is because for the lower LOLE levels, then the system stress events tend to be slightly shorter on average, implying that any limitations of short duration storage are less likely to be revealed and thus the corresponding EFCs can be slightly higher.

Note that an LOLE of 3 hours/year is the target reliability level of the Capacity Market in each delivery year and thus this is the consistent basis upon which to select the EFCs for use as de-rating factors. The table above implies that the impact of a lower LOLE level Base Case is relatively limited in any case.

EFC Results Case #10 - Case #2a with Mean Time to Repair (MTTR) values modified

Table 6 – EFC Results for Storage Duration Classes for Case #10

	MTTR halved	MTTR Base	MTTR doubled
EFCs	Algorithm 1	Algorithm 1	Algorithm 1
Whole Fleet Average	76.2%	76.1%	76.0%
Incremental EFCs	Algorithm 1	Algorithm 1	Algorithm 1
Storage Duration: 0.5	16.2%	14.9%	15.5%
Storage Duration: 1	31.6%	29.9%	30.4%
Storage Duration: 1.5	45.2%	43.2%	43.7%
Storage Duration: 2	57.1%	54.6%	55.4%
Storage Duration: 2.5	68.6%	65.8%	66.7%
Storage Duration: 3	77.2%	74.5%	75.4%
Storage Duration: 3.5	82.4%	79.6%	80.6%
Storage Duration: 4	84.9%	82.3%	83.2%

As part of this study, we have updated the MTTR values used in the UEM simulation based on the most recent system operational data. Details of our assessment are contained in Appendix 2.

The above table would suggest that the incremental EFC is not very sensitive to the values MTTR applied however. The EFC values derived from the MTTR used in the Base Case (middle column above) were not significantly changed when the MTTR was either doubled or halved respectively in Case #2a.This is likely due to the fact that the average outage length of a conventional plant (expected to be a few days) is an order of magnitude longer than the typical CM stress event duration (expected to be a few hours).

EFC Results Case #11 - Case #2a with the MW capacity size of the incremental storage unit modified

Table 7 – EFC Results for Storage Duration Classes for Case #11

	50MW unit	100MW unit	200MW unit	500MW Unit
EFCs	Algorithm 1	Algorithm 1	Algorithm 1	Algorithm 1
Whole Fleet Average	76.1%	76.1%	76.1%	76.1%
Incremental EFCs	Algorithm 1	Algorithm 1	Algorithm 1	Algorithm 1
Storage Duration: 0.5	15.6%	14.9%	15.2%	14.7%
Storage Duration: 1	30.8%	29.9%	29.6%	28.7%
Storage Duration: 1.5	44.0%	43.2%	42.4%	41.7%
Storage Duration: 2	55.7%	54.6%	54.7%	53.8%
Storage Duration: 2.5	67.2%	65.8%	66.4%	65.4%
Storage Duration: 3	75.9%	74.5%	75.3%	74.5%
Storage Duration: 3.5	81.2%	79.6%	80.6%	80.3%
Storage Duration: 4	83.8%	82.3%	83.3%	83.3%

As outlined in Section 2, the capacity size of the incremental storage unit added to the Base Case in order to derive the incremental EFC results for the different storage durations was fixed at 100MW in the Base Case assessments. The above table compares the results for the incremental

EFC values in Case #2a when the size of the incremental storage unit added to the case was varied from 50 MW to 500 MW.

The results would indicate that the EFC parameters are reasonably robust to this assumption regardless, with a slight increase in the EFC for the smaller sized units of short duration.

However, 100 MW as was used in the Base Case is probably a good trade-off between the sizes of various units likely to be successful in the CM auctions, noting the fact that under the proposed rule changes by BEIS in their policy consultation, there is not planned to be any distinction in the de-rating factors of different capacity sized storage plants, only their duration. 100 MW is also probably a good trade-off in order for the incremental EFC values to be numerically robust to various rounding and interpolation operations used in the UEM simulation.

6. Histograms of Stress Event Durations

This chapter outlines the histograms of stress event durations for a range of case studies that are adjusted to 3 hours/year LOLE, the GB reliability standard and the CM target reliability level. Note that the histograms for storage coordination Algorithms 1 and 2 are mainly presented here for clarity and realism purposes, as Algorithms 3 and 4 are but hypothetical studies for now (there is no specific means in the GB CM to coordinate the response of the storage fleet to target either LOL or depth of stress event minimisation).The histogram diagrams below also specify the average stress event duration and also the expected number of events per year, when the system is at this reliability level.

Histogram of Stress Event Durations – 2021/22 Case #4 No Storage

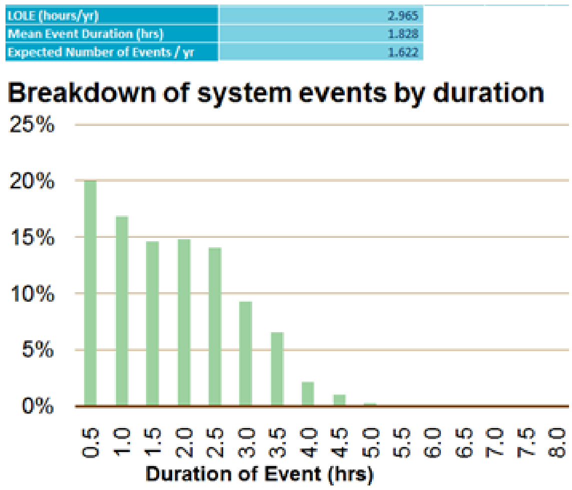

Figure 5 – Histogram of Stress Event Durations 2021/22 - No Storage

The above Figure 5 indicates the histogram of stress event durations in the case where there is no storage at all in the 2021/22 Base Case, furthermore with that case at an approximate LOLE of 3 hours/year. It is worth noting firstly that the overall probability of there being an event in each year is still relatively low – with ~ 1.6 events per year on average, and each event having an average duration of ~ 1.8 hours. It can be observed that the short duration outages are more probable than long ones, with a very low probability of events lasting more than 4.5 hours in duration.

Histograms of Stress Event Durations – 2021/22 Case #2b (1GW)

The histograms in Figure 6 overleaf indicate that the addition of duration limited storage to the same case has the impact of, on average, lengthening the durations of any events that occur (note that the number of events/year is necessarily thus lower as all Base Cases are readjusted to LOLE of 3 hours/year). One can also see that there is a slightly greater probability of longer duration events in the far right hand side tail of the distribution, though it must be cautioned that model accuracy for extended duration events is a challenge as these events are so rare in occurrence and this model is based on the best available 11 years of historical system demand, wind and solar data. As discussed in Section 8, then in future once we have a longer and more comprehensive historical dataset then there can be more robustness of the modelling of this part of the stress event duration histogram.

The impact of the storage Coordination Algorithm can also be seen with a slight change in the distribution shapes.

The reason for the non-linear increase in the storage incremental EFC values (in Tables 2 and 3 earlier) is understandable when observing the shape of these histograms, as the probability of longer events becomes less, and thus proportionally speaking, then the longest duration storage provides incrementally lower reduction in security of supply risk compared to the same amount of short duration storage.

LOLE (hours/yr)	2.966
Mean Event Duration (hrs)	2.141
Expected Number of Events / yr	1.385

LOLE (hours/yr)	2.984
Mean Event Duration (hrs)	2.141
Expected Number of Events / yr	1.394

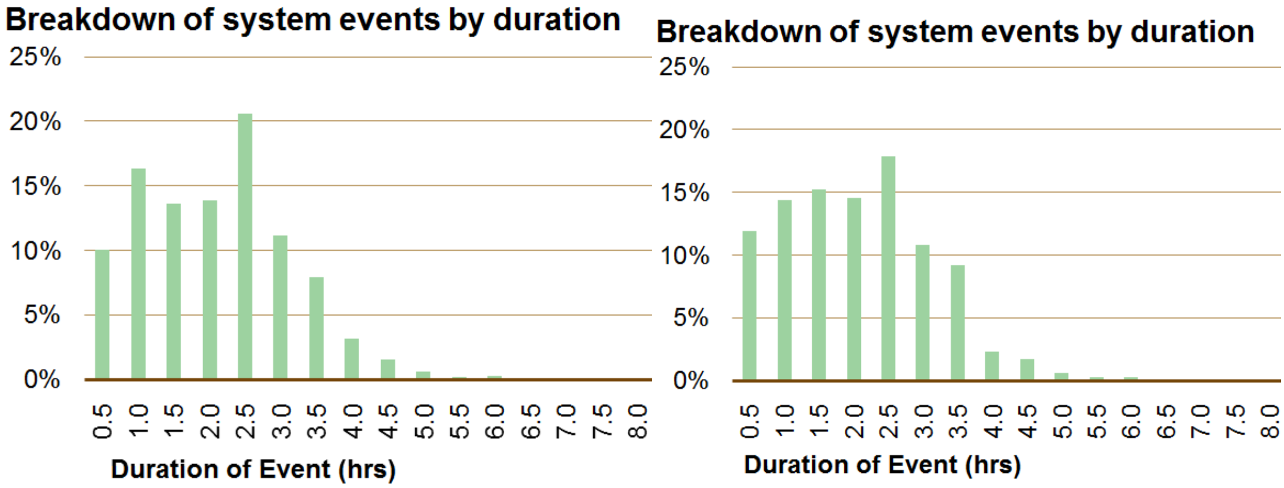

Figure 6 – Histogram of Stress Event Durations 2021/22 – Case #2b (1GW) Storage Coordination Algorithms 1 (on the left), 2 (on the right)

Histograms of Stress Event Durations – 2018/19 Case #1 Storage Coordination Algorithms 1 (on the left), 2 (on the right)

LOLE (hours/yr)	3.088
Mean Event Duration (hrs)	1.967
Expected Number of Events / yr	1.570

LOLE (hours/yr)	3.144
Mean Event Duration (hrs)	1.987
Expected Number of Events / yr	1.582

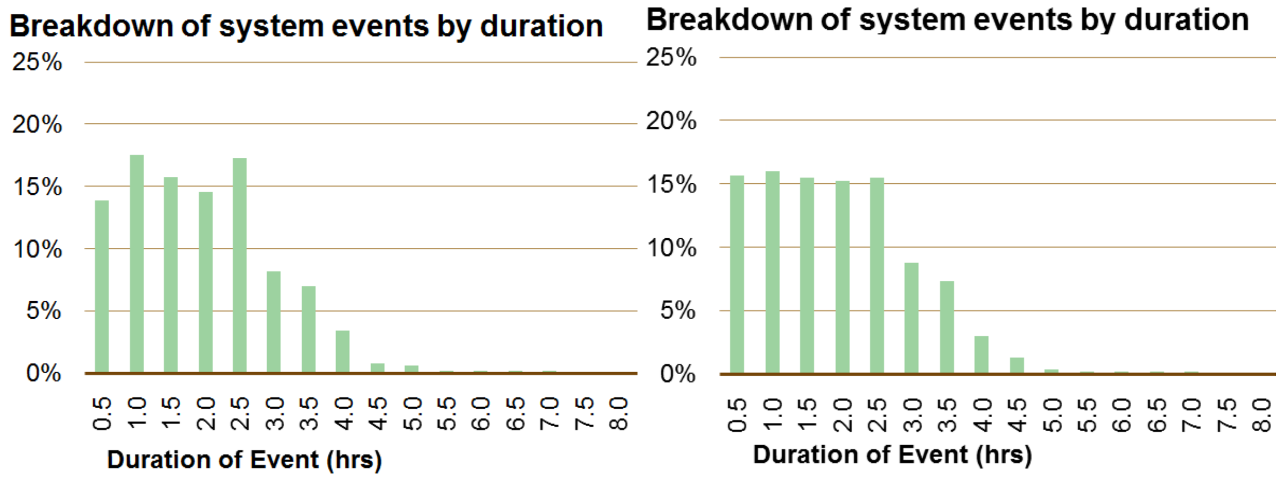

Figure 7 – Histogram of Stress Event Durations 2018/19 – Case #1 Storage Coordination Algorithms 1 (on the left), 2 (on the right)

The impact of addition of storage to the 2018/19 Base Case is broadly analogous to that of the 2021/22 cases discussed above.

Histograms of Stress Event Durations – Important Comments

These histograms indicate the probability distributions of GB system loss of load event durations, for Base Cases that are adjusted to 3 hours LOLE. Industry stakeholders have expressed significant interest in their shape during our EMR methodology consultation, though there are a number of important caveats associated with these distributions that need to be remembered:

- They correspond to EMR Base Cases adjusted to 3 hours LOLE, but the system may have a different reliability level each winter as evidenced by recent Winter Outlook assessments, and the CM auctions may clear at a different capacity level due to a range of factors. The storage de-rating factors must be indexed to the 3 hours LOLE level though, as that is the GB reliability standard target
- The distributions present an indication of the range and probability of stress event durations in the long run, if such an event were to occur. Though there is a small chance of relatively long events, the overall number of events per year is still expected to be relatively low
- Given that the modelling is based on an arguably short 11 years of historical data, there may be an argument made that histogram accuracy in the far right hand side tail can be improved in the assessment in future years as more date becomes available, as these events are quite rare in occurrence in the reliability simulation, and may contain a proportionally higher amount of unserved energy due to their length
- There is no account of emergency system operational actions (e.g. 'maxgen' ancillary service from conventional plants and emergency assistance from interconnectors to other markets) in this modelling framework as they are non-firm, but in the real system they may have some impact in alleviating a tight adequacy margin before a CM stress event might be deemed to have occurred
- The future underlying demand time series used in the UEM tool Monte Carlo simulation are based on historic recorded transmission demand time series scaled up to match the level of underlying peak demand in the future CM delivery year. We can improve on this if NGET-SO gets full access to embedded generation data in future
- These histograms are very much framed by the particular modelling assumptions for storage response in this specific study for the purposes of a storage EFC assessment
- The histograms should thus not be considered as a "forecast" but more so as an "estimate", as the real system behaviour of multiple commercial parties may differ from these necessary modelling assumptions in this particular application
- Industry stakeholders should carefully note these caveats accordingly in the use of these histograms for any decision making.

7. De-Rating Factors Proposed for the 2018/19 and 2021/22 CM Delivery Years

The results in the previous Section 5 indicated the incremental EFC values for a range of different EMR Base Case sensitivity analyses. As discussed in Section 2, then to derive the final de-rating factors for use in the CM auctions, the incremental EFC of the perfectly reliable storage unit on the margin needs to be linearly scaled by the historical technical-availability of the storage class (deemed to be pumped hydro at 96.11% in the 2017 Electricity Capacity Report).

The numerical analyses in Tables 2 and 3 of Section 5 above presented EFC values for a range of storage durations extending up to 8 hours, and the underlying analysis assessed up to 24 hours. The Government Response document published today[15] indicates that storage durations above 4 hours shall receive the traditional de-rating factor based on technical availability alone, hence the de-rating factors proposed in the Table 8 below are presented up to 4 hours. A more detailed discussion of the choice of 4 hours as the cut-off point is included in that Government document.

Note also that a number of Base Case sensitivity analyses for the 2021/22 T-4 CM auction delivery year were presented in the previous Table 3, reflecting a range of uncertainty around the expected capacity and duration penetration of battery storage in the near term GB system. Only one case can be selected for the final de-rating factors however. Therefore, for the reasons outlined in Section 4, we have decided to use the view represented by Case #2b. This corresponds to our updated 'best view' of approximately 1GW of battery storage in total in the case for 2021/22 in addition to the 2.74GW of existing pumped hydro, and 5 MW of liquid air storage.

Table 8 – CM De-Rating Factors Proposed for Duration-Limited Storage Class in the 2018/19 T-1 and the 2021/22 T-4 Auctions

Final De-Ratings Per Duration in Hours	"2018/19"	"2021/22"
Storage Duration: 0.5h	21.34%	17.89%
Storage Duration: 1h	40.41%	36.44%
Storage Duration: 1.5h	55.95%	52.28%
Storage Duration: 2h	68.05%	64.79%
Storage Duration: 2.5h	77.27%	75.47%
Storage Duration: 3h	82.63%	82.03%
Storage Duration: 3.5h	85.74%	85.74%
Storage Duration: 4h +	96.11%	96.11%

[15] Government Response to consultation: https://www.gov.uk/government/consultations/capacity-market-consultation-improving-the-framework-detailed-proposals

8. Discussion and Future Work

This report has presented de-rating factors for duration limited storage in the 2018/19 and 2021/22 CM delivery years based on a new Equivalent Firm Capacity (EFC) methodology that we, the EMR Delivery Body, have consulted with industry upon earlier this summer. From a CM governance process point of view, the introduction of the new storage de-rating factors is a matter for the Secretary of State to decide upon. In that respect, we note the Government Response published today[16] which indicates that these de-rating factors will in fact be used in the upcoming 2018 T-1 and T-4 CM auctions.

The results in this report indicate that storage projects of different durations do contribute in a distinct manner to security of supply – the reasons behind this distinction are clear once the shape of the stress event duration histograms are reviewed. Events are more likely to be shorter than longer, with average stress event duration of circa 2 hours expected when the system is at a reliability state equal to 3 hours LOLE, the GB reliability standard and CM target reliability level. Events of duration > 4 hours are expected to be reasonably rare for the T-1 and T-4 CM delivery years of interest at this time – though such patterns may change in future.

The numerical results also support the proposition that an EFC based assessment methodology is a framework wherein the de-rating factor attributed to each storage duration sub-class is directly consistent with its contribution to security of supply. Usage of de-rating factors based on such a methodology for a range of storage duration sub-classes should help to remove any barriers to short duration storage participating transparently in the CM. It should also allow the CM auctions to continue to be conducted in as efficient a manner as possible, so that all CM participants are fairly rewarded and that security of supply continues to be procured for the electricity consumer at least cost.

We have also conducted a large amount of sensitivity analysis in our numerical modelling studies that indicates:

- The capacity MW amount and constituent duration of the storage fleet assumed in the modelling Base Case has a material impact on the EFC results derived thereafter. For a greater penetration of duration-limited resources in the study case, then the lower the outturn EFC results will be, as stress events will tend to be longer and the incremental contribution of short duration resources to security of supply will begin to saturate
- The use of the 'incremental' EFC of a small storage unit added to the margin at the point which the CM is expected to deliver is a more sensible approach to base the storage de-rating factors upon than using the 'average' EFC of the entire storage fleet overall. This is in keeping with the economic principle of payment in a market being linked to the marginal contribution of supply to meeting demand at the point at which the market is expected to clear. Furthermore it allows direct disaggregation of the contributions of different storage sub-class durations to security of supply.
- The statistical risk metric upon which the EFC values are derived also has a strong impact. Our modelling results show that Expected Energy Unserved (EEU) has a superior performance for conducting an EFC assessment of a duration limited resource than Loss of Load Expectation (LOLE), as EEU is not as sensitivite to the operational strategy of the storage during a stress event, and further, it has a direct link to the customer's economic cost of unreliability in the system by virtue of the GB Value of Lost Load (VoLL) parameter
- The EFC results are only moderately influenced by the underlying adequacy margin (LOLE level) of the EMR 5-year Base Case, the mean time to repair parameters of

[16] Government Response to consultation: https://www.gov.uk/government/consultations/capacity-market-consultation-improving-the-framework-detailed-proposals

conventional supply, and the size of the incremental storage unit applied in the reliability assessment.

The methodology used to derive the new proposed de-rating factors has undergone extensive independent review by BEIS and their Panel of Technical Experts (PTE), and has been endorsed by those parties. The PTE commentary which is broadly supportive of the methodology used and findings of this work is included in Appendix 4.

We have also been guided by expert academic advice from staff at the University of Edinburgh to ensure that the work is reflective of the state of the art in theory and applied modelling of storage, adequacy risk and reliability of power systems – their comments on these issues are included in the Appendix 3 section.

The amount of international experience on this topic to benchmark the GB approach is relatively low. There is only significant capacity market experience with short duration storage in a handful of power systems to date. One useful benchmark that we are aware of would be the recent Irish All Island Single Electricity Market capacity market auction parameters for storage of different durations[17]. Though the exact storage de-rating factor methodology followed by Eirgrid was slightly distinct from ours, and of course it is a different system with different characteristics and a different target reliability level, the end result de-rating factors presented are broadly similar for the short-duration storage categories in both GB and Ireland.

We will update the de-rating factors for duration limited storage for each subsequent year's Electricity Capacity Report using this new methodology, using the best available modelling information we have available in future. There are a number of additional future work directions that we can take to further develop our methodology in the future as it transitions in to an annual process. In particular the development of duration limited storage on the system in future may allow us to get access to real system performance data. At the moment, we are constrained to make a number of modelling assumptions in the absence of any real-life comparisons for validation. In future:

- If we have access to historical data we can update the assumptions around specific storage technology "technical availability"
- We will be able to assess how the storage charge state evolves over time during the course of days where GB system adequacy margins are tight
- We will be able to build in the true response behaviour of duration limited storage resources to both Capacity Market Notices, and indeed CM stress events
- We also may need to further understand how the capacity obligation performance requirements are structured with respect to the new de-rated capacity allocated to successful storage bids in the CM auctions
- We will be able to understand if there is any significant diversity of responses between CM storage that is participating in the Balancing Market, and that which is embedded plant relying on autonomous triggering of power production on the basis of e.g. NGET control centre operational warning notices
- We can better understand how the impact of ancillary service provision and other commercial services overlaps with the CM obligation requirements
- We can further refine our methods and sources of historical data to ensure that statistically rare long duration outage events are appropriately captured in the modelling

[17] Eirgrid/SONI Initial Capacity Market Auction Parameters: http://www.sem-o.com/ISEM/General/Initial%20Auction%20Information%20Pack.pdf

- We can review the historical transmission demand time series used in the modelling. If we can acquire accurate time series data on the demand met by embedded generation we could generate a historic underlying demand time series
- We can improve our modelling of solar PV contribution to system adequacy during “shoulder” hours so that storage charge state is appropriately reflected, etc

We look forward to engaging with industry and other stakeholders in due course as the storage de-rating factor modelling approach evolves.

Appendix 1 - Solar PV Modelling Assumptions in Detail

Power system reliability simulation in GB has traditionally focused on model accuracy at time of peak demand (usually after dark in winter) as this was considered as the time of day where capacity shortage risk is most concentrated. Therefore, wind power was the main variable/uncertain renewable resource incorporated in the models and solar PV effects were understood to be negligible and were thus ignored.

In recent times, the amount of solar PV in GB has been growing and is expected to grow further, both in terms of rooftop residential PV but also utility scale embedded PV projects. The EMR 5-year Base case includes ~ 13.5GW of solar PV in 2018/19 and ~ 16.5GW in 2021/22. Solar PV has an impact on GB LOLE and EEU modelling in terms of

- Correctly estimating the historical underlying demand time series shape/stochasticity
- Correctly accounting for all supply availability in future risk model target years

Initially this project did not include solar PV however early model results challenged this assumption. The Figure A1.1 below shows a histogram of stress event durations for the 2021/22 Case #2b **without any solar PV in the model**. It can be observed that there is a small probability of very long duration events ~ 10-11 hours under this assumption. This would imply that some events could span daytime hours where solar PV could be important.

Duration of Event (hrs)	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5
Expected Quantity	0.135	0.185	0.076	0.104	0.227	0.144	0.105	0.062	0.039	0.012	0.013
Proportion of events	11.6%	15.9%	6.6%	8.9%	19.6%	12.4%	9.1%	5.3%	3.4%	1.0%	1.1%
LOLE (hours/yr)	3.035										
Mean Event Duration (hrs)	2.616										
Expected Number of Events / yr	1.160										

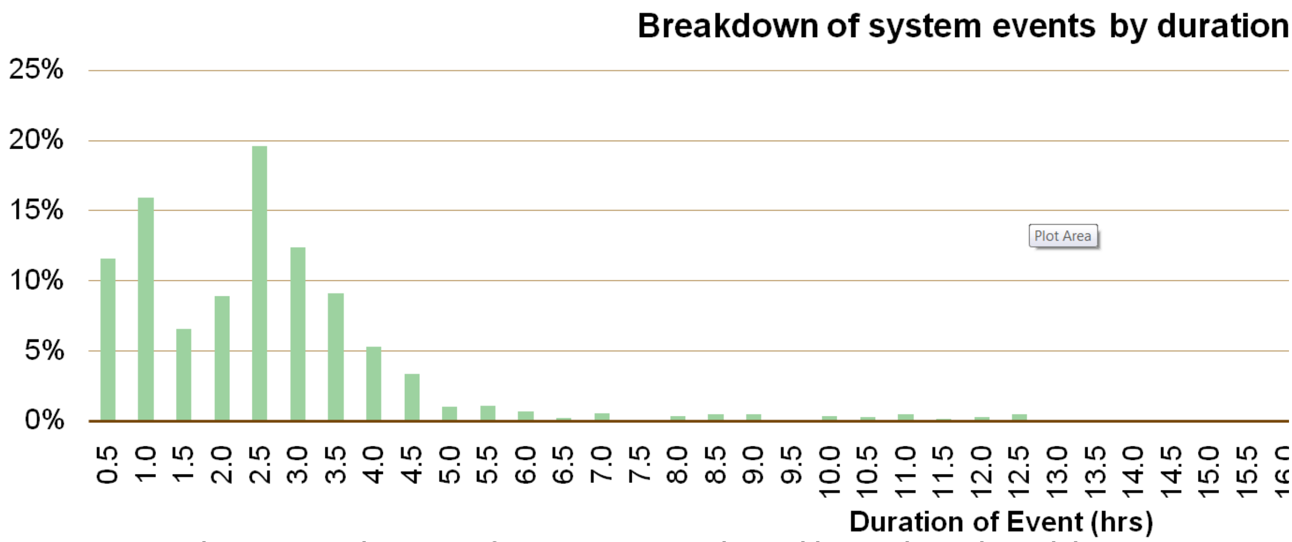

Figure A1.1 – Histogram of Stress Event Durations Without Solar PV in Model

The reason why there are a few loss of load outage events of very long duration may be driven by the shape of the load time series on peak demand days (see a very stylised depiction below in Figure A1.2). There is sometimes a slightly higher demand at ~ 11am than at 3pm, and further, the demand between 11am and 4pm can be relatively flat, on peak demand days.

Therefore, if there are a number of conventional plants simultaneously on forced-outage, then an event could in theory last much longer than the peak demand hours and thus span the daytime – where solar PV may be available.

We therefore decided to include an estimate of solar PV in the modelling, with the best available information we had within the short time constraints of this project. Ideally we would use a historical solar dataset that is of 11 years of length and is furthermore time-stamp coincident with our existing demand and wind historical time series.

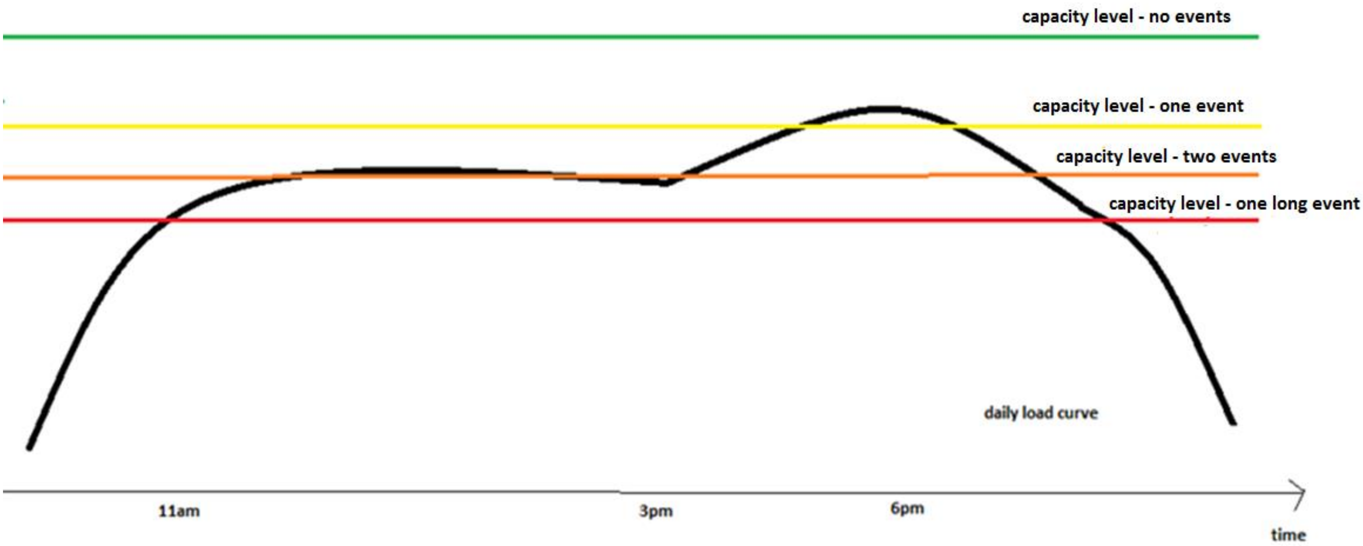

Figure A1.2 – Indicative Daily Load Curve Shape Influencing Stress Event Duration

Only a 2011-2017 estimated embedded solar data set was easily available for use however from the NGET-SO operational databases, which upon visual investigation, seemed most reliable in the 4 years of 2013-2017. We added this time series to the load demand time series to get a better estimation of historical underlying load demand time series. We also “stitched together” 12 years of solar data by looping the 4 reliable years of data 3 times so that we could have a model of solar PV contribution to supply adequacy in the future model target years. This will accurately represent solar/load/wind statistical dependency for 4 years of the historical dataset, but will implicitly assume independence for the rest of the years. This is a noted area of improvement for future years though is much better than the “do nothing” scenario.

The following diagram in Figure A1.3 shows the histogram of the loss of load outage events for 2021/22 Case #2b which shows that the probability of long duration events decreases substantially **when solar PV is included in the model.**

Duration of Event (hrs)	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	5.5
Expected Quantity	0.176	0.205	0.099	0.116	0.262	0.211	0.117	0.061	0.033	0.015	0.005
Proportion of events	13.4%	15.6%	7.5%	8.8%	19.8%	16.0%	8.9%	4.6%	2.5%	1.2%	0.4%
LOLE (hours/yr)	3.040										
Mean Event Duration (hrs)	2.304										
Expected Number of Events / yr	1.319										

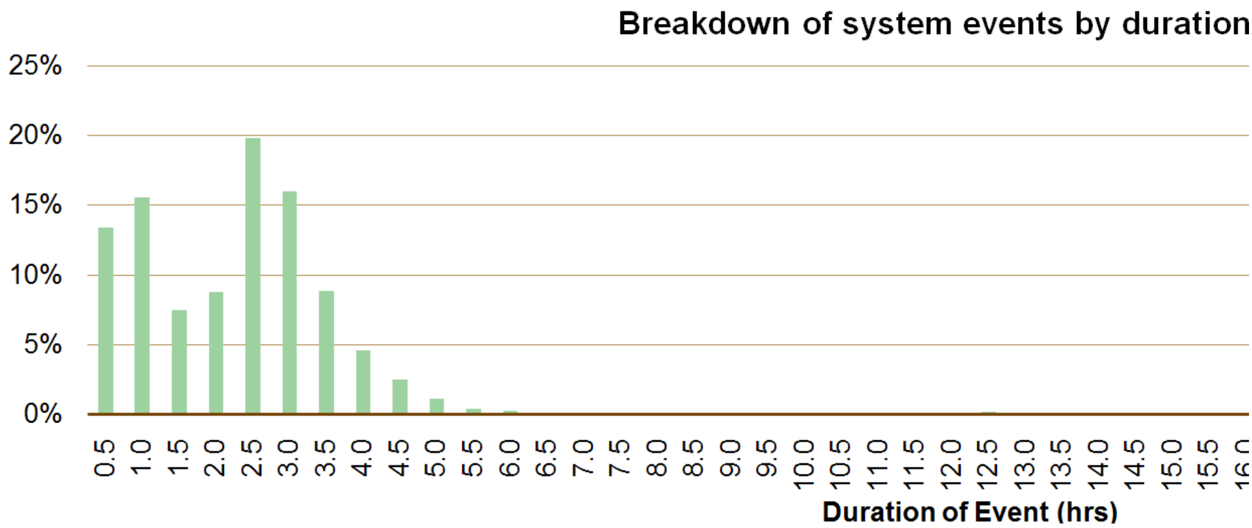

Figure A1.3 – Histogram of Stress Event Durations With Solar PV in Model

The following Figure A1.4 of half-hourly solar PV capacity-normalized output in Jan/Feb 2016 indicates that there may be days in winter where the fleet reaches 30-40% of available power in the daytime hours, though further work is required to assess the contribution of this to security of supply in shoulder hours that may impact the value of a duration-limited storage EFC the most. While solar PV may have a material impact on the storage EFC calculation, we do not expect it to have a material impact on the calculation of the capacity to secure figures published in the Electricity Capacity Report.

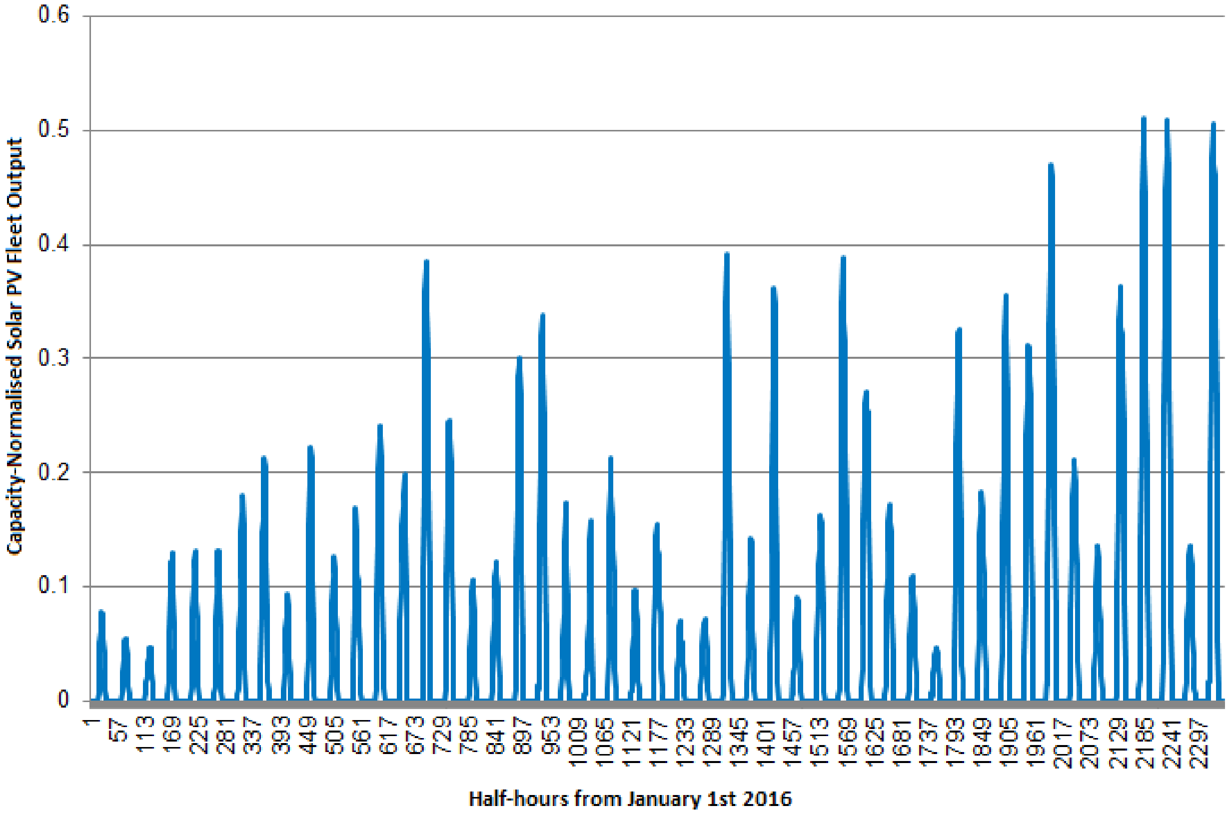

Figure A1.4 – Solar PV Time Series Over ~ 2 months in Early 2016

The solar PV data we used is described in the "*Quarterly Forecasting Report - June 17*" published at the NGET demand data page[18] which has the following description of the solar PV forecasts. The solar generation forecasts, arise from an internal PV generation forecasting model, based on a number of parameters:

- Estimated capacities,
- Weather forecasts from a weather provider at 28 geographical locations and
- Empirically derived models connecting radiation and national generation using data from the collaboration with Sheffield Solar.

The solar PV data that we used in this project was from the "*DemandData_2011-2016*" file on this website link in the footnote.

[18] NGET demand data:
http://www2.nationalgrid.com/UK/Industry-information/Electricity-transmission-operational-data/Data-explorer/

Appendix 2 – Plant Mean Time to Repair Update

As part of this project we also carried out an update on the previous 2012 analysis of the Mean Time to Repair (MTTR) values for the LCP UEM sequential simulation. As per the usual EMR convention for de-rating factors we used historical data based on last 7 years Maximum Export Limit (MEL) for overall technology types, with some embedded plants being mapped to “closest” transmission system technology.

We used the following outage definition rule in the data assessment:

- *Outage starts when real time MEL < 20% of maximum MEL*
- *Outage stops when real time MEL > 70% of maximum MEL*

The maximum MEL is calculated for each Balancing Market Unit (BMU) for each winter, though with the following plant exclusion rules :

- Outages longer than 6 months (assumed to be mothballed)
- Outages starting outside winter months (November to March)
- Closed BMUs
- Supplemental Balancing Reserve (SBR) sites from 2016/17 winter
- BMUs with less than 1000 data point observations.

An approach was also required to treat data outliers in the dataset. For each BMU a 99 percentile MEL was calculated. If the “real time” MEL > 1.15 * 99 percentile MEL then this data point was assumed to be an outlier. It was therefore:

- Replaced by “gate closure” MEL if this was less than 1.15 * 99 percentile MEL
- Otherwise “real time” MEL substituted with 99 percentile value

To calculate mean time to repair, the outage durations were converted to days and percentages of days and the final MTTR value calculated for each transmission system technology class is indicated below in Figure A2.1:

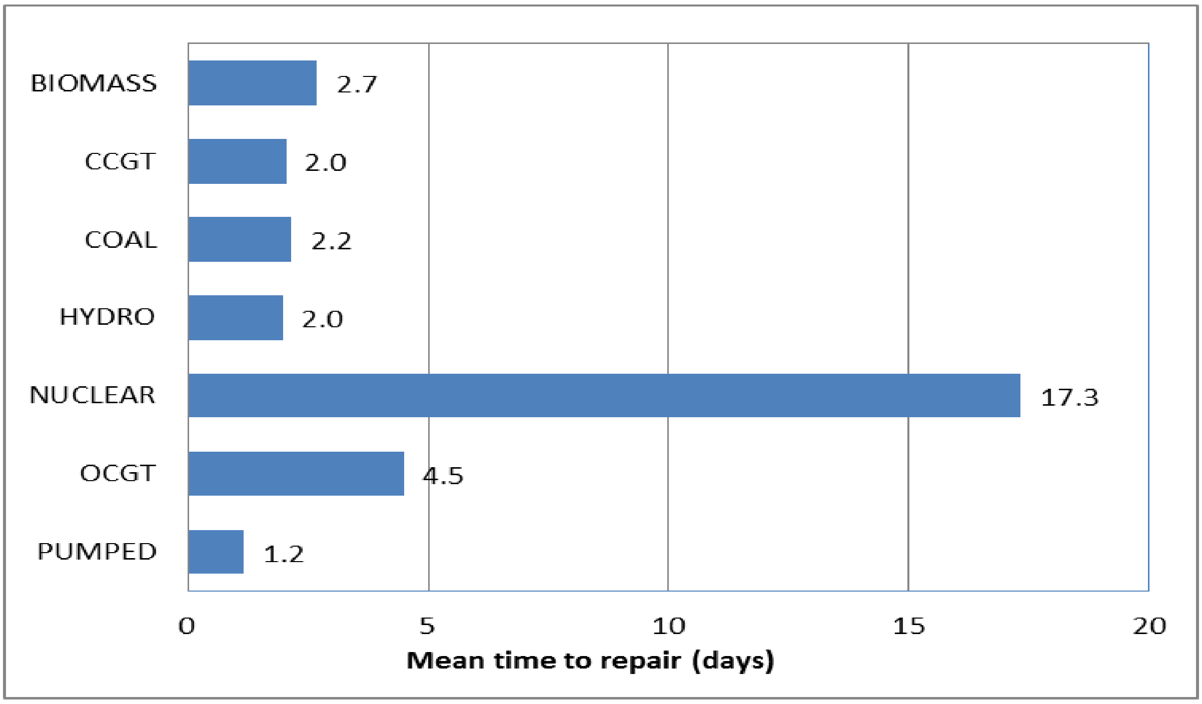

Figure A2.1 – Updated MTTR Values Derived for Transmission Technology Classes

The following Figure A2.2 shows real time MEL values on 3 specific plants (not named for confidentiality reasons) at the top with outage periods below. Light grey coloured MEL data points are where an outage has been identified and black coloured MEL data points indicate when the unit is assumed to be on. The red diamonds show the period the unit returns from outage.

These examples show that there are still some minor anomalies caused by the selected outage rule. However the impact on storage de-rating factors is not material as shown by the earlier sensitivity analysis results in the main report on Page 24.

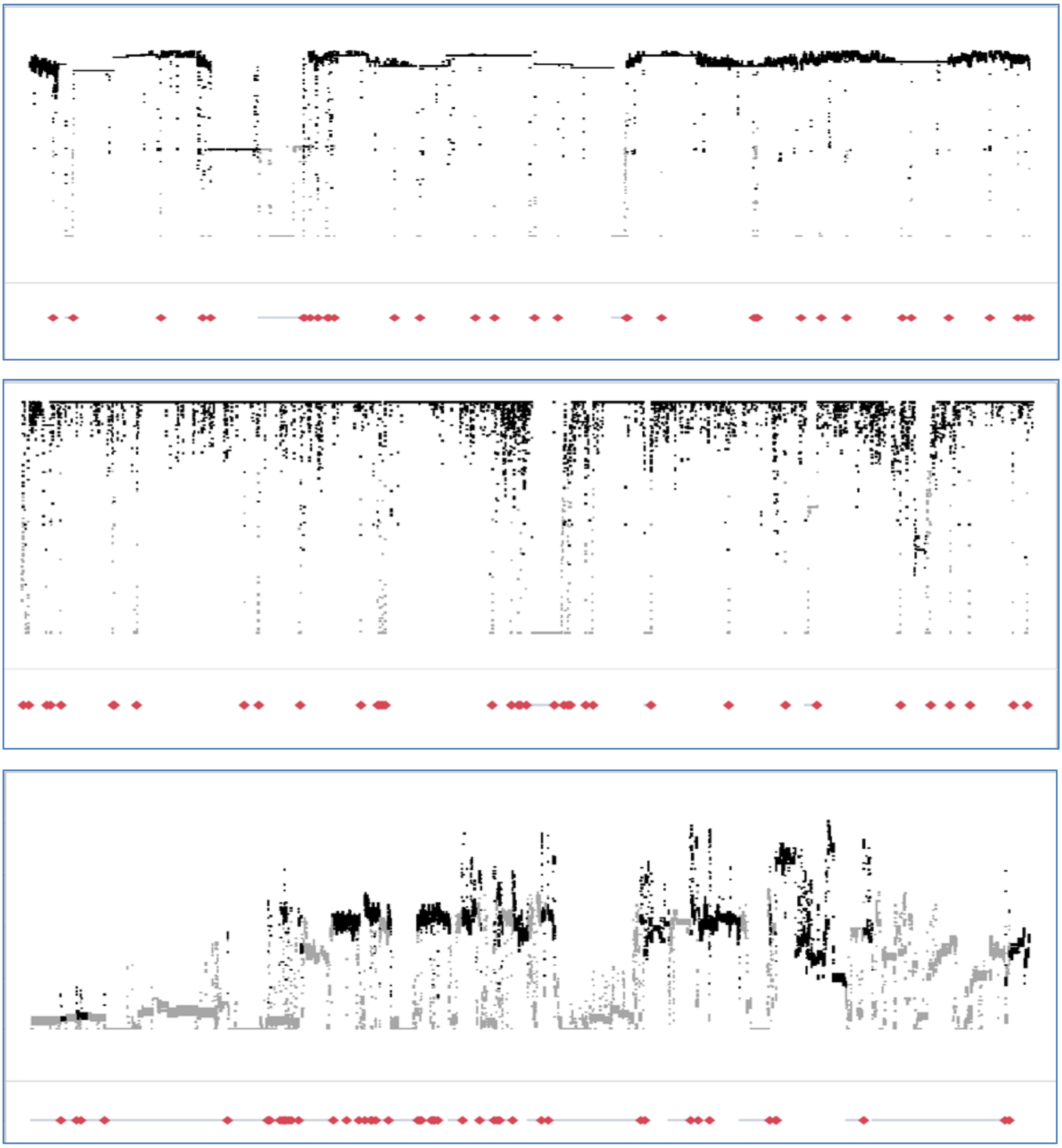

Figure A2.2 – Indicative MEL Time Series of Plant Availability

As indicated in the following Figure A2.3, then no clear trends were evident in the MTTR values calculated between years of the dataset so all years were thus included in 7-year averaged calculations for Figure A2.1.

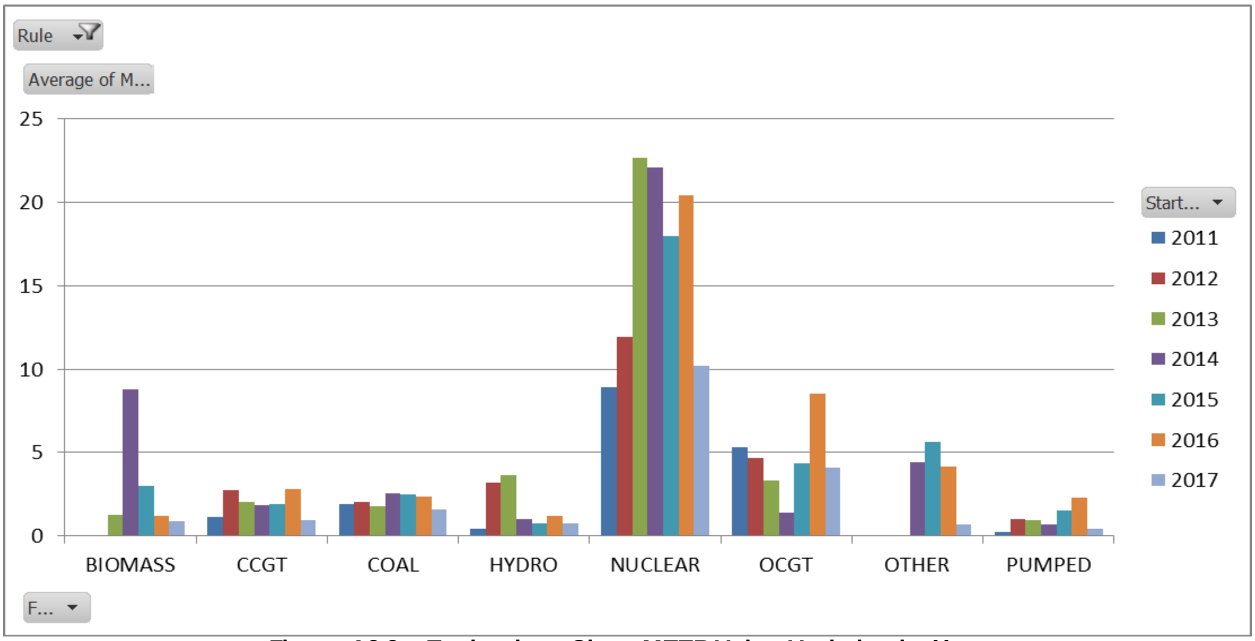

Figure A2.3 – Technology Class MTTR Value Variation by Year

Appendix 3 – Academic Expert Methodology Commentary

Valuation of storage for capacity markets

Chris Dent, Amy Wilson and Stan Zachary

University of Edinburgh

October 27, 2017

In the operation of the GB capacity market, it is necessary to assign an *equivalent firm capacity* (EFC) to storage facilities. We comment on two issues here.

Choice of risk metric against which to calculate EFC

The two risk metrics in primary use for measuring security-of-supply are *loss-of-load expectation* (LOLE) and *expected energy unserved* (EEU).

EEU is the more direct measure of economic damage due to a shortfall in supply, while LOLE takes no account of severity of shortfall (which is clearly important). However, the use of LOLE as a measure of security-of-supply is traditional. In the case of electricity systems in which the supply is provided solely by fully dispatch-able generation, it makes very little difference which of these two metrics is used. This is because, in any such system, as the level of generation is varied there is essentially a one-one correspondence between values of EEU and values of LOLE, so that either might be used to set a standard for that system. Further the EFC of variable generation (e.g. wind) will be the same no matter which metric it is calculated against. If EEU is the more direct economic measure, then, for both traditional and variable generation, LOLE is nevertheless a perfectly reliable proxy measure.

In the case of storage, this correspondence between EEU and LOLE no longer holds. The reason for this is the many different ways in which the limited energy content of storage may be used. While given storage makes very much the same contribution to reduction in EEU no matter how it is used (provided its energy content is not wasted, and on the assumption that it is able to fully recharge between supply shortfall periods), its contribution to reduction in LOLE varies greatly.

As an illustration we give three examples of ways in which storage with given capacity and maximum rate (power) might be used to mitigate some given period of continuous shortfall in supply. We assume that this storage is insufficient to entirely eliminate this shortfall. Assuming that none of the storage is wasted (e.g. through failing to correctly predict the shortfall profile), in each case the achieved reduction in unserved energy is simply the initial volume of energy in store. However, the reduction in *loss-of-load-duration* varies greatly between the examples.

A. If the storage is used so as to achieve the *maximum reduction in loss-of-load-duration* (minimise the length of the residual shortfall period) then it will be deployed at those times for which the prior shortfall in supply is least (so as to enable it to eliminate this shortfall for the longest possible period) and may well not contribute anything at those times at which it does not eliminate the shortfall entirely. This is in contrast to the firm capacity which would achieve the same reduction in loss-of-load-duration, as the latter would continue to contribute to reduction in unserved energy throughout the *entire* shortfall period. Such firm capacity would therefore contribute the same reduction in LOLE as the storage, but would contribute a greater reduction in EEU. Here the LOLE metric clearly *overvalues* the contribution of storage relative to that of firm capacity.

B. If the storage is used so as to *minimise the maximum shortfall* over the considered period, then there may be no reduction in *loss-of-load-duration* whatsoever (the storage simply shaving off the peak of the shortfall profile). Nevertheless the storage continues to make a significant contribution to the reduction in unserved energy. (Indeed one might argue that here the reduction in unserved energy is achieved at the times when it is most needed. Here the LOLE metric clearly *undervalues* the contribution of storage.

C. If the storage is used at as constant a rate as possible throughout the entire period of shortfall, then the reduction in *loss-of-load-duration* is intermediate between those achieved in the preceding two examples, and is the same as would be achieved by that firm capacity which also resulted in the same reduction in unserved energy. Here the LOLE metric *equitably* values the storage relative to firm capacity.

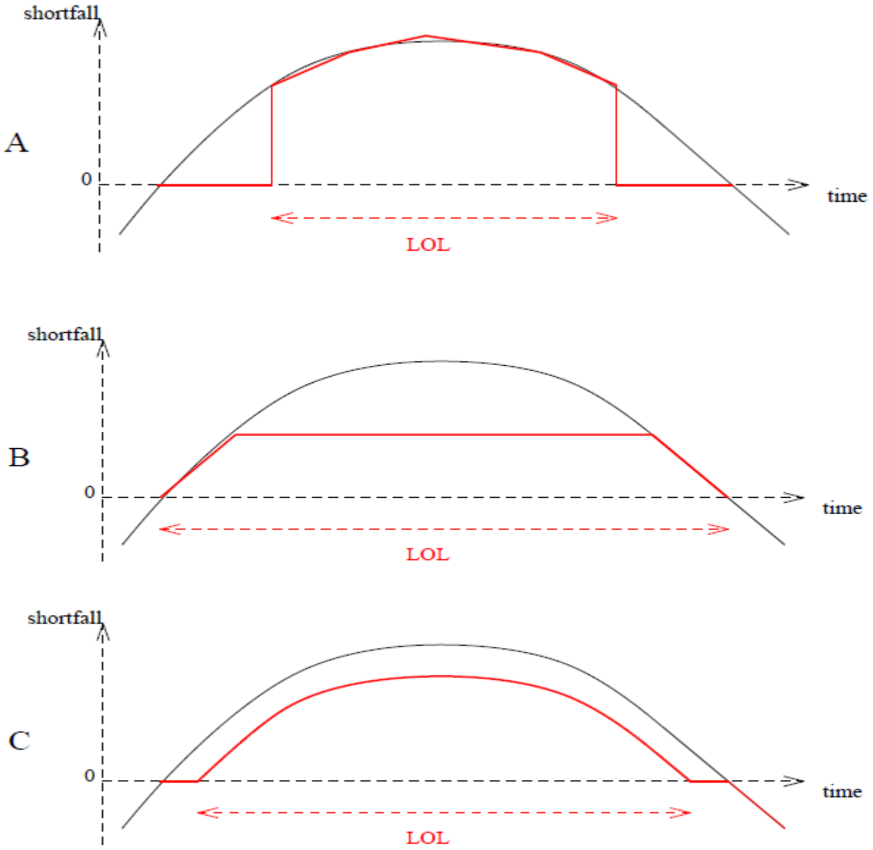

Figure 1: Three possible ways in which stored energy may be used to mitigate shortfall. The labels A, B, C correspond to the three examples discussed.

The three examples are illustrated in Figure 1, where in each case the black and red curves plot (depth of) shortfall against time respectively before and after the use of the given storage. The area between the black and red curves corresponds to the reduction in unserved energy due to the use of the storage and is here equal to the volume of energy available in the store, and is hence the same in each case. As described above, the reduction in loss of load is different in each of the three examples.

Thus, for the reasons discussed above, we believe that EEU is the correct metric against which to value the contribution of storage.

The above conclusion does not—at least at present—affect the suitability of LOLE (e.g. the current standard of 3 h/y) as a target metric for the determination of the capacity-to-procure within the capacity market itself. The amount of storage currently in existence is insufficient to affect the one-one correspondence between system LOLE and EEU referred to above, and further this overall relationship is not significantly affected by the way in which currently available storage might be used. Thus either metric might in principle be used to set a target, and it is LOLE which is currently sufficiently well understood for this purpose. As discussed above, the importance of EEU is in the correct determination of the relative contributions of different types of capacity provision.

Determination of EFC for participation in a capacity auction

We assume that the objective of a capacity auction is to obtain the required level of security-of-supply (as measured by whichever risk metric is chosen for this purpose) at minimal cost. For this to happen it is necessary that the contributions of the various facilities participating in the auction—as measured by their EFCs—are correctly valued relative to each other *at the point where the market clears*. This is, at least in principle, the point at which it is considered that sufficient capacity has been obtained to meet the required standard—and is at present that defined by the current standard of 3 h/y LOLE.

To see this, note that typically a capacity auction accepts bids in ascending order of their cost to EFC ratios until sufficient capacity is obtained. At this point it should *not* be possible to identify an unsuccessful participant providing capacity at lower unit cost than a successful participant—otherwise the former should be swapped for the latter both so as to reduce the overall cost of providing sufficient capacity, and also of course as a matter of market fairness. The measure of capacity contribution is thus the *incremental* (or *marginal*) EFC evaluated at the point where such transactions might be considered, i.e. at the point where the market is expected to clear. Thus the appropriate measure of capacity contribution of any facility participating in the auction is that level of firm capacity which results in the same reduction in the risk metric at this point.

Note that, since shortfall periods are very rare and typically of relatively short duration (of the order of a small number of hours) the incremental contributions of individual facilities are typically quite small at the market clearing point (and would become smaller were the security-of-supply standard to specify a lower level of risk). However, the EFC is a measure of *relative* contribution, and at the point where the market clears the EFC of storage in particular is typically higher than would be the case were shortfall periods significantly longer. Thus a store able to provide power of 10 MW for 3–4 hours during the day may, at the above margin, be nearly as useful as 10 MW of generation able to run continuously.

Finally, we remark that when designing capacity auction processes, it is important to remember that the EFC of an ensemble of units of resource (generation or storage), defined relative to the required target risk level, will not in general be equal to the sum of the *incremental* EFCs of the individual units defined relative to that risk level. Thus when using incremental EFC in a capacity auction, care must be taken to ensure that the outcome of the auction is a system with the intended risk level.

Appendix 4 – BEIS Panel of Technical Experts Review Comments

Panel of Technical Experts[19] – Conclusions on Storage De-rating Methodology Proposed by National Grid

The Panel of Technical Experts (PTE), having reviewed National Grid's (NG) proposed Storage De-rating Methodology, is content with the proposed approach and notes that NG has undertaken a convincing piece of work. The analysis to determine de-rating factors for storage is very thorough and based on appropriate fundamental principles in the context of the present CM framework, leading to the proposed approach to derating storage being robust.

The approach may not be fully optimal but is clearly a big improvement on the status quo. We have not identified anything better at this stage in development of the Capacity Market and storage business.

An emerging caveat is whether the ability of storage to tap into other revenue streams may affect the ('fully charged') hours it could prudently commit and deliver to the Capacity Mechanism. We do not believe this affects the recommendations at this stage but it is an area that should be subject to further analysis, and evaluation of actual performance of storage registered under the CM.

[19] The list of PTE members:
https://www.gov.uk/government/groups/electricity-market-reform-panel-of-technical-experts