# M&V & Flexibility Application Guide

February 2026.

Version 11.

Acknowledgements

Many thanks go to the following people and companies that have greatly contributed to the writing of this guide.

<table>
<tr>
<td></td>
<td>Rajvant Nijjhar (Flex Committee Chair, iVEES & BankEnergi)<br>www.bankenergi.com : and www.ivees.co.uk<br>LinkedIn: (25) 💎 Rajvant Nijjhar 💎 PMVE | LinkedIn<br><br>Industry M&V professional, trainer and entrepreneur, working as UK M&V Expert for the British Standards Institute and UK trainer for PMVA courses since 2014 having trained over 2.500 persons globally. Worked with Technical Committee 301 developing ISO standards in energy management (ISO 50001 revision), M&V (ISO 50015), energy baselines (ISO 50006) and energy savings within projects (ISO 17742). To date, I have developed many M&V Plans and savings reports in both public and private sectors. I have also acted as a third-party independent M&V specialist on several projects. Currently also working on start-up, BankEnergi, related to energy and flexibility trading as a Flexibility Services Provider to local and national markets</td>
</tr>
<tr>
<td></td>
<td>Ana Soares (INESCC / University Coimbra).<br>Ana Soares | LinkedIn<br><br>Assistant Professor at the University of Coimbra, where I teach topics related to energy management and electronics. I'm also Member of the board of Directors and Researcher at INESC Coimbra. I was a member of the Coordination Committee of INESC Coimbra's Scientific Council between November 2021 and December 2024. Ana has a PhD in Sustainable Energy Systems from the University of Coimbra (MIT Portugal Program) and two master's degrees: one in Electrical Engineering and Computers (Energy) and another in Energy for Sustainability (Energy Systems and Policy).</td>
</tr>
<tr>
<td></td>
<td>Touray, Haddijatou (SSE plc)<br>Haddijatou Touray MIET | LinkedIn<br><br>A certified Performance Measurement and Verification Analyst with over 3 experiences in Sustainable Energy Systems design, Measurement and Verification, Energy Systems Optimisation (energy systems modelling) and Building Services Audit. Currently working on building more industry experience in Power Systems Engineering specifically, HVDC. I have acquired invaluable engineering and management skills through the various roles I have served within the energy industry. I am particularly passionate about tackling challenging and innovative technical problems related to energy sustainability.</td>
</tr>
</table>

| | |
|---|---|
|  | Sergio Leon (ESCo, Mexico)<br>Sergio Leon Marquez \| LinkedIn<br><br>*Sergio is a high-performance energy manager with deep expertise in optimizing electricity use and enabling demand response for HVAC and industrial refrigeration systems. He has a strong track record of creating measurement and verification plans that help commercial and industrial users reduce energy costs. His work reflects a clear commitment to advancing sustainable, data-driven energy solutions. As a certified Performance Measurement and Verification Analyst (PMVA) by EVO, he brings the experience and analytical insight needed to develop practical strategies for quantifying and applying energy flexibility in support of decarbonization, electrification, and digitalization goals.* |
|  | Thomas Yeh (NYSERDA)<br><br>*Thomas Yeh is a recognized expert in energy management, energy efficiency, and utility incentive programs. He currently serves as a technical advisor and consultant for NYSERDA. Throughout his career, Thomas has led the development of innovative technologies in areas such as data analytics, wireless lighting control, embedded gateway design, IoT device networking, and energy efficiency in data centers. He holds 20 U.S. and international patents and is certified by the Association of Energy Engineers in several areas, including as a Certified Energy Manager (CEM), Certified Measurement and Verification Professional (CMVP), Certified Building Commissioning Professional (CBCP), and Certified Demand Side Manager (CDSM). Additionally, he is accredited by the International Geothermal Heat Pump Association. Thomas earned a bachelor's degree and a master's degree in Electrical Engineering, as well as a master's degree in Computer Science.* |
|  | Maggie Selig<br>Maggie Selig, PE, PMVA, CMVP \| LinkedIn<br><br>*Maggie is a proven expertise and leadership in the energy services industry for over 35 years. Driven to deliver high performing energy projects for positive impact to our customers, our companies and our world. My career started in the mechanical design and construction field, then moved into the energy consulting and project development arena in the late '80s to focus primarily on energy efficiency. I have worked for several nationally known ESCOs and consulting firms in various roles, including engineering management/quality assurance, M&V system design and implementation, and corporate technical risk management.* |

| | |
|---|---|
|  | Vibhor Bhatnagar (Abu Dhabi National Energy Company)<br>[Vibhor Bhatnagar \| LinkedIn]<br><br>*Vibhor is highly qualified engineer and PMV, a member of Indian Green Building Council (IGBC) and competent in managing & monitoring projects with MEP Engineering Design, Building Energy Modelling & Simulation, Measurement & Verification (M & V) Data Analytics, Operational Performance Analytics, Project Management, Energy Management (ISO 50001) & Physical Asset Management (ISO 55001) practices. Building Performance Analysis (using Design Builder, IES VE & RETScreen), MEP Engineering Design (utilizing AutoCAD & Revit MEP), Energy Auditing and Operational Risk Analysis of Mechanical & Electrical systems.* |
| | Javier Espinoza (Grupo INGESA)<br><br>*Javier Espinoza is a Costa Rican electrical engineer and founder of Grupo Ingesa*, a firm specialized in energy efficiency and ISO 50001-based energy management systems. With over 18 years of professional experience, he has led more than 65 successful ISO 50001 implementations across diverse sectors, including healthcare, food and beverage, hospitality, retail, and logistics.*<br>*He holds a Master's degree in Business Administration from the Costa Rica Institute of Technology and is a Certified Energy Manager (CEM) by the Association of Energy Engineers (AEE). He is also internationally accredited as a Performance Measurement and Verification Analyst (PMVA).*<br><br>*Grupo Ingesa is a Costa Rican Energy Services Company (ESCO) specialized in energy efficiency and the design and implementation of energy management systems under ISO 50001. With over 65 projects executed and a 100% success rate in certification processes, the firm provides technical solutions, energy audits, and performance monitoring and verification services. Grupo Ingesa combines innovation, sustainability, and profitability to turn energy use into a competitive advantage.* |
|  | Sandeep Kachhawa (AEEE)<br>[Sandeep Kachhawa \| LinkedIn]<br><br>*Sandeep is an energy professional and energy auditor for large commercial and mixed-use buildings. He works with different stakeholders from the government, civil society organizations, academia, and the private sector on various technical, policy and regulatory matters related to energy efficiency uptake and implementation. He helps bring focus on the energy efficiency and sustainability aspect in the national cold chain discourse. He has authored/co-authored multiple research papers, including making a case for adaptive thermal comfort, decoding India's residential and commercial building stock characteristics, rethinking ECBC implementation in states and cities, evaluation framework for assessing state-level energy efficiency implementation, and policy & technology roadmap for retrofitting/modernizing cold storages in India.* |

| | |
|---|---|
|  | Chandan Sarkar (Sustainability & ESG Consultant<br>Chandan Sarkar \| LinkedIn<br><br>*Sandeep is a highly experienced and results-driven sustainability, and ESG professional with over 15 years of expertise in spearheading and implementing innovative solutions for energy efficiency, water efficiency, waste management (circularity), and green procurement projects. Adept at working and collaborating with diverse stakeholders, including corporations, governments, and non-profit organizations, to drive sustainable practices and achieve measurable environmental and social impact.* |
| | Bill Campbell EQ Energy |
| | Leiven Van S – Belesco<br>Mattia Barbero 0 Bamboo Energy |

Reviewers

- UKPN
- NYSERDA
- Natasha Mills, ADE via Flex Assure.
- Lind, Leandro, (via Sergio)

Table of Contents

## Contents

# 1. Background and Introduction

This guide, developed by the EVO Committee on Measurement and Verification (M&V) and Flexibility, aims to establish globally recognized best practices in energy and flexibility trading.

## 1.1 Approach:

The approach to the development of the guide is via international committee of experts with experience in both M&V and flexibility services. Reference is made to global best practices and to create parallels to the IPMVP, via the use of international case studies.

Many DNOs/DSOs/TSOs are currently developing own guides to baselines, with a global standardized approach to be delivered. We reference those organisations that have advanced in baselining and explain the practical similarities and differences to the baselines created under Options A to D of the IPMVP.

## 1.2 Who is this guide for?

This guide is primarily to support energy auditors and M&V professionals that might need to create baselines for the purposes of flexibility trading but can also be used by:

- Market operators such as DNOs, DSOs, ISOs and TSOs in the promotion of their markets and supporting market participation
- Energy co-operatives supplying energy at scale.
- Energy bill payers interested in participating in energy and flexibility markets
- Energy managers and energy auditors
- EVO subscribers
- M&V Professionals ascertaining the flexibility delivered for settlement purposes.
- Software providers of Virtual Power Plants
- Aggregators or Flexibility Service Providers participating in energy and flexibility markets

## 1.3 Fit to IPMVP

The IPMVP Core concepts already lays out four options that can be used to evaluate energy savings from energy efficiency projects, energy performance contracting or from an energy management system. When it comes to flexibility trading, baselines tend to be shorter and can be either:

1. At the whole facility level using utility meters.
2. At a retrofitted level using sub-meters using a nomination baseline approach.
3. Deemed savings approach (not IPMVP) for EVs and other smaller energy consuming items. This is a modelled approach albeit not in the traditional sense of Option D of the IPMVP.

Baselines are also considered in terms of:

1. Macro scale or explicit events such as response to a signal from a marketplace, or
2. Micro or implicit event in which energy consumption is shifted during the day to respond to an energy price signal. This is also sometimes referred to as load shifting or behind the meter optimisation

The key difference between energy baselines developed traditionally as part of an energy management system etc, is that an energy baseline in flexibility trading can be as short as 5 days or 10 days, and prior flexibility events are often removed as non-routine adjustment. In an energy management system, baseline tends to reflect the most recent full energy cycle, typically one year under a whole facility approach.

When considering placing a trade in flexibility markets, it is important to consider whether there will be a negative impact on energy prices during this trade. Often, the energy supply contract is directly between the facility owner and the energy supplier; a flexibility contract might be with a third party and therefore a scenario could occur in which there are financial gains being made via a flexibility trade but financial losses in terms of energy prices.

## 1.3 Why do flexibility?

DNOs/ DSOs/ TSOs procure flexibility (demand turn up or turn down) to manage local grid constraints, thereby deferring the need for immediate capital-intensive infrastructure upgrades. To make financial settlement of a trade, baselines provide proof of delivery of flexibility. Reasons for flexibility include:

- Matching supply to demand is increasingly becoming difficult including the over-supply of renewable energy at the wrong time of day to match demand.
- EVs, intermittent generation from Solar PV and wind means that the grid is struggling to balance both over production of energy and increased energy demand.
- Organisations would like to include flexibility in their pathway to Net Zero.
- Growing economies means that the grid is further constrained, and grid connection agreements might be restricted by DNOs
- EU directives such as Smart Readiness Indicator are also encouraging change and directing organisations to make use of buildings as Virtual Power Plants.
- There is physical limitation to the wires and transformers that made up the power gird. It is quite expensive to design these systems to operate at maximum load during peak times. It is more effective to optimise the use of the network my time and places.

According to the Carbon Trust (2025), the size and scale of the UK flexibility markets is alone is estimated to be 1 to 2 billion pounds by 2030.

## 2. Energy Markets and Flexibility Trading

This next section explains the various global markets that are currently either in development or developed when it comes to energy and flexibility trading. An energy supplier needs an energy supply licence to trade energy; an FSP does not require a licence to trade energy assets into flexibility markets albeit the landscape around this is changing.

Strategic insight is often provided by an FSP as to which are the best energy and flexibility markets to trade, and which can maximise revenue. Note: Private wire agreements or Power Purchase agreements are bi-lateral agreements between a supplier and off-taker. For purposes of simplicity, we do not discuss these, since they are very specific to the nature of the contract.

### 2.1 Energy Trading

Energy trading markets are organized sequentially, progressively narrowing the uncertainty between generation, demand, and network availability from long-term planning down to real-time operation. Energy Trading typically involves:

- Long-term procurement contracts.
- Shorter-interval spot market trading.
- Balancing mechanisms used by system operators to match real-time supply and demand.
- Market structures including wholesale, bilateral, and retail trading.

Figure 1 provides some insights into energy markets, normally under the remit of the TSO.

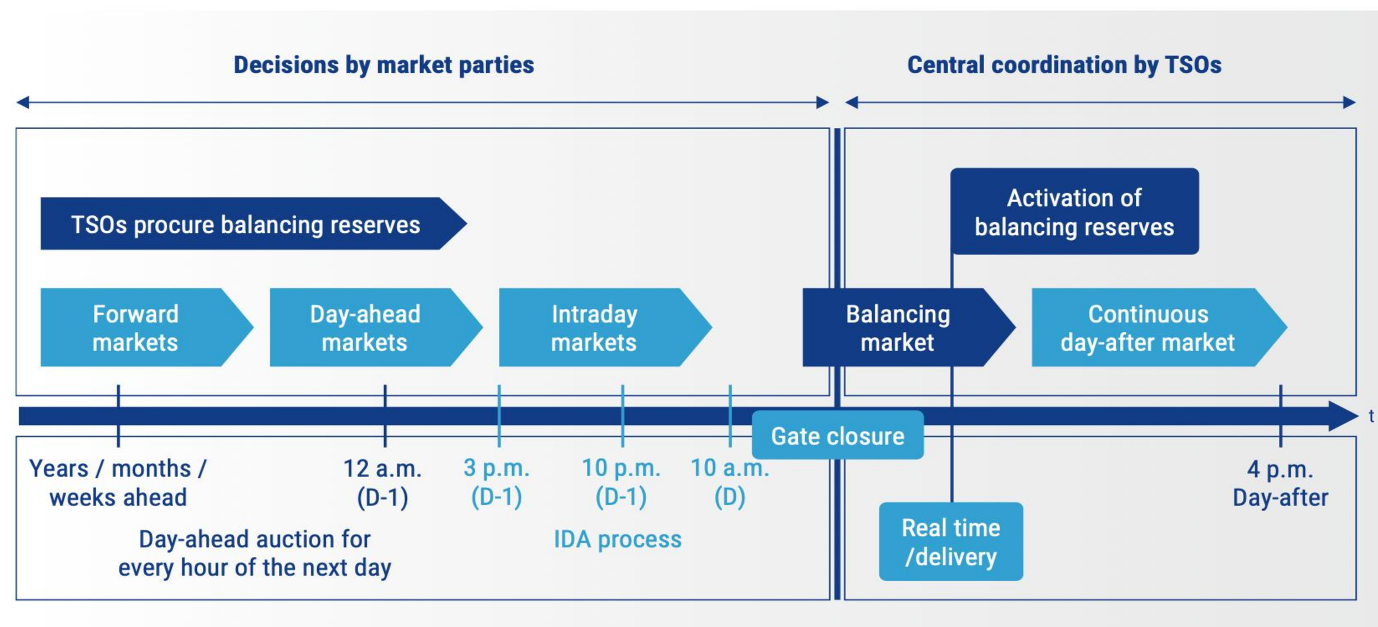



**Figure 1:** Overview of different time frames of the wholesale and balancing markets — ENTSO-E Market Report 2025

### *2.1.1 Typical energy trading markets and timeframes:*

When it comes to the value derived in energy trading, these normally are focused on the short-term (bearing in mind, the bulk of energy supply might have already been procured by an energy supplier 3 months to years ahead). These markets (limited to energy suppliers and licensed operators) include:

- **Day-Ahead Market:** One day prior to delivery (D-1), Balancing Responsible Parties (BRPs) submit hourly energy bids. The result is a binding economic dispatch schedule based on forecasted demand and the merit order of bids. Transmission System Operators (TSOs) ensure sufficient cross-border capacity is available to support these trades.
- **Intraday Market:** Following the day-ahead clearance, TSOs recalculate cross-zonal transmission capacities (e.g., IDA2 and IDA3). Power exchanges use these updates to enable trading adjustments across bidding zones up to one hour before delivery.
- **Balancing Mechanism:** After gate closure, TSOs assume responsibility for real-time supply and demand matching to keep system frequency within the synchronous operating range. Balancing Service Providers (BSPs) offer reserves to activate based on system needs.

## 2.2 Flexibility Trading

Flexibility trading leverages the capacity of Distributed Energy Resources (DERs)* and can cover:

- Demand-side response (DSR) mechanisms (turn up means using excess network energy to increase energy uses in the DER/energy assets or charging a battery; turn down means reducing demand or discharging a battery). This can be at either the DNO/DSO level or TSO level in markets such as Demand Flexibility Services.
- Participation in ancillary services (e.g., fast frequency response, reserve services, voltage response) to respond to other grid constraints such as network voltage or frequency drops. This is normally responded to within sub-second timeframes and requires battery energy storage systems directly controlled in response to these markets.

** E.g. residential EV chargers (8–10 kW), commercial HVAC systems, and industrial refrigeration.*

### *2.2.1 Role of Aggregators or FSPs*

Middleware entities, commonly referred to as aggregators, can act as facilitators gathering individual flexible demand-side. Acting as facilitators between end-users and markets / grid operators, aggregators can collect and manage the end-user's flexibility in the utilization of energy resources and combining it into a system asset. The scope can vary from participation in the provision of ancillary services, including load frequency control, which can be used by other power system entities, to the scheduling of resources while considering economic and environmental objectives. Since individual resources cannot contribute to flexibility services on their own due to limited capacity and controllability, this aggregation plays an essential role.

Once a portfolio of flexible energy resources is built, the aggregator can adopt different approaches to manage those resources to address various objectives. The main difficulties lay on the definition of appropriate regulatory frameworks and the creation of conditions for the massification of energy flexibility services (Mlecnik et al., 2020).

There are two main definitions of aggregators:

- demand aggregators - which negotiate more advantageous contracts with retailers on behalf of a large group of consumers by making the most of the scale effect
- load aggregators - which gather the load flexibility of consumers of different types, namely residential ones, or production aggregators, which group small generators as virtual power plants.

These service providers may aggregate portfolios to maximise revenue gains since certain markets have minimum entry thresholds of 1MW or 3MW etc. These entities group multiple behind-the-meter assets (e.g., residential loads) to form coordinated portfolios. They enable smaller assets to participate in markets through demand response programs offering upfront or performance-based incentives.

### *2.2.2 Grid Management*

It is important to understand voltage levels when providing flexibility services since the TSO operates at 400kW to 132 kV sub-station level; and the DSO/DNO at 132kW or below – usually, 11kV. At this level, energy assets are usually large-scale battery or DERs or commercial and industrial customers. At the domestic level, large-scale aggregation of home EV chargers, as an example, whilst at the 240/ 250 Voltage level, can have impact at the 11kV in aggregation.

## 2.3 Case studies

The variety of energy and flexibility markets is illustrated in the following global case studies.

### *2.3.1 Costa Rica:*

In Costa Rica, a tariff Structure exists for industrial/commercial sectors that segments the day into Peak, Off-Peak, and Night hours to incentivize load shifting. In addition, aligning with the National Decarbonization Plan, companies certifying an Energy Management System under ISO 50001 receive preferential rates (e.g., TMT-b Tariff).

### *2.3.2 India:*

Demand side response primarily exists as Time of Use (ToU) tariffs, with full flexibility market replication limited by current grid development - flexibility markets are still in their infancy.

### *2.3.3 United Kingdom:*

The DFS is a National Level (TSO) service introduced in Winter 2022/23, albeit local DNO level markets had become into operation a few years earlier. This service requests energy users to scale back consumption during grid constraint periods. Registered providers submit weekly baseline and settlement data via an assessment platform, with operational baselines calculated using the P376 industry methodology.

The ancillary services market, also operated by the UK National Grid ESO, where the TSO procures these services to maintain safe, efficient grid operation. Participants apply through tenders (years or day ahead), with technology eligibility dictated by service requirements. Common services are:

- Frequency Response maintains system frequency at ~50 (or 60) Hz (deviations >1% risk damage). Services operate on a second-by-second basis using fast-ramping assets like batteries, controlled by inverters.
- Capacity markets are where reliable sources of energy such as battery and demand side response measures can be provided as an alternative to power stations
- Voltage Regulation keeps voltage within 5% of standard levels (e.g., 400 kV, 275 kV, 132 kV) to ensure stability.

FURTHER COUNTRY EXAMPLES

## 2.4 Settlement of trades - considerations

Flexibility events might occur once a day, several times a day, or not at all, according to the market procurement mechanism.

### *2.4.1 Procurement contracts*

Much like energy markets, or indeed, financial markets, flexibility trading is moving towards shorter timeframes and windows. Trades can be set a day ahead and delivered on the day within a 2 - 4 hours window for call out. When a DNO/DSO/TSO models or predicts that there are many market providers that can compete, they are likely to use this mechanism.

Other contracts can be longer term in which the procurer pays for both availability and utlisation when a gird event is required. E.g. a short-term summer contract might require availability from Monday to Friday 9am to 5pm, with only say 10 actual events being called upon (usually when the weather is predicted to an at an extreme).

DERs that are traded into one market, but whose trade price is not accepted, might be placed into other energy and flexibility markets.

This is still an emerging marketplace and there is likely to be greater collaboration between the DNO/DSO and TSOs to allow multiple market stacking without causing conflicts.

### *2.4.2 Settlement of trades in multiple markets and with multiple parties*

A consequence of multiple events stacked in multiple markets and with potentially multiple parties in aggregated trading portfolios, is robust baselines to use to settle trades.

Each procurer (DSO/DNO/TSO) will have their own methods and methodology for baselines and non-performance penalty clauses. In this guide, we can only make the reader aware of the baselines being used by the parties and ask the reader to consider consequences of stacking DERs in multiple energy and flexibility markets.

Usually, allowance is made for non-routine adjustments once a flexibility event has occurred. I.e. the drop (or increase) in kW delivered can be extracted from the baseline.

Shorter baseline periods are also used as a result – such as last 5 or 10 days (longer data may need to be retained since this is the last typical 5 or 10 days. E.g. the last 5 Mondays).

Gaming is also a consideration to procurers, where a provider of Flexibility deliberately increases the baseline prior to an event. However, this might be a legitimate reason for doing so, I.e. you are over-cooling a building prior to turning down the cooling at the required delivery time. Or indeed, once you have turned down cooling for the period of half an hour (or the required delivery timescale), then you have to increase the cooling after the event, to bring the room temperature conditions back to operating normal conditions.

# 3. Implementing Flexibility

This section provides users of this guide with some practical guidance on identifying and implementing flexibility in energy and flexibility markets.

## 3.1 Identifying Flexible Loads

A flexible asset / DER is an electrical energy using plant or equipment whose use can be time-shifted or reduced without compromising a facility's needs. The type of flexible loads depends on the facility type e.g. residential, commercial or industrial, hospital, data centre, leisure centre etc. It is common for facilities, especially residential buildings, to be aggregated into the total demand of a set of facilities with flexible loads.

Typical DERs and assets might include:

- Residential – EV chargers and home battery energy storage devices. To have impact at an 11kV level or above, then series of EV chargers can be aggregated. These assets are usually maintained and controlled by energy suppliers that can place strategic tariffs to encourage users of EV chargers to either charge or stop charging.
- Commercial offices – Typical DERs are AHUs, electric heating and cooling, refrigeration, chillers, pumps and motors (non-critical to a facilities operation) etc.
- Industrial – large-scale industrial plant, HVAC, chillers, cooling, electric heating, etc.
- Hospitals – back-up generators (that need to be run for maintenance purposes, can be run at the same time as a flexibility event), battery energy storage, non-critical pumps/motors/ AHUs and HVAC.
- Leisure centres – Heating and cooling, AHUs, battery energy storage, ASHP, WSHP etc.

The following are suggested good sources of information that can be utilised for identify flexible loads for trading DERs in energy and flexibility markets.

a) Half hourly data will support realistic assumptions around flexible loads
b) Loads to be analysed at the summer and winter peaks should include recovery times
c) Using weather files, the seasonal impact of flexibility can be established.
d) Design philosophies, nameplate ratings, meter drawings, building data, AutoCAD drawings, O&M, plant schematics, metering schematics, operational control philosophies can be used.
e) Modelling assumptions or building thermal models.
f) Information on Building management systems
g) Assessment of monitoring and controls needed or additional metering
h) Review of any battery energy management systems or other renewable energy systems and their control philosophy.

Sources of flexibility should be non-critical to the building operations E.g. Battery, AHUs, refrigeration, thermal energy storage etc. Energy assets in this context are the various equipment, systems, and resources which are owned or accessed by end-users to produce, consume, or store energy that can be used for flexibility trading.

It's important to understand that flexibility can be:

1. Turning down non-critical items or equipment so that energy consumption is reduced. E.g., a DNO-procured service for 200kW of flexibility might comprise of 120kW of battery used behind the meter, and 80kW of equipment being turn off at the same time. This drop will be measured against a baseline or against a metered usage.
2. Turning up demand through the absorption of excess energy from the grid. I.e. heating the pool a little warmer or charging the battery when there is excess grid energy.
3. Using battery 'behind the meter' so that grid generated energy can be offset with generation at site. This is also sometimes referred to as load shifting where the battery is charged overnight at a cheaper tariff and used during the day when it is more expensive.
4. Using battery 'in front of the meter' for more pooled or wider balancing reasons.
5. Exporting to the grid, energy generation from battery where export licences exist. This may require connection upgrades to be made.
6. Demand Peak response to avoid costly peak demand charges.

## 3.2 Flexibility metrics

To place a trade in flexibility markets, it is important to understand key flexibility metrics. Many DERs and energy assets could be assessed for the potential to trade, based on these criteria, as part of an energy audit of a facility. This would enable facilities to become VPPs.

1. **kW Flexibility:** The amount, in kW, that can be provided as flexibility e.g. demand turn up or demand turn down. This value may change over seasons, for example, an ASHP might have a different coefficient of performance over the summer compared to the winter.
2. **Response time** – how long do need to respond to a market signal (day ahead, 4 hours, two hours etc.).
3. **Run rate** – how long the flexibility will need to be run for. The run rate is determined by how long an energy asset can be turned off, or indeed, generate energy onto the grid. E.g. a battery might have a one hour run rate, an Air Handling Unit might have half an hour. This all depends on the size of the battery (can be longer run rates), and the criticality of the supply. An Air handling unit might be calculated to be able to be turned off for only half an hour (usually the minimum in a market), before the Carbon Dioxide conditions erode and does not meet building air quality standard.

4. **Recovery time** – how long does the facility have to recover to normal building operations before the next flexibility event can be run. The recovery time will determine time periods between one flexibility event, and the ability of another. E.g. a battery might need to be re-charged overnight with a cheaper tariff or be charged up directly by Solar PV if connected to such a system. An AHU might need say a period of 4 hours to bring the building up to normal operating conditions. During this time, the energy asset is unavailable for trading. Beware of the bounce back effect – which is the additional make-up energy used to bring the facility back to normal operating conditions.

## 3.3 Revenue Potential

To determine the revenue potential, consider:

- The availability of energy markets (are you in a grid constraint zone)
- How many flex events are expected in the period the markets are open
- What are the market prices and number of competitors
- The development of a strategy that allows stacking without compromising delivery.
- Market deployment strategies to deploy into the next market, if an asset is not fulfilled.
- Choosing the markets are most attractive (local or national) since it might not be possible to enter into multiple markets at the same time
- The impact on the baseline if entering multiple markets - baselines might be reduced in one market following a flexibility event and therefore when averaging the last 5 days, this baseline will be lower than usual.
- Your ability to track baseline changes if an asset is fulfilled, and to ensure that flexibility can be delivered against baseline.

## 3.4 Trading in flexibility markets

To trade in a flexibility market, the procurer normally will ask for pre-qualification criteria to be achieved (hence an aggregator is normally used). There must also be a process in place, either manual or automatic, to respond to a market call. If operating in the Fast Frequency Response sub-second market, this will need to be a direct signal directly to the frequency inverter of a battery, since response time needed is sub-second. Otherwise, a manual process can also work if there are site staff available to turn off equipment (potentially turn up) in the timeframe needed.

Accompanying energy data is needed to demonstrate the kW drop or turn up during the flexibility event for settlement.

## 3.5 Summary of Flexibility Approach

To summarise, a good place to start is in the assessment of non-critical loads, and their ability to be controlled / responded to for a flexibility event call-up. The assessment of loads on site, could be as part of an energy audit. Once the loads are identified, assess their turn up / turn down potential, any seasonal variations, and an estimate of the total site kW potential. Some assumptions might need to be made on use to provide estimates of "High", "Medium", and "Low" flexibility over the seasons, as well as run rate and recovery times. Once these metrics have been identified, either work directly with the energy markets, or use the power of aggregation, an aggregator or FSP. Figure 2 below depicts this process.

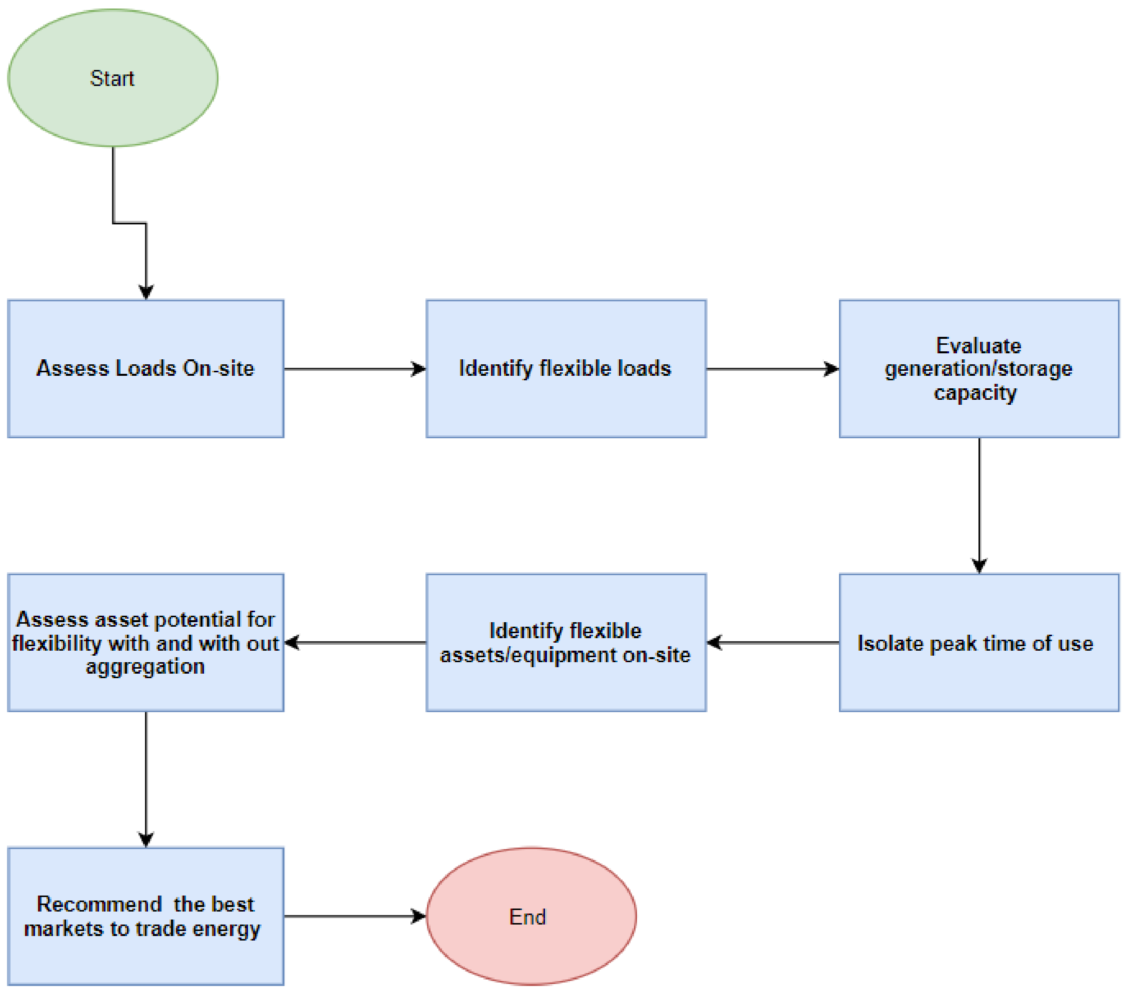



Figure 2: A typical process to apply for identifying flexibility potential for trading.

## 3.6 Integrating Flexibility studies with Energy Audits and ISO standards

An energy audit is a structured, systematic process used to evaluate energy consumption and patterns within an organization, aiming to identify opportunities for improved energy efficiency and reduced operational costs.

EN16247 – part 1 and ISO 50002 are international standards for Energy audits. They follow the ISO 50001:2018 standard that establishes the framework for an Energy Management System (EnMS), enabling organizations to monitor, measure, and continuously improve their energy performance using Energy Performance Indicators (EnPIs) and Energy Baselines (EnBs).

ISO 50002 complements this by providing practical guidelines for conducting energy audits, helping organizations identify, prioritize, and track opportunities for energy savings as part of their EnMS.

ISO 50047 further supports these efforts by outlining methodologies for energy savings determination. It helps organizations ensure that energy savings are calculated correctly, without double counting or errors, while integrating both operational and design improvements.

Together, these standards allow for comprehensive planning, monitoring, and verification, which is crucial for managing energy performance effectively and participating in energy markets and flexibility programs. These standards can support in:

1. **Assessment of Flexible Loads during an Energy Audit**: An energy audit involves evaluating non-critical loads that can participate in demand response activities. During the data collection phase, facilities must identify flexible loads—such as HVAC systems or lighting—that can be adjusted to respond to energy market signals without disrupting operations.
2. **Measurement and Verification of Flexibility Impact:** The ISO standards help integrate energy performance metrics (like EnPIs) with demand response actions. An energy audit ensures that baselines and indicators accurately reflect changes in consumption when flexibility events occur. This avoids double-counting savings and ensures reliable tracking of the impact of energy trading or demand-side management activities.
3. **Energy Savings Validation and Double Counting Prevention**: ISO 50047 further supports this framework by providing guidelines on determining and validating energy savings in a way that avoids double counting. This is particularly critical when flexibility activities impact energy use, as it ensures that each improvement action is measured and accounted for correctly. ISO 50047 helps organizations to consistently and accurately report energy savings while participating in demand response and energy trading, providing a clear and reliable picture of the energy performance improvements.
4. **Planning and Control through an EnMS:** Within the framework of ISO 50001, the creation of an Energy Management System (EnMS) involves continuous monitoring, which supports

effective participation in flexibility programs. The identification of run times, recovery periods, and response times for energy assets is integrated into the operational planning processes outlined during an energy audit.

## 3.7 Case studies

The following case studies are provided by committee members and provided a breadth of processes and markets that their projects touched.

### *3.7.1 Belesco (Belgium) Commercial offices in Portugal*

The flexibility was derived from Energy Efficiency Measures - PV (existing and new), LED lighting, installation of variable speed drives, demand response - standby optimisation of ventilation units, smart heating and cooling. Flexibility was based on active control and using ToU tariffs (4 periods: Peak, Shoulder, Normal Off Peak and Super Off-Peak hours).

Electricity baselines were calculated using the ABEPeM tool (Active Building Energy Performance Modelling platform), which has a so-called flex model to create a baseline model and a reference model. They were checked/calibrated against 5-year historical measured data (from the meters).

Market driver was using extending existing Energy Performance Contracting business models for higher cost and CO2 savings, allowing customer projects to become more efficient and performant (incl. better TCO) and ESCOs to become flexibility aggregators and thus develop new business, using Time-of-Use (ToU) dynamic pricing.

Energy savings were calculated using standard national C02-intensity parameters. Pricing did not yet contain direct CO2-based components but only indirect ones (over the 4 time periods).

€ savings were +/- 8% from EE measures and an additional +/- 6% from Demand Response (2% by optimizing heating and cooling operation, while maintaining comfort requirements).

The unique feature was the integration into EPC models (with performance guarantees) turning it into an Active building Energy Performance Contracting or AEPC model with reward-penalty mechanism. This is an example of flexibility being used implicitly – without responding to a market signal from a DNO/DSO/TSO, but internal changes and controls to optimise revenue savings based on energy prices and tariffs.

### *3.7.2 Cold Chain food storage (Mexico)*

This was a cold chain storage project where demand peak response and shifting the peak demand was primary purpose. Since demand charges were not so high, savings were on energy consumption. An ESCO style contract with incentives built in rates - Lower demand on peak hours with a recovery time for temperature being 2 hours.

Flexible load identification - critical equipment was not touched but HVAC bearing in mind a 4-hour demand response event. Facility therefore must have enough thermal cooling capacity. The supermarkets had comfort policy of 24 deg C, so not all HVAC was used. In total the following flexibility with seasonality was identified:

- Example 1 -Processing plant refrigeration -The maximum kW flexibility was either 247 kW or 350 kW depending on which compressor was turned off. Summer: run rate 2 hours, recovery time 0.5 hrs. Winter: run rate: 4 hours, recovery time 0.75 hrs.
- Example 2 - Distribution center refrigeration - The average kW flexibility was 148 kW in the summer and 232 kW in the winter. Summer: run rate 2 hours, recovery time 1.5 hrs. Winter: run rate: 4 hours, recovery time 1.5 hours.
- Example 3 - Supermarket – HVAC -The average kW flexibility was 86 kW both in summer and winter. Summer: run rate was 2 hours. Winter 4 hours. Recovery time doesn't apply because the store closes.

Flexibility was derived by initial estimates based on nominal power and the fraction of equipment that were to be turned off simultaneously during flexibility. Flexibility was expected to operate only Monday to Friday at peak hours: summer 8 to 10 pm (2 per day) and winter 6 to 10 pm (4 per day).

The following baselines were derived:

- Example 1: Sub-meter measurements for two compressors were taken for 5 days to determine the maximum demand and average consumption on each time-of-use-period. For the demand baseline, maximum demand of the compressor was measured and added it to the metered maximum demand on peak hours during that month reported by the utility. Verification of which compressor was actually turned off was made using ON/OFF status data. For the energy baseline, this was the total time the compressors were turned off by the system during the month multiplied by the average consumption. The result was adjusted based on the difference between the CDD modeled total energy consumption during the reporting month and when the initial measurements were made (e.g. if we measured in the summer, we adjusted down for winter months).
- Example 2: Similarly to example 1, retrofit isolated measurements for each condensing unit (22 units) were made over a one day. For the demand baseline, 15 min periods were used, at which the maximum monthly demand. This was added to the measured maximum demand of every condensing unit that was turned off during that period and to the metered maximum on-peak monthly demand reported by the utility.
- Example 3: Fixed measurements of aggregate HVAC energy consumption (all packaged units) were observed before and after demand response began. The energy baseline was determined by the previous year’s HVAC energy consumption data.

The main incentive was financial savings by reducing the retail rate demand charge for capacity that is based on the maximum monthly demand on peak hours. Rates were 17 to 20 £$ per kW. The energy savings are also considered at the peak hours rate of 0.1 £$ per kWh. Savings were also reported using the annual national carbon emission factor 0.423 $tCO_2e$ / MWh. (Carbon savings from demand response project were low compared to the demand charges).

In total, the following savings were achieved:

- Example 1: 3,500 kW/year (maximum demand reduction) - 300,000 kWh/year (energy savings) – 120 $tCO_2e$ - £ $75,000/yea
- Example 2: 2,250 kW/year (maximum demand reduction) - 200,000 kWh/year (energy savings) - 80 $tCO_2e$ - £ 50,000/year
- Example 3: 1,000 kW/year (maximum demand reduction) – 100,000 kWh/year (energy savings) – 40 $tCO_2e$ - £ $25,000/year

Note: This time-of-use retail rate with demand charges applies to most end-users that have a demand of at least 100 kW. The utility (CFE) also participates in the wholesale markets as a generator (used to be a public monopoly) and controls transmission and distribution.

### *3.7.3 BankEnergi (Consortio Limited) and a UK City Council*

BankEnergi was contracted to undertake a study on 5 leisure centres with a City Council.

The aim of the study was to identify and understand flexible assets, costs entailed to make them tradeable in flexibility markets (local constraint and National markets), and what if any were the limitations and to ultimately provide the potential revenue that can be generated during different times of the year.

Such opportunities include:

1. Peer to peer trading on capacity headroom

2. DNO/DSO -procured flexibility (the local distribution network operator, SSEN procuring flexibility during a grid constraint event to prevent a brown out)

3. National energy markets (a minimum aggregation size in a Balancing Mechanism of 1MW is needed)

The further aim of the study is to understand the recovery times of these flexible assets in the winter and summer peaks, given a 1 degree drop in temperatures. Whilst most facilities have wet and dry side (indoor pools), there was an outdoor pool that only operates from April to Sept and the Ice Rink is an Olympic standard ice rink with long operating hours late into the night (2 am on occasion) and a 6am start.

The flexibility that was identified was:

1. Ice rink (chillers 65 – 102kW) and AHUs (32-51kW).
2. Outdoor pool (April to Sept only) Battery 120kWe, WSHP 120- 161 kW.
3. Leisure centre 1 with wetside –Battery 120 kWe, AHUs 47-70kW, ASHP 19-28 kW.
4. Leisure centre 2 with wetside - Battery 300 kWe. AC units 7kW, ASHP/AHU combined 33-77kW.
5. Leisure centre 3 with wetside – Battery 180 kWe, AHUs 96- 118 kW. ASHP 88-108 kW.

kW flexibility was determined through thermal models using IES software and Transys for the pools and ice rink. With these models it was possible to determine the following run rates and recovery times:

1. Ice rink (30/180 mins chillers, 120/60 AHUs).
2. Outdoor pool – battery 60/90 mins. WSHP 120/300 mins.
3. Leisure Centre 1 - battery 60/90 mins, AHUs 120/90 mins, ASHP 120/330 mins.
4. Leisure centre 2 - battery 60/90 mins, AC units 60/90 mins. ASHP/AHU 120/90 mins.
5. Leisure centre 3- battery 60/90mins, AHUs 120/90, ASHP 120/90.

Baselines were not really considered. The kW flexibility would have been noted against submeters as an explicit to a signal from the Market Platform of the DNO. Gaming was not considered, nor on the radar of the Council in the questions they asked for the consultancy work.

Due to the range of possible call up events, the total revenue for 30 events/ 30 hours were estimated to be £57k. Up to a max number of events of 150 events/ 150 hours of £140k per annum.

Flexibility revenues could only be derived when the flexibility markets were open. There was not enough aggregation to enter National markets at 1MW since the sites were split across 3 Grid supply points.

### *3.7.4 Bamboo Energy – Large Communities (residential)*

The aim of their study was to analyze a new use case where flexibility can play an important role in energy communities. An energy community is a legal entity that empowers citizens, small businesses and local authorities to produce, manage and consume their own energy.

Case studies were implemented in Spain, Netherlands, Sweden and Greece in the framework of commercial and innovation projects. Flexibility can help the sustainability of energy communities by:

1. Reducing costs performing prices arbitrage
2. Reducing costs maximizing (collective) self-consumption
3. Reducing costs reducing power peaks
4. Increase benefits by allowing their participation in balancing and local flexibility markets

At the end, energy communities are a particular use case of what flexibility can do and can be extrapolated to individual use case. The advantage of energy communities is that their value is socialized, and more assets can be aggregated in a unique fronter point. In addition, energy community members used to be innovative and sustainable.

Flexibility (kW) was derived, as the sum of all flexible assets connected to the platform (industries, heat pumps, batteries and electric vehicle). Assets are still being connected. Seasonal variability of flexibility was not considered. The platform calculates the available flexibility in real time considering flexible asset consumption and constraints, among other variables.

Run rate and recovery time depends on the asset and user's preferences in their software platform. E.g. for batteries it depends on Power/Energy ratio and has no recovery time. Thermal loads depend on the acceptable comfort range and usually the recovery time is equal to the delivery period.

For explicit demand response, baselines are provided by the TSO/DSO. Bamboo also calculates its own baseline and uses it to optimize flexible activation keeping into account the DSO/TSO expected baseline.

Bounce back is also considered in the optimization algorithm and depends on the asset and user's preferences.

In Spain, energy prices are the driver. In other projects the drivers are balancing markets or local flexibility markets.

Usually, Bamboo provides an optimization based on all these drivers to maximize the value of the available flex.

Carbon savings are considered for explicit demand response as that we are substituting gas power plants (traditional flex providers), while for implicit demand response we keep into account the national energy mix.

Bamboo does not consider that flex activation implies a reduction in energy consumption, but a shift of them. Thanks to this shift, an industry of 500 kW flex can save about 5-10 % of its energy bill and generate between 50 and 100 k€ in flexibility markets.

### 3.7.5 SSE CHOP project description

The CHOP Project stands for Combined Heat and Power Optimization. This project is designed and executed by the SSE Digital Services team for SSE Heat Networks to optimize the heat network assets like Combined Heat and Power Plants (CHPs).

The project aims to optimize energy assets specifically CHPs to lower energy costs and reduce net carbon grid intensity.

Hardware is installed on the CHPs to upgrade its performance to remotely access operation and performance data. The process also involved the development of a digital twin through the application of the latest machine learning techniques to achieve a good quality data flow between the physical and digital version.

Through this, the new digitally optimized operating schedules of the CHPs are created with the objective function of reducing costs and lowering net carbon grid intensity. This optimized schedule is implemented on- site to improve operations.

The kW flex is converted to cost savings and is estimated based on digital twins as £350,000 per year for all 4 sites. The CHPs run at the following two schedules: 07:00 – 00:00, and 16:00 – 00:00.

The baseline was calculated using the historical generation and demand data of the operating CHP. This real data is used against the digitally optimized version to create dynamic week-ahead operating schedules.

The market drivers are mainly the historic and forecasted day ahead pricing and grid constraints specifically grid carbon content. These are included as inputs to the optimization to minimize site operation costs and carbon.

Carbon savings are calculated by comparing the site carbon content for the digitally optimized schedule with the manual schedule. Digital twin based on hourly calcs. Carbon emissions factor, based on real-time values from National Grid website (API call).

Projected savings of £350,000 per year across all 4 sites.

### 3.7.6 Demand Flexibility Service - Retail sector

The UK National grid introduced DFS to reduce the demand at certain times during winter when commercial actions are insufficient to cover demand and upwards margins.

SSE Energy Solutions serves as a provider to deliver DFS. Last winter, we had several customers participate in the DFS where the customer is required to scale back their consumption when the energy supply is anticipated to be strained to alleviate pressure and maintain stability on the grid.

This case study summarizes the outcome from one of our department store customers over a 1-hour participation.

The DFS methodology is an explicit process where baseline methodology strictly follows ESO standards. The participant delivered the DFS by turning off non-essential systems, mainly HVAC for the duration of the DFS. No site energy generation, the site was fully dependent on import.

This example includes a total of three sites, each with 2 import MPANs, and a DFS participation of two 30 minutes intervals a day.

This was under the solutions team, not the distribution side of the business.

A total of 865.5 kW was delivered to the grid for a duration of 1 hour (from 17:30 to 18:30).

Participation in this service is solely during the winter hence no seasonal variation is considered.

The participant delivered the DFS by turning off non-essential systems, mainly HVAC for the DFS duration. No site energy generation, fully dependent on import. Reduction (settlement) is calculated by measuring actual demand relative to a baseline. The difference between the baseline and the actual is credited as the delivery.

The delivery period was from 17:30 to 18:30 with a 30-minute participation window.

The baselines were explicit (responding to an external signal). The process strictly adheres to the guidelines and standards published by the ESO for selecting baseline and submitting settlement data. The baseline import meter read is taken a week before the DFS participation window to calculate the operational baseline according to the P376 industry methodology.

Providers submit customer Unit Meter Points daily or when there are changes to the portfolio, to the ESO via an API.

### *3.7.7 Demand shifting only*

End-users can also benefit from time-shifting their energy consumption to avoid higher energy prices during peak hours, that can be at more expensive rates, as defined by their electricity supplier. Sometimes demand charges are also specific to peak hours which may drive consumers to reduce their demand more during that time. This can include shifting to time periods where more locally generated energy is being supplied – or within local energy markets.

A primary incentive therefore for end users is to change their demand to offset energy bills by using a different price structure that is cheaper during specific time periods in the day. This is sometimes known as load shifting where energy consumption usually remains the same volume1, but energy is used at a different time of the day.

In the UK, the energy prices are determined by adding up the flat tariff and the Distribution Use of System Charge (DUoS). The DUoS charge varies by time of day and hence, are divided into three zone; Green zone being the cheapest, followed by the Amber zone, and then the Red zone being the most expensive. The green zone starts from midnight and ends at 6:30 am. This is followed by the amber zone which starts from 7:00 am up to 15:30 and the start again at 19:00 and ends at 23:30, and finally the Red zone starts from 16:00 ends at 18:30. These charges influence the electricity rates and hence, motivate energy users to render services like the DSR.

A flex service being provided by a BESS can rely on the electricity rates to charge the BESS or use the Heat Pump at times of low tariff and delivery flex services to the grid when the trading prices become high. The process of buying energy at cheaper rates and storing it for future usage when the tariff is high is known as an Energy arbitrage process.

Price adjustments from baseline period to flex delivery period might be required to determine the cost savings for any flex provided depending on the contract agreement for the flexibility service being provided.

Occasionally, in Western Europe, we have experienced the grid being supplied by 100% renewables and therefore the energy price has been incentivized for users to take energy from the grid. E.g. during the day-time, due to the higher penetration of renewables, the spot price could be significantly lower, which makes the overall electricity price, including taxes, lower compared to non-peak periods.

[1] In case of electric water heaters, there can be a shift but if that shift is to prior hours, then losses can be bigger and the total energy consumed increases.

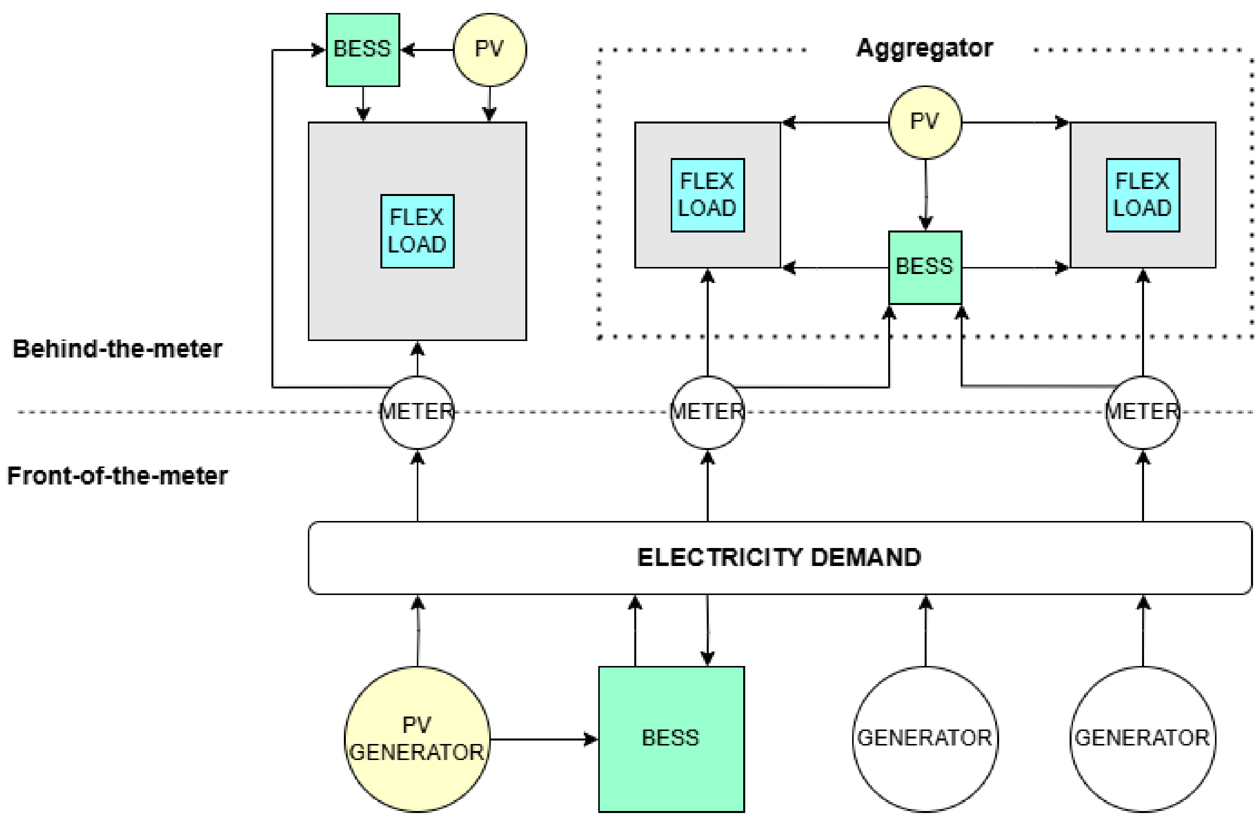
# 4. Measurement and Verification Approaches

There are two main perspectives when it comes to energy baselines for demand response during flexibility events:

1) **Implicit baselines** are when end-users adjust DERs to optimise energy usage according to energy prices. Sometimes this classification is also known as 'behind the meter'', indicating that the activity is designed to improve efficiency and reduce internal operational costs, e.g. energy costs. These changes are not in response to a DNO/DSO/TSO flexibility event where turn down or turn is measured.

2) **Explicit baselines** are in response to a market driver or signal from a DNO/DSO/TSO where the facility owner will benefit from revenues for providing flexibility. This is sometimes known as 'in front of the meter' flexibility that impacts sub-stations and supply at 11kV or above (note: aggregated residential flexibility also falls into this category).

Figure 3 highlights the difference of behind-the-meter and front-of the meter flexibility assets. Individual facilities are represented in grey boxes, and each one has its own electricity meter. Flexibility assets can be generalized into the concept of flexible loads and battery energy storage systems, represented in blue boxes and green boxes respectively.



The financial value of flexibility is determined by the market mechanism:

- Implicit (Cost Avoidance): Focuses on shifting load from high-tariff periods.
  - Calculation: (Baseline Cost) - (Actual Cost)
- Explicit (Revenue Generation): Focuses on responding to external market signals.
  - Calculation: (Delivered Flexibility kW) × (Duration) × (Contract Price)

Following a flexibility event, it might be necessary to also evaluate the increase in energy use or consumption post flexibility event to understand the full impact of energy and demand savings on the facility level. This is also known as the rebound or bounce back effect.

Implicit baselines would traditionally use IPMVP Options A to D to evaluate savings from behind the meter events. For the purposes of this guide, we focus mainly on **explicit baselines**.

## 4.1 Standard Baselining Methodologies for Explicit events

For this section we reference the Flexible Market Rule -Standard Baseline Methodology, draft 2025, designed to go live 1 July 2026. The document provides UK industry with a means to define baselines, and this is being adopted by the DNO/DSOs/TSOs in the UK.

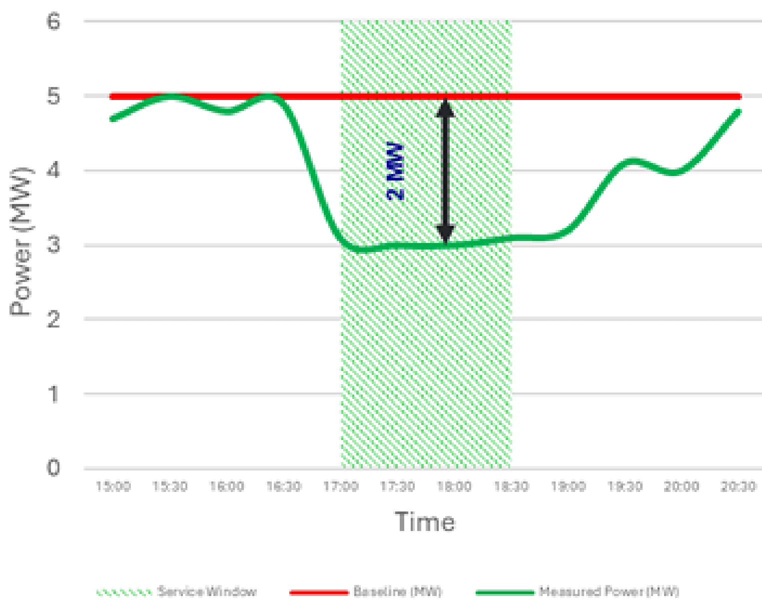
### *4.1.1. Core Baselining Methodologies*

The framework defines five distinct methods to calculate the baseline (what energy usage would have been) versus actual performance during an event:

1. **Asset Capacity**: Assumes the participating DER operates at its full rated capacity during the flexibility event, irrespective of actual operational conditions.
2. **Fixed Reference**: Uses a fixed daily profile derived from the observed energy behaviour of a representative group of DERs that represents normal operating behaviour or a DNO’s planning scenarios.
3. **Nomination baseline**: Uses an estimated/expected set of meter values, provided by the Flexibility Services Provider (FSP), based on the expected or normal operation of the site / DER without any flexibility event, for a future availability window that then becomes a utilisation event.
4. **Recent History**: Uses a set of recent, eligible days (i.e. excluding Event Days such as a day providing a flexibility service) preceding a utilisation event. This may also include correction factors such as within-day adjustments for weather variables or removal of outlier values.
5. **Zero**: The baseline energy consumption or generation is set to zero, meaning no energy usage or production is assumed in the absence of the utilisation event.

### 4.1.1.1 Examples (taken directly from FMR-SBM 2025 draft)

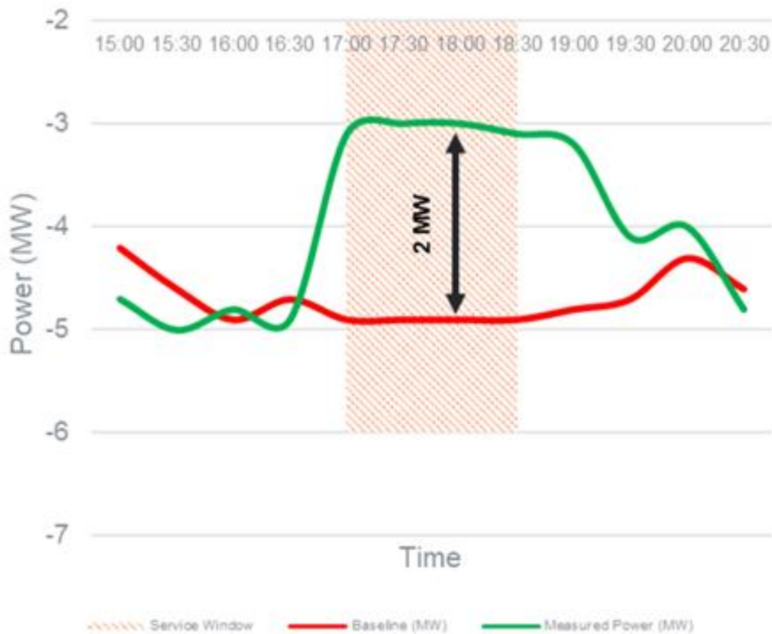
1. **Asset Capacity**: This methodology assumes the participating DER operates at its full rated capacity (no de-rating) during the flexibility event, irrespective of actual operational conditions. Worked examples of these principles are shown below for a Schedulable Generation DER which has a full rated export capacity of 5MW. In the example, the Generation DER has a Generation Turn Down contract for 2MW. The DER turns down generation from 5MW to 3MW and is paid for 2MW as the baseline was set at 5MW (i.e. the full rated capacity) and the measured power during the Service Window was 3MW.




Asset name plate rating can typically be found on the motor itself of the operating device, or it the case of a battery, its kWh rated capacity. Care should be taken that the asset was operating at its full load conditions, or the asset capacity baseline will not be a true reflection of kW flexibility provided. I.e. an Air Source Heat Pump might be operating at half capacity in the summer. This methos is akin to a deemed savings approach.

**2. Fixed Reference**

This methodology uses a fixed daily profile derived from the energy behaviour of a representative group of DERs that represents normal operating behaviour or from a DNO's planning scenarios. In the example shown below, the baseline value represents a reference, non-smart charging behaviour for a large group of domestic EV charging points. It is this representative behaviour that is used to set the baseline value of -5MW to assess the level of Demand Turn Down provided by the DER. The measured output during the Service Window is -3MW, therefore it is deemed to have delivered a dispatched capacity of 2MW.




Typical examples would be an energy supplier providing flexibility from a residential area. The flexibility potential has been assessed using modelling and historic behavior consumption patterns. This method is akin to deemed savings approach.

### 3. Nomination baseline

This methodology uses a proposed estimated/expected set of meter values/data, provided by the FSP, based on the expected or normal operation of the site / DER without any flexibility event, for a future availability window that then becomes a utilisation event.

Nomination values may differ over the Service Window or be a single value that is applied for the whole Service Window. As nomination baseline methodologies rely on FSPs to propose the values used, there will be accuracy standards that the DNO requires the FSP's nominations to adhere to, to ensure the nominated values accurately reflect normal operating behaviour.

In the example shown below, a variable nomination value has been proposed by the FSP (-5MW for the earlier part of the Service Window and -4MW for the later part). In this scenario, the DER is paid for a dispatched capacity of 2MW of Demand Turn Down for the first part of the Service Window and a dispatched capacity of 1MW for the latter part.


The above figure illustrates the variable delivery that has been nominated by an FSP.

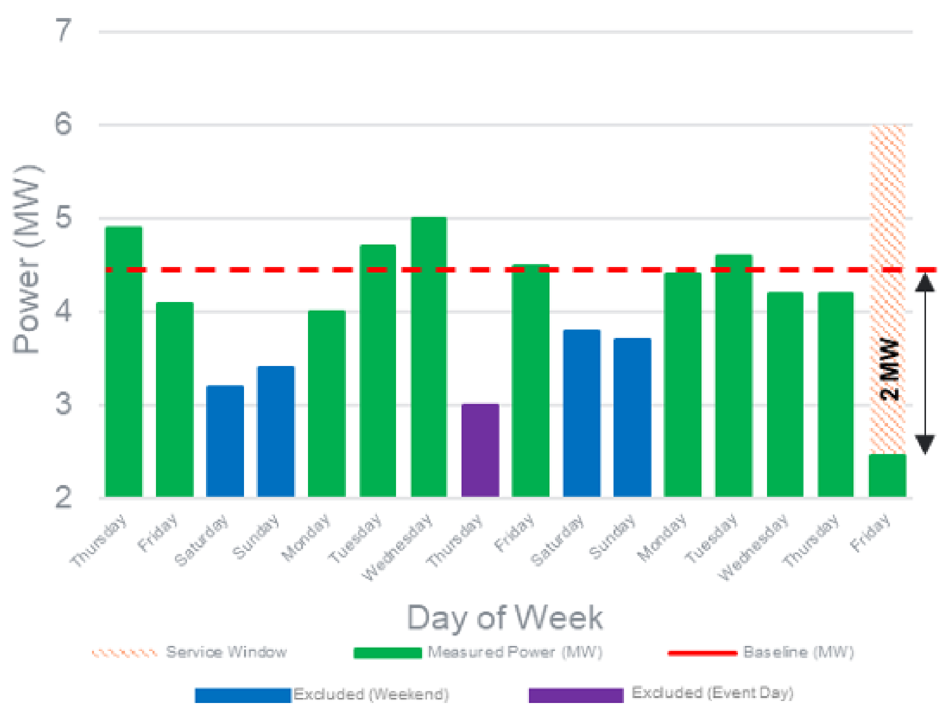
## 4. Recent History

This methodology uses a set of recent, eligible days preceding a utilisation event and will complete an assessment based on the same time of day (i.e. 30-minute window) and type of day (i.e. working day or non-working day). This could be the last 5 or 10 similar days (e.g. the last 5 Mondays). This is the more typical methodology used and is based on actual meter data. This can be under IPMVP Options A, B, or C.

It is the most used mechanism that is also used by the Balancing Settlement Code (BSC) for the Balancing Mechanism in the wholesale market and includes within-day adjustment. It is also used by the UK's TSO (NESO) for Demand Flexibility Service (DFS) and also for the Capacity Market for Demand Side Response (DSR) Capacity Market Units (CMUs).

Within day correction factors can include weather variables or removal of outlier values (typically defined as Event Days). Event Days can include activities such as delivery of Balancing Services, Virtual Trading, Site Outage, Equipment and/or Site Disconnection.

In the example shown below, a DER baseline of 4.5MW for a Service Window on a Friday is set based on the mean average of the 10 most recent working days, which weren't Event Days. Weekends and the Purple coloured Thursday (which met the criteria to be defined as an Event Day) are therefore excluded from the calculation. The measured power during the Service Window was 2.5MW of generation, and therefore, the DER is paid for a dispatched capacity of 2MW of Generation Turn Down.



A note on adjustment factors:

Flexibility baselines could be affected by changes to the values of independent variables and static factors. Usually, the 5- or 10-days baseline period does not necessitate adjustments to be made, however, it is important to consider if independent variables and static factors could significantly affect a baseline, especially if the flexible energy asset performance itself is impacted by either. If there are no influencing factors, then measurements of a baseline as a counterfactual and then the metered energy during a flexibility event, takes on the IPMVP principle of adjacent measurements, where no adjustments are made.

However, if the DER is impacted by the weather, then it may deliver less apparent flexibility if consideration is not given. This could be an ASHP, whose electrical efficiency changes during winter and summer. The kW flexibility should be considered based on CoP of the asset (which varies over the seasons) to provide more accurate forecasting. Extreme weather events during the baseline period (an outlier is plus or minus 3 standard deviations), can be a candidate for removal.

Other factors that can change during the nomination baseline period are static factors. These are non-routine events; and typically, this can be a flexibility event. When there are numerous markets being traded (e.g. at DNO/ DSO level, as well as TSO level), adjusting the baseline for static factor events can become an issue.

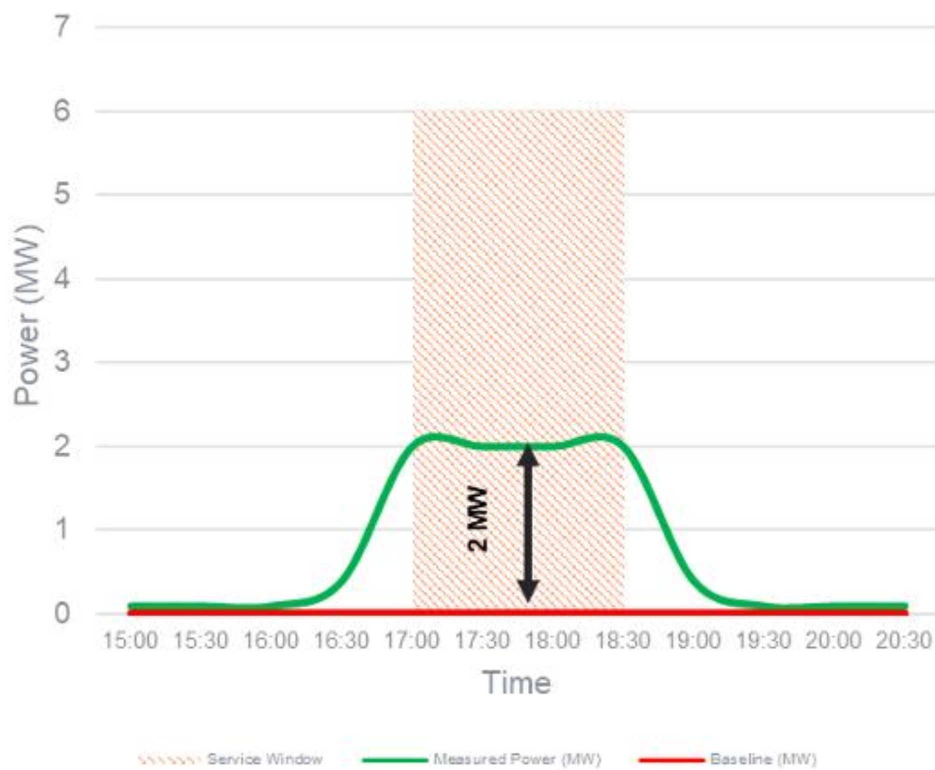
Other static factors can include the facility or generation size, the design, and operation of installed equipment, the production and occupancy level. Most of the information on static factors can be derived from facility logs, or EMS systems. For use in the development of baseline, only static factors that incur major changes in the preceding baseline period need to be identified, and evidenced in order to make non-routine adjustments.

To avoid gaming (please see section 4.2), the baseline period should ideally represent the typical operating mode of the facility, equipment or system where there is no flexibility event. I.e. this should be within the normal operating period immediately before the flexibility event happened.

5. **Zero baseline**:

The baseline energy consumption or generation for this methodology is set to zero, meaning no energy usage or production is assumed in the absence of the utilisation event. In the example shown below, the DER is deemed to have provided a dispatched capacity of 2MW through a Generation Turn Up service as the measured output during the Service Window was 2MW.

Typically, this could be a battery energy storage system that is fully charged and waiting to be discharged to provide flexibility turndown. Or, vice versa, a battery that is fully discharged, waiting to be charged up during a demand turn up event.



### 4.1.2 Typical DERs and Baselines to consider.

In the next section we align the typical baselines used in flexibility markets to that familiar with the IPMVP. It is important to note that there are two distinct sectors: residential and commercial and industrial when it comes to trading DERs in energy and flexibility markets. Hence, the deemed approach taken for smaller EV chargers of rated capacity, e.g. 7kW, but when aggregated in larger residential areas can amount to MW delivery of flexibility.

The following table groups DERs as examples, and the typical baseline/ IPMVP options.

| ***Sector*** | ***Typical assets*** | ***Typical baseline / IPMVP options*** | ***Example of effect of flexibility*** |
|---|---|---|---|
| Cold generation e.g. in Supermarkets and food industry.<br><br>Cold chambers and distribution chains.<br><br>Food, chemical or industry. Plants and flower industry. | Fridges<br><br>Freezers | Recent history/ nomination.<br><br>Option B, all parameter measurement. | Changes in the temperature of the assets can affect the energy consumption. Consider also the effect of an increase in energy use and consumption post flexibility event.<br><br>This is otherwise known as rebound effect. |
| HVAC in commercial offices. | ASHP<br><br>Electric heating or cooling | Recent history/ nomination.<br><br>Option C or Option B. | The temperature of the building changes.<br><br>Consider rebound effect.<br><br>Consider if carbon savings are being made if the carbon grid intensity is lower during time of actual use, and higher during the flexibility event. |
| Temperature control e.g. in leisure centres swimming pools, Industry processes | Pumps, cooling. | Recent history/ nomination.<br><br>Option C or Option B. | The temperature of the water / media changes. This might need to be raised to optimal temperature during the recovery phase. |

| | | | |
|---|---|---|---|
| Pumping and water movement activities, e.g. Irrigation systems<br>Water pumps<br>Treatment plants and desalination and other shiftable processes. | Pumping. | Recent history/ nomination.<br><br>Option C or Option B. | Use of cheaper tariffs – no change in overall energy consumption<br><br>Consider peak measurements to supplement energy consumption changes over a longer period of time . |
| Batteries and EVs | Storage via battery or EV battery. | Zero baseline, asset capacity or fixed reference.<br><br>Option B or C. | Manage the charge and discharge to adapt the consumption with the electricity prices. |
| Residential – large scale aggregation of e.g. 25,000 homes with devices enabled with IoT to control usage. | Washing machines, EV chargers. | Zero baseline, asset capacity or fixed reference.<br><br>Option C for each individual home, or modelling, Option D for a geographical area. | Consider the end uncertainty in the result when modelling large scale aggregation. |

## 4.2 Operational Verification

Operational verification could be a useful tool in the prevention of gaming, especially by recording the conditions in the baseline period and comparing them to the reporting period or flexibility event.

This can evidence that other key equipment and significant energy uses within the measurement boundary did not change during the flexibility event. I.e. that equipment was not deliberately turned to higher uses and consumption to raise (or lower) the baseline (lower in the case of demand turn up). Caution should be exercised however, if for example, pre-cooling actually was used purposefully prior to a flexibility event, that it is not considered gaming.

The same methods used in the IPMVP Core concepts for operational verification can also be used in this case.

Spot measurements can support the process to prove that assets were either turned on or off during baseline period development and during flexibility events, especially if there is concern around gaming.

| OV approach | Typical flexibility application | Activities |
|---|---|---|
| Visual inspection | EV chargers not charging | View and verify that the vehicles are not charging. |
| Sample spot measurements | Individual assets not performing especially if the kW flexibility has been calculated incorrectly, or changes over seasons e.g. the kW electrical savings from an Air Source Heat pump in the summer is different to that in the winter. | Spot measurement of individual assets that make up an aggregated load in a building to determine actual kW rather than assumed. Like IPMVP Option A. |
| Short-term performance trending | Performance of assets may change depending on interoperability or interactive effects if multiple assets impact each other. E.g. The turn down of lighting in a building, increases heating demand, which then impacts the kW flexibility delivered from an Air Source Heat pump, | Short term metering on systems to identify system level changes over the year and in accordance with actual weather conditions. Like IPMVP Option B. |
| Data trending and control logic review | The overall building performance may vary according to actual heating or cooling demands. Component or system being monitored through a BMS can be monitored using independent meters to check back with BMS outputs. | Measurement periods may last from a few days to several months. Sub-meters, independent meters might be required. |

### *4.2.1 The use of facility logs*

Building Management System (BMS) software can serve as a useful mean to prove facility performance in operational verification and thereby support baseline and nullify gaming. E.g. a DER's operational status (ON/OFF Running status), and typical operating hours of plant and equipment are all usually recorded in a BMS.

In addition to the control and monitoring, the data access and storage feature of an BMS facilitates the acquiring of both site and system baseline information. EMS systems can store site data in various resolutions. This data can be in the form of energy data in kWh, battery charge and discharge times, site and system operating temperatures and many other parameters.

These features all facilitate the recording and storage of data for the purposes of baselines and settlement.

## 4.3 Examples of implicit and explicit baseline calculations

### *4.3.1 Implicit baseline*

To calculate the cost savings generated by flexibility event 'behind the meter', the energy price of the energy consumption/demand in the baseline period should be subtracted from the energy price of the consumption/demand in the reporting period.

E.g. In Load shifting, the key reason for flexibility is to avoid energy use when the energy price is high and then use energy later in the day at a lower price. Total energy consumption might not change; however, cost savings will be achieved. In this case:

***Cost Savings = Baseline consumption* baseline energy price – Reporting period consumption* reporting period energy price.***

### *4.3.2 Explicit baseline*

The first step to determining the baseline for explicit events is the selection of the reference days. Representative days (Y days) that precede the day 'A' of the actual flexibility event, should be of the same "category" as day A ( e.g. weekday, weekend day, public holiday etc.). Excluded days prior to the flexibility day 'A' might be:

- where another flexibility event has occurred
- an extreme of weather or a major change in the facility where a non-routine static factor adjustment might be necessary.
- Or dissimilar days (e.g. the baseline is built up of 5 working days, and the flexibility event is on a weekend).

The second step is to select $X$ reference days among $Y$ representative days. They correspond to the X days for which the average net offtake of active power during the corresponding period of day A is the highest.

The third step is the determination of the quarter hourly baseline. The baseline value for each quarter hour is calculated as the average of the X values of the active power of the considered delivery point, measured at the same quarter hour over the X reference days.

The last step is the determination of the hourly baseline. To calculate the hourly baseline, the average of the quarter-hourly baseline profile is calculated.

#### *4.3.2.1 Example of nomination baseline under the NESO DFS market, provided by SSE*

**DFS Operational Baseline Calculation**

The baseline energy consumption is calculated using the eligible days – *days when there have been no DFS used*. In calculating the baseline, the days are categorised into **Working days** (Monday – Friday) and **Non-working days** (Saturday, Sunday and bank holidays).

For working days: 10 most recent eligible days are selected to calculate the baseline.

For non-working days: 4 most recent eligible days are selected and the mean average of the 2 median days will be used to calculate the baseline.

Based on the selected days, the **Actual Usage (kWh)** for each selected eligible day is calculated as follows:

*Day 1 import MPAN – Day 1 Export MPAN*

Based on the actual usage, the unadjusted baseline is calculated by taking the average usage of the eligible days.

*Unadjusted Baseline j = ∑n Actual Usage / n*

*where j is the jth Settlement Period and n is number of Eligible Days*

| Day type | From | To | Day 1 MPAN Total | Day 2 MPAN Total | … | Day 10 MPAN Total | Unadjusted Baseline (kWh) |
|---|---|---|---|---|---|---|---|
| **Working Day** | 00:00 | 00:30 | | | | | = Average (Day 1 MPAN Total: Day 10 MPAN Total) |
| **Working Day** | 00:30 | 01:00 | | | | | |
| **Working Day** | .. | .. | | | | | |

#### *4.3.2.1.1 General commentary*

X-of-Y and rolling average baselines are common methods that rely on historical metered data from end-use customers with flexible loads, managed by flexibility service providers. Control group baselines, which consist of load profiles from similar customers not participating in flexibility programs, can enhance the integrity of baselines for aggregated loads. However, improving baseline accuracy often requires a trade-off in simplicity, as more complex models, such as regression or other machine learning techniques, may be needed.

With the availability of sub-metering for distributed generation resources, meter-before-meter-after baselines are also used. In this approach, a static baseline is defined based on a meter reading taken just before flexibility activation. For non-controllable technologies, like solar PV and wind, a control group approach may improve baseline integrity but may not necessarily enhance accuracy. A more precise alternative is the comparable day baseline, selected by the flexibility service provider to reflect conditions similar to the activation day. However, given the potential for selection bias in this method, buyers may prefer more complex models that incorporate weather data to enhance accuracy.

For controllable technologies, such as Combined Heat and Power (CHP) systems, the approaches mentioned above may be vulnerable to manipulation by the operator, making them less reliable. In these cases, non-regression-based machine learning techniques can offer highly accurate and trustworthy estimates. Although more complex, these models are justified, particularly as the number of distributed CHP plants is relatively small, making model development feasible. Alternatively, a fixed zero baseline could be applied for backup generators, especially if these units are more numerous and provide flexibility.

For storage systems, the zero-baseline approach can also be applied. However, a more accurate estimate may be obtained by using a machine learning model to assess the counterfactual state of the battery.

Self-reported baselines are not recommended.

***Reference:*** *Lind, Leandro, et al. "Baseline methods for distributed flexibility in power systems considering resource, market, and product characteristics." Utilities Policy 86 (2024): 101688.*

#### *4.3.2.2 Example, NYSERDA DLM program*

**Example 2 – NYSERDA**

Con Edison's Dynamic Load Management Program – A Flexible Load Model for Global Adoption

Program Overview

Consolidated Edison (Con Edison) has introduced the Dynamic Load Management (DLM) program, a strategic initiative designed to enhance grid reliability and efficiency through innovative demand response mechanisms. This case study explores the DLM program's structure, its differentiation from conventional demand response programs, and its significance in the global context of flexible load management.

Program Structure

The DLM program consists of two primary components:

1. Term-DLM Program: A day-ahead peak shaving initiative where participants commit to reducing load during specified periods with at least 21 hours' notice. Contracts span 3 to 5 years, offering participants long-term price certainty.

2. Auto-DLM Program: A rapid-response program requiring participants to provide load relief with just a 10-minute notice. This component is designed for critical peak load scenarios and also involves 3 to 5-year contracts.

Key Differentiation from Conventional Demand Response Programs

Traditional demand response programs, such as Con Edison's Commercial System Relief Program (CSRP), typically operate on shorter-term contracts and provide participants with less revenue certainty. The DLM program differentiates itself through:

· Long-Term Contracts: Offering 3 to 5-year agreements provides participants with predictable revenue streams, encouraging investment in advanced technologies like energy storage systems.

· Enhanced Performance Requirements: Particularly in the Auto-DLM program, the requirement for a 10-minute response time necessitates the deployment of sophisticated energy management solutions, ensuring rapid and reliable load relief.

· Higher Compensation but Stricter Performance Criteria: DLM programs provide higher incentives than CSRP or Distribution Load Relief Programs (DLRP), but they also impose stricter performance standards, requiring more reliable load reduction.

Performance Evaluation and Non-Performance Penalties

A critical component of the DLM program is its Performance Evaluation and Payments structure, which ensures that participants meet their contractual obligations. Participants are assessed based on their ability to deliver the committed load relief during events.

Penalties for Non-Performance:

· Annual Reservation Payment Reduction: The annual reservation payment is directly tied to the participant's Average Seasonal Performance Factor. If performance falls below expectations, the reservation payment is reduced accordingly.

· Financial Penalties for Poor Performance: If an aggregation's Average Seasonal Performance Factor falls below zero, the participant may face financial penalties, which could be deducted from their earnings or invoiced separately.

· Early Exit Fees: Participants wishing to reduce their committed load relief can do so by paying an Early Exit Fee, calculated as:

Early Exit Fee = Incentive Rate($/kW)×Deficient Quantity(kW)×Remaining Contract Years×10%

This fee ensures that participants remain accountable for their commitments and cannot withdraw from the program without financial consequences.

· Termination for Chronic Underperformance: If a participant's performance remains consistently low, Con Edison reserves the right to terminate their aggregation contract early, forcing them to pay exit fees and forego future participation.

The DLM Program and Its Potential Role in Global Standards for Flexible Load Verification

Con Edison's DLM program provides valuable insights for utilities worldwide seeking to develop or refine flexible load programs. However, while its Measurement & Verification (M&V) framework is robust and well-structured, it does not represent an official standard or align with any existing global standard. Instead, the program serves as a potential candidate for study, offering key elements that could inform the development of a more widely adopted global standard for flexible load verification.

Key takeaways from the DLM program's M&V approach that could inform future standardization efforts:

· Clear Performance Metrics: The program's methodology for assessing load relief at an aggregation level ensures accountability and reliability.

· Incentive-Driven Compliance: The link between financial compensation and measured performance provides strong motivation for participants to deliver on their commitments.

· Comprehensive Event-Based Evaluation: The structured performance assessment across multiple demand events provides a replicable model for evaluating flexible load resources.

· Integration of Advanced Technologies: The program encourages the use of battery storage, automated controls, and real-time metering, aligning with modern grid modernization goals.

While the DLM program does not serve as a de facto industry standard, its structure and enforcement mechanisms provide a valuable reference point for utilities, regulators, and policymakers worldwide looking to establish robust measurement and verification frameworks for flexible loads.

Conclusion

Con Edison's Dynamic Load Management program represents a significant advancement in demand response strategies. Its innovative approach to long-term contracting, rigorous performance standards, and strict penalties for non-performance differentiates it from traditional programs. More importantly, while it does not constitute a formal standard for flexible load verification, it offers a strong foundation for future discussions on standardization in demand-side management.

By analyzing the program's successes and challenges, utilities and regulatory bodies worldwide can extract key insights to shape globally applicable M&V frameworks, helping accelerate the integration of flexible loads into modern grid systems.

Example 3 – Communities  AIDA

Example 4  ...

## 4.4 Common issues

The following issues can be encountered when defining a fixed baseline period for demand:

1. Lack of historical data due to lack of meters
2. Inadequate time resolution e.g. half hourly data
3. Cost of acquiring data
4. Whether facilities are aggregated, or there are remote / isolated flexibility assets such as in farming.
5. User behaviours
6. Gaming
7. Rebound / interactive effect
8. Negative carbon savings

The first 4 items on the list can be managed via metering (section 5). We will look the last four items in this section.

### 4.4.1 User behaviour

User behaviour can sometimes jeopardize technological solutions placed for flexibility trading. This is mainly related to user comfort in a commercially heated space (e.g., day-to-day operational decisions, like adjusting thermostat settings or turning off unused lights).

One method of minimising this risk is to engage with the facilities staff and to ensure that the sustainability and carbon management teams/ senior management is on board. Comfort levels should be tested prior to engaging in energy and flexibility trading to avoid complaints.

Indeed, user behaviour can also be influenced by the placement of tariffs that support e.g. EV charging at more favourable rates, to encourage users to take excessive energy off the grid. Energy suppliers that manage large portfolios of residents, could influence user behaviour.

### 4.4.2 Gaming

Gaming can be viewed on several levels:

- Sometimes a strategy used by the electricity system operators that consists on the modification their consumption patterns to obtain profit, by taking advantage of the system rules for balancing supply and demand on the grid. To maintain the grid balanced all time, the grid operator accepts bids from generators and consumers to increase or decrease the generation in case of shortage or excess of electricity. This is akin to influencing user behaviours. Market participants that benefit from gaming, take advantage of the difference in prices between the day ahead market and the real time market. In the first one, where the prices are more predictable, the generator underbid the generation, so when more generation is needed during the real time market, the same generator can submit another bid at a higher price, obtaining profit from this price difference.
- An FSP might deliberately increase baselines to support a kW drop in flexibility during a flexibility event. This is by increasing their energy consumption before a scheduled demand response event. If they do so, during the activation hour it appears as if they have reduced their consumption more than what they have done. Conversely, there could also be a logical reason for increasing the baseline, such as pre-cooling a building prior to a known flexibility call.

Gaming has a negative impact in the market, as it can create an artificial congestion on the grid, as their contribution in the real-time market does not solve the congestion. This leads to higher market costs, as these participants are being paid for reductions or increases in their generation, but due to the baseline modification, the real energy consumption is not changed. So, overall, gaming generates market inefficiencies, as no real flexibility is added to the grid when it is needed.

Improvements to the baseline calculation methods with the inclusion of operational verification can support / evidence where increases in baselines are justified.  The pre-cooling events can also be considered as outliers and removed from the baseline calculations, with prior agreement with the market procurer.

### 4.4.3 Rebound/ interactive effect

The rebound effect is a phenomenon where there is an increase in energy consumption following a flexibility event, also known as recovery period. This is to bring the facility back to normal operating conditions. Interactive effects can also occur at a whole facility level such as turning off a DER, impacts the performance of other equipment.

- E.g. turn off a chiller plant during a flexibility event, chilled water temperature drops, causing the room temperature to drop, and then local AC units then start up in response, negating the impact of site-wide kW drop. This is of concern when using asset capacity, fixed reference or zero baselines (deemed approach).
- E.g. turn of a chiller plant, and temperature drops. Post the flexibility event, the chiller needs to stay on for longer to make up the room temperature back to normal conditions.

### 4.4.4 Carbon savings

Generally, carbon savings are defined as either:

- Scope 1 emissions - e.g. on-site use of natural gas in gas boilers
- Scope 2 - e.g. use of electrical energy produced off-site power stations, or
- Scope 3 (supply chain) emissions.

Carbon savings under Scope 2 emissions might be evaluated by using real time carbon grid intensity factors to understand grid intensity during a flexibility event in countries where this data is available. In general, most countries produce at least annual carbon grid intensity factors that are regionalized based upon the nature and type of energy production, e.g. a state in Canada might produce hydropower and have lower emissions than a state that relies on coal-fired power stations.

Carbon savings may need be negative if the primary purpose is to respond to a financial market signal. However, increasingly, the grid is being constrained by an overproduction of Solar PV and Wind or other renewable energy. As such, strategies might be developed where energy is used at times of low carbon grid intensity, and to turn off equipment or have flexibility during these times where the carbon grid intensity is high. This type of flexibility could become important to a Net Zero Carbon future and is known as carbon flexing.

# 5.0 Metering and controls

Effective metering and controls are fundamental for ensuring reliable measurement, verification, and optimization of flexible loads. Advanced metering infrastructure (AMI), interval metering, and automated controls facilitate real-time monitoring, demand response participation, and compliance with industry best practices. This section establishes best practices for metering and control strategies that align with load management programs, market settlements, and regulatory compliance.

## 5.1 Minimum Metering Standards for Load Management

To enable accurate measurement and verification of demand-side flexibility, metering infrastructure should ideally meet the following criteria:

- **High-resolution data collection:** Minimum 15-minute interval data is required for flexibility participation, while some programs may require data resolution down to sub one-minute intervals for frequency response and ancillary services.
- **Bidirectional energy tracking:** Meters must be capable of measuring both energy consumption and energy export for participants utilizing distributed energy resources.
- **Revenue-grade accuracy:** Meters should adhere to standards such as P375 revenue grade meters, ANSI C12.20 or IEC 62053-22 standards to ensure compliance with billing and settlement requirements.

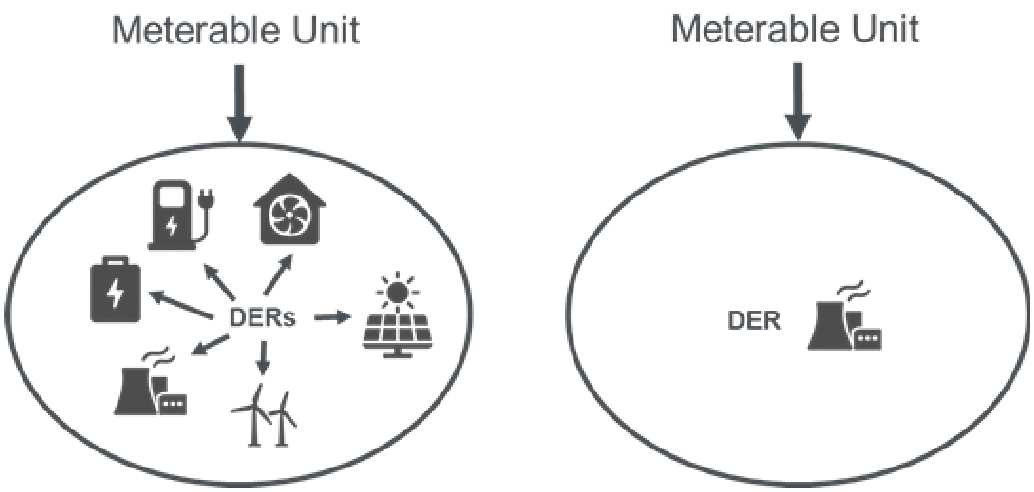
## 5.2 Types of Meters for Flexible Load Applications

The choice of metering technology should align with the level of precision required for program participation and settlement calculations:

1. **Smart Meters (AMI Meters)**
    - Provide automated reporting and remote communication capability.
    - Support demand response program participation through near real-time usage tracking.
2. **Interval Meters with Communication Capability**
    - Measure energy usage in increments of 15 minutes or less.
    - Required for participation in demand response and flexibility programs where real-time data is necessary for verification.
3. **Submetering for End-Use Flexibility Tracking**
    - Allows for granular measurement of specific flexible loads (e.g., HVAC systems, refrigeration, battery storage).
    - Supports baseline adjustments for more accurate demand reduction calculations.

## 5.3 Aggregated portfolios metering

Baselining will be completed at a Meterable Unit level rather than individual DER level. A Meterable Unit is the aggregated level used for settlement and may consist of either multiple DERs or a single DER, as shown in the examples below:




Ref: FMR-SBM, draft 2025.

Under the FMR- SBM draft, the following controls will be implemented:

- Non-Nomination baselined DERs (using the baselines of Asset Capacity, Zero or Fixed Reference) can be grouped into a single Meterable Unit, as all methodologies make use of static values, which are known in advance and are set by DNOs.
- Non-Nomination baselined DERs cannot be grouped into a Meterable Unit containing Nomination baselined DERs, as Nomination baselines use values which are submitted by the FSP and therefore, a Meterable Unit containing Nomination baselined DERs are subject to accuracy standards.
- If an FSP wishes to combine Non-Nomination baselined DERs with Nomination baselined DERs into a Meterable Unit, the FSP would be required to submit an Alternative Baseline Methodology requests to the DNO for all Non-Nomination baselined DERs to become Nomination baselined.

Note: In aggregated meterable units that contain multiple DERs, some DERs might be e.g. increasing in use, others might be decreasing since user behaviour is unpredictable especially at domestic level.

## 5.4 Data Communication and Integration Requirements

For load relief programs, metering data must be seamlessly integrated with market settlement systems and energy management platforms. Best practices include:

- **Secure data transmission:** All metered data should be encrypted and transmitted over secure networks using protocols such as TLS or MQTT.
- **Standardized communication interfaces:** Ensure interoperability with existing grid management systems through industry-standard protocols like Modbus, BACnet, or IEEE 2030.5.
- **Real-time data accessibility:** Participants should have access to their meter data through online portals or API integrations, facilitating performance tracking and verification.

## 5.5 Controls for Flexible Load Management

### 5.5.1 Automated Demand Response and Load Control Strategies

Efficient load control mechanisms enable participants to respond dynamically to grid events and market price signals. Common strategies include:

- **Automated Demand Response (ADR):**
    - Enables direct load control in response to program dispatch signals.
    - Supports OpenADR 2.0b standards for seamless integration with demand response aggregators.
- **Building Energy Management System (BEMS) Optimization:**
    - Load scheduling based on real-time pricing and event notifications.
    - Includes HVAC and lighting controls to shift or curtail demand.
- **Battery Energy Storage Dispatch Optimization:**
    - Dynamic charging and discharging based on day-ahead and real-time market conditions.
    - Supports participation in frequency regulation and peak demand reduction programs.

### 5.5.2 Use of Metering and Control for Verification and Settlement Processes

- **Metered Event Comparison:**
    - Load reduction is assessed by comparing actual metered consumption during an event with the pre-established baseline.
- **Automated Performance Validation:**
    - Control systems should log timestamps for load reductions, ensuring compliance with event participation requirements.
- **Audit and Compliance Checks:**
    - Periodic verification ensures continued adherence to program requirements and validates reported reductions.

### 5.5.3 Regulatory Compliance and Market Integration

For successful participation in flexibility programs, metering and control systems must comply with:

- **Energy baseline and verification standards:** Ensure alignment with industry guidelines.
- **Program-specific settlement rules:** Follow protocols for demand response program performance payments and penalty structures.
- **Grid operator communication requirements:** Ensure that control signals from system operators are received and executed in accordance with program specifications.

## **5.**5.4 Cost of data acquisition

Storage of data from multiple trading portfolios can be expensive and unnecessary. Data storage and archiving strategies should be considered as part of the overall strategy for flexibility; to distinguish between what is required e.g. in particular for nomination baselines, where the last 5 typical days or 10 days is required (meaning, last 5 to 10 weeks' worth of data) and what can be achieved.

In general, a well-designed metering and control strategy is essential for ensuring accurate demand response performance, compliance with regulatory requirements, and successful participation in flexibility programs. By leveraging real-time metering, automated control mechanisms, and standardized verification processes, market participants can maximize operational efficiency and optimize their contributions to grid reliability.

# 6.0 Uncertainty

Uncertainty is an inherent challenge in Measurement and Verification (M&V) of flexible loads, impacting the accuracy, reliability, and credibility of reported performance in energy management and market settlements. Effective uncertainty analysis ensures that flexibility providers, utilities, and market operators have confidence in the measured outcomes of demand response (DR), load shifting, and energy efficiency programs.

## 6.1 Sources of Uncertainty

Uncertainty in the measurement of flexibility delivered can arise and include:

1. **Measurement Uncertainty:**
    - Errors in metering devices, sensors, and data acquisition systems.
    - Instrumentation drift over time affecting data reliability.
    - Non-ideal placement of meters leading to erroneous load readings.
    - Communication and transmission issues in **Advanced Metering Infrastructure (AMI)** and submetering systems.
2. **Baseline Uncertainty:**
    - Variability in pre-event energy consumption patterns due to weather, occupancy, or operational changes.
    - Selection of inappropriate baseline models that fail to account for external influencing factors.
    - Adjustments to historical energy data leading to potential **over- or under-estimation** of flexibility contributions.
3. **Operational and Behavioral Uncertainty:**
    - Human decision-making affecting flexible load performance (e.g., facility operators overriding control strategies).
    - Variability in process loads that depend on production schedules or external demands.
    - Misalignment between **demand response dispatch signals** and the actual operational capacity of flexible loads.
4. **Market and Settlement Uncertainty:**
    - Differences in the interpretation of flexibility event rules by aggregators and system operators.
    - Inconsistencies in M&V methodologies used for energy market settlements.
    - Changes in regulatory requirements impacting compensation and compliance standards.

## 6.3 Quantifying Uncertainty in Flexible Load Measurement

Uncertainty can be quantified through the following techniques:

1. **Meter Accuracy Ratings:**
    - Use of revenue-grade meters compliant with **ANSI C12.20, IEC 62053-22, or equivalent standards** to minimize device-related uncertainty.
    - Periodic calibration of meters to maintain measurement integrity.
2. **Propagation of Uncertainty (ISO GUM - Guide to the Expression of Uncertainty in Measurement):**
    - The **root sum square (RSS) method** is used to combine uncertainty contributions from multiple sources.
    - Sensitivity analysis is performed to evaluate the impact of metering errors on energy reduction estimates.
3. **Monte Carlo Simulation for Load Forecasting:**
    - Probabilistic models simulate different energy consumption scenarios to estimate confidence intervals for measured flexibility.
    - Helps identify **outlier events** that may distort settlement calculations.

## 6.3.2 Evaluating Baseline Uncertainty

Baseline uncertainty is assessed using statistical and analytical techniques, including:

1. **Coefficient of Variation (CV) in Baseline Models:**
    - A high CV in baseline energy consumption data suggests increased variability and reduced reliability.
    - A **CV threshold of <20%** is recommended for strong baseline stability.
2. **Baseline Adjustment Factor (BAF):**
    - Adjusts historical baselines for temperature, occupancy, or production variations.
    - Recommended BAF techniques include:
        - **Regression-based weather normalization** using degree-days.
        - **Adaptive learning algorithms** that continuously refine baselines based on new data.
3. **Cross-validation Against Metered Data:**
    - Comparison of predicted vs. actual consumption using a **Mean Absolute Percentage Error (MAPE) threshold**.
    - **MAPE <10%** is generally accepted for high-confidence baseline models.

### *6.3.2.1 Example of uncertainty in a nomination baseline*

**(Ref FMR – SBM, Draft 2025)**

Accuracy assessments will be completed every month by the DNO, and the results will be reported to the Market Facilitator under the new rules to be implemented in the UK, July 2026.

The accuracy calculation assessment has been designed to account for the maximum usable swing in either direction (import to export and vice versa), constrained by any declared operational limitations (the Available Capacity), and assesses error levels in proportion to the Meterable Unit's capacity, to ensure the measure reflects the relative scale of the Meterable Unit. These will be assessed against the Net % Error and the Absolute % Error.

1. Net % Error – Identify, over an assessment period, the average net error of a Meterable Unit to highlight whether on average, a Meterable Unit's behaviour is consistently over or under an FSP's Nomination values. A negative net error indicates Nominations are higher than the Meterable Unit's normal behaviour, while a positive error indicates the Nominations are under the Meterable Unit's normal behaviour.

$$Net\ \%\ Error_{u,m} = \frac{\sum_{s\in S(m)}(Nomination_{u,s} - Metered\ Output_{u,s})}{\sum_{s\in S(m)} Effective\ Capacity_{u,s}}$$

2. Absolute % Error – Identify performance of a Meterable Unit over an assessment period in absolute terms to mitigate the risk of positive and negative Net % Errors netting to give a false indication of accurate Nominations. The absolute error assessment provides a clear indication of accuracy over an assessment period.

$$Absolute\ \%\ Error_{u,m} = \frac{\sum_{s\in S(m)} \left|Nomination_{u,s} - Metered\ Output_{u,s}\right|}{\sum_{s\in S(m)} Effective\ Capacity_{u,s}}$$

Where Effective Capacity is:

$$Effective\ Capacity_{u,s} = \min(\max(\left|Import\ Capacity_{u,s}\right|, \left|Export\ Capacity_{u,s}\right|), \left|Available\ Capacity_{u,s}\right|)$$

Where:

u = The Meterable Unit being assessed.
m = The month being analysed.
$s \in S(m)$ = Each settlement period "s" within the month "m"
$Nomination_{u,s}$ = The Nomination submitted by the FSP for the Meterable Unit in MW at time "s"
$Metered\ Output_{u,s}$ = The metered value in MW at time "s"
$Import\ Capacity_{u,s}$ = The fully rated import capacity of the Meterable Unit in MW at time "s"
$Export\ Capacity_{u,s}$ = The fully rated export capacity of the Meterable Unit in MW at time "s"
$Available\ Capacity_{u,s}$ = The declared available capacity for the Meterable Unit at time "s"

Thresholds for the measures will be set by the DNOs. Upon failure to meet these thresholds, the DNO will have the right to initiate a Nomination Performance Review.

E.g. a battery energy storage system (BESS) operating bi-directionally over the assessment period illustrates how the assessment of accuracy would work in practice. In the example, there are five periods under consideration (s1 to s5) in the table below, and the standard sign convention of export being a positive value and import being a negative value has been used.

| Interval | Nomination (MW) | Metered Output (MW) | Import Capacity (MW) | Export Capacity (MW) | Available Capacity (MW) |
|---|---|---|---|---|---|
| s1 | -3 | -5 | -5 | 5 | -4 |
| s2 | 2.5 | 1 | -5 | 5 | -6 |
| s3 | -1 | 1 | -5 | 5 | -3 |
| s4 | 3 | 2.5 | -5 | 5 | -5 |
| s5 | -2 | -1.5 | -5 | 5 | -2.5 |

The Nomination (MW) less Metered Output (MW) for each period is aggregated to give a total of 1.5MW of the net error calculation.

| Interval | Nomination (MW) – [A] | Metered Output (MW) – [B] | [A] – [B] (MW) |
|---|---|---|---|
| s1 | -3 | -5 | 2 |
| s2 | 2.5 | 1 | 1.5 |
| s3 | -1 | 1 | -2 |
| s4 | 3 | 2.5 | 0.5 |
| s5 | -2 | -1.5 | -0.5 |
| **TOTAL** | | | 1.5 |

The absolute value of Nomination (MW) less Metered Output (MW) for each period is aggregated to give a total of 6.5MW of the absolute error calculation.

*Table 8: Aggregated absolute value of Nomination (MW) less Metered Output (MW) for s1 to s5*

| Interval | Nomination (MW) – [A] | Metered Output (MW) – [B] | ABS([A] – [B] (MW) |
|---|---|---|---|
| s1 | -3 | -5 | 2 |
| s2 | 2.5 | 1 | 1.5 |
| s3 | -1 | 1 | 2 |
| s4 | 3 | 2.5 | 0.5 |
| s5 | -2 | -1.5 | 0.5 |
| **TOTAL** | | | 6.5 |



The aggregated Effective Capacity is -19.5MW (Table 9).

*Table 9: Aggregated Effective Capacity for s1 to s5*

| Interval | Import Capacity (MW) – [A] | Export Capacity (MW) – [B] | Available Capacity (MW) – [C] | Min (Max(\|A\|, \|B\|), \|C\|) (MW) |
|---|---|---|---|---|
| s1 | -5 | 5 | -4 | 4 |
| s2 | -5 | 5 | -6 | 5 |
| s3 | -5 | 5 | -3 | 3 |
| s4 | -5 | 5 | -5 | 5 |
| s5 | -5 | 5 | -2.5 | 2.5 |
| **TOTAL** | | | | 19.5 |

This gives a Net % Error of 7.69% (i.e. 1.5MW / 19.5MW) and an Absolute % Error of 33.3% (i.e. 6.5MW / 19.5MW).

## 6.4 Mitigating and Managing Uncertainty

Best Practices for Reducing Measurement and Verification Uncertainty include:

1. **Enhanced Metering and Data Collection Standards:**
   - Require **high-frequency data collection (≤5-minute intervals)** for fast-acting flexibility resources.
   - Implement redundancy in data acquisition (e.g., dual-meter setups for critical loads).
2. **Adaptive Baseline Modeling:**
   - Use **dynamic baselining techniques** rather than static historical averages to reduce baseline drift.
   - Incorporate **machine learning algorithms** to refine predictions based on real-time data.
3. **Statistical Confidence Intervals in Settlement Calculations:**
   - Introduce **95% confidence interval analysis** to account for variability in reported energy reductions.

    - Require **data smoothing techniques** (e.g., moving averages) for event performance evaluations.
4. **Independent Third-Party Verification:**
    - Require audits and verification from certified M&V professionals to enhance credibility.
    - Ensure compliance with **ISO 50015 guidelines** for energy performance validation.

## 6.5 Integrating Uncertainty Considerations into Energy Markets

Market settlements must incorporate standardized uncertainty thresholds to avoid **overcompensation or underpayment**. Recommendations include:

- **±5% error tolerance** for metered load reduction verification.
- **±10% baseline variation allowance** before requiring model adjustments.
- Application of **penalty/reward structures** based on deviation from expected performance.

(see example prior).

DNOs might implement performance reviews and sliding scale penalties for failure to meet certainty requirements.

## 6.6 Risk Mitigation Strategies for Market Participants

1. **Contractual Provisions for Uncertainty Allowance:**
    - Demand response contracts should define **acceptable uncertainty limits** and establish **performance adjustment factors**.
    - Include force majeure clauses to account for uncontrollable operational disruptions.
2. **Regulatory Compliance Alignment:**
    - Ensure settlement methodologies comply with **grid operator rules, regulatory frameworks, and M&V protocols**.
    - Regularly update **M&V plans** to reflect evolving energy market conditions.

Uncertainty is an unavoidable aspect of **demand-side flexibility measurement**, but through **rigorous data collection, robust baseline modeling, and statistically sound verification methods**, it can be effectively managed. A structured approach to **quantifying, mitigating, and integrating uncertainty into market settlements** ensures that flexible load participants are fairly compensated and that system operators maintain grid reliability.

By adhering to **best practices** and leveraging **advanced data analytics**, stakeholders can **enhance transparency, reduce settlement disputes, and improve the reliability of demand response and flexibility programs**.

# References


1. International Performance Measurement and Verification Protocol (IPMVP), 2022.
2. ISO 50006:2014 - Energy Management Systems – Measuring Energy Performance.
3. ISO 50015:2014 - Measurement and Verification of Energy Performance Improvements.
4. EU Emissions Trading System (ETS) Reports.
5. U.S. Environmental Protection Agency (EPA) eGRID Database.
6. Asia-Pacific Carbon Reporting Standards.

7. Gade, P. A. V., Skjøtskift, T., Bindner, H. W., & Kazempour, J. (2022). Ecosystem for Demand-side Flexibility Revisited: The Danish Solution. Electricity Journal, 35(9), 107206. https://doi.org/10.1016/j.tej.2022.107206

8. Lampropoulos, I., van den Broek, M., van der Hoofd, E., Hommes, K., & van Sark, W. (2018a). A system perspective to the deployment of flexibility through aggregator companies in the Netherlands. Energy Policy, 118, 534–551. https://doi.org/10.1016/j.enpol.2018.03.073

9. Lampropoulos, I., van den Broek, M., van der Hoofd, E., Hommes, K., & van Sark, W. (2018b). A system perspective to the deployment of flexibility through aggregator companies in the Netherlands. Energy Policy, 118(July 2017), 534–551. https://doi.org/10.1016/j.enpol.2018.03.073

10. Mlecnik, E., Parker, J., Ma, Z., Corchero, C., Knotzer, A., & Pernetti, R. (2020). Policy challenges for the development of energy flexibility services. Energy Policy, 137(November 2019), 111147. https://doi.org/10.1016/j.enpol.2019.111147

11. Ribó-Pérez, D., Heleno, M., & Álvarez-Bel, C. (2021). The flexibility gap: Socioeconomic and geographical factors driving residential flexibility. Energy Policy, 153(March). https://doi.org/10.1016/j.enpol.2021.112282

12. Flexibility Market Rule, Standard Baseline Methods, draft 2025, produced by Energy Networks Association, UK.

Metering References:

1. International Performance Measurement and Verification Protocol (IPMVP), Efficiency Valuation Organization, 2022.

2. ISO 50006:2014 - Energy Management Systems – Measuring Energy Performance.

3. ISO 50015:2014 - Measurement and Verification of Energy Performance Improvements.

4. IEEE 2030.5 - Standard for Smart Energy Profile Application Protocol.

5. OpenADR 2.0b - Automated Demand Response Standard.

Uncertainty references:

1. International Performance Measurement and Verification Protocol (IPMVP), Efficiency Valuation Organization, 2022.
2. ISO 50006:2014 - Energy Management Systems – Measuring Energy Performance.
3. ISO 50015:2014 - Measurement and Verification of Energy Performance Improvements.
4. Guide to the Expression of Uncertainty in Measurement (GUM), JCGM 100:2008.
5. Monte Carlo Methods for Risk Analysis in Demand Response Programs, IEEE Transactions on Power Systems, 2023.

## Appendix 2: Glossary of terms

### Glossary of acronyms

| | |
|---|---|
| ASHP | Air Source Heat Pumps |
| AHU | Air Handling Unit |
| BACnet | Building Automation Control and Network |
| BESS | Battery Energy Storage System |
| BSP | Bulk supply point – see also GSP |
| BM | Balancing Mechanism |
| BMS | Building management system |
| CHP | Combined Heat and Power |
| DC | Dynamic Containment |
| DSR | Demand Side Response |
| DNO | Distribution Network Operator |
| DSO | Distribution System Operator |
| DFS | Demand Flexibility Service |
| ESO | Electricity System Operator e.g. NG ESO |
| GSP | Grid Supply Point |
| HHD | Half hourly data |
| ISO | Independent Systems Operators. |
| MPAN | Meter Point Administration Number |
| NMF | Neutral Market Facilitator |
| PV | Photovoltaic |
| P2P | Peer to Peer |
| RPA | Revenue Potential Assessment |

| | |
|---|---|
| SSEN | Scottish and Southern Energy Networks |
| SCADA | Supervisory control and data acquisition |
| TSO | Transmission System Operator |
| WSHP | Water Source Heat Pump |

## Glossary related to energy markets

**Bilateral Trading** - Bilateral transaction means a transaction for exchange of energy (MWh) between a specified buyer and a specified seller

**BM: Balancing mechanism** - Trading in the last half hour and to balance the grid. This would be energy trading on markets such as Nordpool and EPEX. Trading can be Intra-day (during the day) or DA – Day ahead.

**BM start up** - On-the-day access to additional generation Balancing Mechanism units (BMUs) that would not otherwise have run.

**BMU** - Balancing Mechanism Unit set in units of 1MW.

**BSC**: Balancing and Settlement Code - Code of operation (like a memorandum of understanding) that is administered by Elexon for the settlement of trading post event.

**CM levy** - Introduced in 2017 - a charge based on consumption during the biggest winter peak.

**CM or Capacity market** - This the amount of available capacity (kW) according to the UK's power stations, that is available for the country. Auctions for CM are held 4 years ahead.

**Elexon** - Quasi NGO affiliated to the National Grid that set industry codes and regulations such as the BSC.

**Energy Trading** – Intraday, Day ahead or longer term (e.g. 3 months, 2 years) energy deals between one or more parties. Trading is usually bi-lateral. Longer term trading might be accompanied by insurances by a third party.

**FFR and EFR: Firm or Enhanced frequency response** - Service to provide frequency response within seconds to keep the Grid frequency s close to 50Hz (UK) (Note: other countries may operate at 60Hz). FFR is post fault with a three second response and EFR is pre-fault with a one second response – also known as Dynamic Containment or DC.

**Imbalance settlement** - Process of reconciling an individual party's imbalance positions following the end of a trading period and post National Grid ESO's use of balancing mechanism tools to balance the electricity system in real time.

**LEO Sustain peak management** – demand down and or generation up during Mon - Fri 15.00 - 19.00. Mid May to Mid Sept. 12 hours' notice period.

**LEO Secure DSO constraint management** - demand down and or generation up during Mon - Fri 09.00 - 17.00. Mid July to Mid Sept. 4 hours' notice period.

**LEO Dynamic DSO constraint management** - demand down and or generation up. 24 hours. 30 mins notice period.

**LEO Sustain Export peak management** – demand up, generation down (taking excess from grid or using more at site level). Mon to Fri 10.00 - 14.00 and Mon - Sun 02.00 to 06.00. 12 hours' notice.

**Peer-to-Peer** – trading bilaterally with two counterparties. In the LEO context, this is on headroom or capacity.

**Private wire** – direct electrical connections made between two parties to facilitate the flow of electricity. These wires and connections are not maintained by the DNO or DSO but by the private entity. There are therefore also no network charges to be bear.

**PPA or Power Purchase Agreement** - bi-lateral agreement to purchase energy directly from a generator to a user.

**Route to Market Provider (RTMP)** – organisations such as FSPs that can trade and access energy and flexibility markets. This could also be direct access to the Neutral Market Facilitator.

**Short term operating reserve (STOR)** - Service that provides additional active power from generation or demand reduction. A STOR provider must be able to offer a minimum of 3MW or more of generation or steady demand reduction, deliver full MW within 240 minutes or less from receiving instructions from National Grid and provide full MW for at least 2 hours when instructed.

**Sleeving** - Direct agreement between an electricity consumer (importer) and a generator (exporter). This is also known as 'virtual network' or 'third party netting'.

**Time of Use (ToU) Tariff** - electricity tariff where the charge made for using electricity changes depending on the time of day. A "static" ToU Tariff has set times of day when charges differ and is typically used to encourage people to reduce their electricity use at times of regular peaks in demand. A dynamic ToU Tariff does not have set times of day and the charges depend on how much it costs to generate electricity at any given point.

**Wholesale markets / Energy trading** - Requiring a membership fee for participation - this works on day to day trading of the capacity markets a few months ahead of anticipated demand. Both this and the BM work best when a site can support batteries of a few hundred kW.