TU DELFT

MASTER THESIS

Day Ahead Capacity Limitation for Distribution Grid Relief: Plant Level Modelling and Evidence from a Dutch WWTP

Author:
María Alexandra BOBOCESCU

Supervisors:
Prof.dr.ir. Laurens DE VRIES
Dr. Simon TINDEMANS
Sugandha CHAUHAN

Company Supervisor:
Menno DALMIJN

A thesis submitted in fulfillment of the requirements for the MSc in Sustainable Energy Technology

Energy and Industry
Engineering, Systems and Services

September 25, 2025

TU DELFT

Abstract

Technology, Policy and Management

Engineering, Systems and Services

Sustainable Energy Technology

Day Ahead Capacity Limitation for Distribution Grid Relief: Plant Level Modelling and Evidence from a Dutch WWTP

by María Alexandra BOBOCESCU

The Netherlands faces strong congestion on the distribution grid, which makes it a good setting to study practical congestion management for medium voltage users. This thesis examines contract based capacity limitation with day ahead notice, and develops a plant level optimisation framework for a real wastewater treatment plant. The model represents the site's energy assets, gas plant, a biogas tank, photovoltaic generation and, later on, the addition of a battery. It also includes real Dutch tariffs, metering and verification rules, and the contract logic.

The method uses two variants built on the same core. A perfect foresight set up provides clear upper bound benchmarks and supports design screening. A rolling horizon set up mirrors day ahead operation with forecast error. Experiments focus on a demanding stress week with Monte Carlo call schedules, and include sensitivity tests on prices, demand and tariffs. We evaluate feasibility, operational cost, and the compensation needed to cover the additional costs of participating in the program.

Results show that reliable delivery under contract based capacity limitation is possible without changing core processes, by coordinating on site energy assets and, where appropriate, adding a modest battery sized to the typical call profile. Under imperfect information, feasibility depends mainly on the asset state at the start of a call and on forecast quality. We see that simple operating rules help manage this risk. The economic analysis indicates that asset based flexibility can be delivered at activation prices lower than those linked to production shifting or shedding, as long as compensation covers the extra costs of delivery and account for the bill savings the site might get.

The contributions are a transparent framework that follows Dutch market practices, facility scale evidence on feasibility and economics for a medium voltage wastewater treatment plant, and practical guidance for programme design and site operations. The reasoning extends to other process driven utilities and it points to a scalable path for distribution system operators to procure dependable congestion relief at a competitive cost.

Acknowledgements

I am deeply grateful to my supervisors, Professors Laurens de Vries and Simon Tindemans, for their guidance, sharp questions, and encouragement throughout this project. Their insight and high standards shaped the scope and rigor of this thesis. I also thank Sugandha Chauhan for her thoughtful support, practical advice, and patience when it mattered most.

This thesis was conducted in collaboration with Repowered. I am sincerely thankful for the opportunity and the warm welcome I received. In particular, I would like to thank Menno Dalmijn and Jeroen Jansen for their daily supervision, generous feedback, and many lively conversations on CBCs that kept this work grounded in reality. The broader Repowered team's openness and willingness to share their knowledge made this research far better.

To the friends I made over the past two years: you made my time at TU Delft unbeatable. Thank you for the laughs, late nights, and mutual encouragement.

To my family and friends across the distance, thank you for your steady support. To the Hendricx's: thank you for making me feel at home in another home. A special *thank you* to Tygo, whose support, in so many ways, was crucial.

Any remaining errors are mine alone. I hope the findings are useful to practitioners and researchers working to make congestion management more effective, fair, and practical.

Contents

List of Figures

List of Tables

List of Abbreviations

ACM	Autoriteit Consument en Markt
ATR	Alternative Transmission Rights
ATO	Aansluit- en Transportovereenkomst
BAU	Business as Usual
BESS	Battery Energy Storage System
BRP	Balance Responsible Party
CBC	Capacity Limitation Contract
CC MPC	Chance-Constrained Model Predictive Control
CAPEX	Capital Expenditure
CHP	Combined Heat and Power
CSP	Congestion Service Provider
DA	Day-Ahead
DSO	Distribution System Operator
EPEX	European Power EXchange
FRR	Frequency Restoration Reserve
GDS	Gesloten DistributieSysteem
GOTORK	Gebruik Op Tijd Of Raak Kwijt
GTO	GroepsTransportOvereenkomst
HV	High Voltage
IGDT	Info-Gap Decision Theory
KGG	Ministerie van Klimaat en Groene Groei
LAN	Landelijk Actieplan Netcongestie
LCCR	Levelised Cost of Congestion Relief
LV	Low Voltage
MC	Monte Carlo
MLOEA	Meerdere Leveranciers Op Één Aansluiting
MPC	Model Predictive Control
MV	Medium Voltage
NFA	Non-Firm ATO
O&M	Operations & Maintenance
PCC	Point of Common Coupling
PF	Perfect Foresight
PTDF	Power Transfer Distribution Factor
PV	PhotoVoltaic (solar)
RH	Rolling Horizon
SoC	State of Charge
TBTR	Time-Block Transmission Rights
TDTR	Time-Dependent Transmission Rights
TSO	Transmission System Operator
UIOLI	Use It Or Lose It
WWTP	WasteWater Treatment Plant

Chapter 1

Introduction

This chapter sets the scene for the study. It defines the problem (Section 1.1), highlights the knowledge gap (Section 1.2), states the research objectives (Section 1.3) and questions (Section 1.4), and ends with a short guide to the thesis structure.

1.1 Problem Definition

The transition to a low carbon energy system is a central priority to achieve global climate targets, as defined in the Paris Agreement and the European Green Deal. The Paris Agreement seeks to prevent an increase of the global temperature beyond two degrees Celsius above pre industrial levels by reinforcing greenhouse gas reduction [1], while the European Green Deal aims to make Europe the first carbon neutral continent by 2050 [2]. The consequences of failing to achieve these targets are severe and pose significant threats to human life, ecosystems, and economic stability.

The Netherlands has shown a strong commitment to renewable energy integration and decarbonisation through the implementation of national climate strategies, including the Dutch Climate Act and the National Climate Agreement [3, 4]. However, the country faces significant challenges in meeting its 2030 climate goals due to growing grid congestion [5]. The rapid electrification of industries, businesses, and households, combined with the accelerated integration of variable renewable energy sources, has resulted in capacity constraints and a sharp increase in flexibility needs, particularly in highly populated and energy intensive regions [6].

Grid congestion in the Netherlands presents a twofold challenge. It compromises progress toward climate targets, and it hinders economic growth by preventing the expansion of existing industries and delaying new grid connections. Traditional solutions, such as grid reinforcement or grid expansion, require significant investment and long implementation timelines, which makes them difficult to deploy as short term responses [6]. Alternative congestion management strategies are therefore required.

Among the proposed solutions, industrial demand flexibility has emerged as a promising approach to mitigate congestion. By adjusting electricity consumption in response to network constraints and signals, industrial consumers can unlock significant potential, improve system stability, and benefit from economic incentives [7]. There are many mechanisms available but In practice, participation rates remain low.[8]. Smaller industries often do not perceive themselves as potential providers of flexibility for congestion management, yet their aggregate potential is valuable [9]. The Netherlands is one of the countries most affected by grid congestion, which makes it a critical case to understand barriers and opportunities associated with industrial flexibility. Lessons from the Dutch experience can inform other countries as they progress toward greater electrification and renewable energy integration [10].

In this context, and in collaboration with *Repowered*, a Dutch *smart-energy* provider that helps large connection owners optimise energy flows and costs by making assets smart, coordinating demand and supply, and enabling access to energy markets, this thesis evaluates capacity limiting contracts for demand-side congestion management for industrial users with largely inflexible demand and existing connections to the Dutch regional grid. A municipal wastewater treatment plant (WWTP) with these characteristics serves as the case study.

1.2 Knowledge Gap

As we will explore in the chapters below, the threat of grid congestion to the stability and the development of the electricity system has been documented by Dutch and international sources. Studies consistently highlight that congestion is an urgent and growing challenge, yet it remains tractable if suitable measures are implemented [11]. Despite industrial participation in congestion management (off-take) being recognised as a useful tool, actual participation in the Netherlands is still limited [7]. The barriers include uncertainty over financial outcomes, limited internal capability to quantify benefits, and operational concerns in process driven facilities [8].

In response, Dutch actors have introduced a range of congestion mechanisms and flexible contract structures aimed at different user groups. The variety of options has expanded quickly, and new mechanisms continue to appear. However, many industrial consumers, particularly those with process driven operations, do not perceive themselves as active participants and opt out [6]. The national framework provides the enabling conditions, but facility scale evidence that speaks directly to industrial operators is scarce.

Two gaps emerge from the literature and from practice. First, there is a lack of facility scale, data driven assessments of contract based congestion mechanisms at a single connection point that is both process driven and connected at medium voltage. In particular, there is limited work that uses real industrial time series and Dutch tariff structures to represent call based capacity limitation, and to assess feasibility and economics under realistic operational constraints [12]. Secondly, there is limited quantitative analysis of compensation levels that make participation viable for the industrial user while delivering reliable reductions for the distribution system operator (DSO), taking into account metering rules and the alignment with market processes [7, 8].

This thesis addresses these gaps by developing a real world, **data driven model** that evaluates **industrial demand flexibility** in a **process driven** setting, aligned with Dutch congestion challenges. In contrast to purely theoretical models, the study integrates **measured load profiles**, contract call signals and settlement elements to provide practical insight into feasibility and **economic implications** for both the site and the DSO. The modelling is staged to separate clean capability benchmarking from design exploration and from rolling horizon realism, which matches the way call based capacity limitation is operated in practice.

1.3 Research Objectives

The objective is to assess whether, and under what conditions. an industrial site with a rigid load profile, in this case a waste water treatment plant, can support congestion management on the Dutch medium-voltage grid using only on-site assets, without altering core process operations. The evaluation combines mechanism selection, site level modelling, and economic assessment. Each objective below is formulated to be achievable with the methods and analyses documented in the current thesis, and is mapped to the chapter or section where it is addressed.

Objective 1, Mechanism positioning and design cues. Synthesise the congestion management programmes that are relevant for a single industrial site in the Netherlands.

Objective 2, Site level core model and benchmarks. Formulate a plant level optimisation model that captures the wastewater treatment plant's controllable resources and operational constraints, and that values operation with Dutch price and tariff components.

Objective 3, CBC representation and design sweep. Represent call based CBCs in the model with day ahead notices and capacity caps over selected hours, aiming to mimic reality as close as possible.

Objective 4, Economic viability and compensation. Quantify the CBC compliance at the site across different experiments and infer the compensation that would make participation economically neutral or attractive to the facility.

1.4 Research Questions

To make the objectives operational, the thesis is guided by one main question and five sub questions. The main research question is:

What is the potential for an industrial facility with inflexible electricity demand to support congestion management on the Dutch medium voltage grid, and under what conditions can participation be economically attractive for both the facility and the DSO

Building on this, the sub-questions are structured to follow the flow from mechanism choice to modelling, feasibility and costs, and sensitivity.

SQ1. What congestion management programmes and demand response mechanisms are currently available in the Netherlands, and how do they operate?

SQ2. Which of these programmes are technically and operationally suitable for a facility with a fixed, process driven demand profile such as a wastewater treatment plant, considering notice, metering, and enforcement?

SQ3. How can the technical and economic impact of participation be modelled and assessed for such a facility using measured data and Dutch tariffs?

SQ4. What compensation is required to make participation viable from the perspective of the facility?

SQ5. How do external conditions, and investments in flexible assets, influence the integration of the congestion management programme and its economic attractiveness?

1.5 Thesis Structure

The remainder of the thesis is organised as follows. **Chapter 2** introduces the case-study WWTP, its assets and layout, grid interface, and the process constraints that shape flexibility. **Chapter 3** frames Dutch grid congestion: core concepts, drivers, and regional patterns; it then summarises the regulatory setting and available congestion-management instruments. **Chapter 4** synthesises the literature regarding mechanism landscape, DSO economics, WWTP-specific flexibility, and state-of-the-art modelling approaches. **Chapter 5** details the modelling framework, staged experimental design, data and decision criteria. **Chapter 6** reports findings for the experiments, **Chapter 7** reflects on method, strengths and limitations and answers the research questions. Finally, **Chapter 8** summarises the contributions and highlights avenues for future work.

Chapter 2

System Description

This chapter sets out the case facility and how it operates in practice. Section 2.1 describes the physical layout, key assets, and baseline parameters. Section 2.2 details the grid connection, contracted limits, metering and tariff structure. Finally, Section 2.3 explains the process constraints that shape flexibility.

2.1 Current Layout

The industrial facility under study is a wastewater treatment plant (WWTP) located in a congested region within Liander's electricity grid, and serves as a critical infrastructure for the region. For the purposes of this study, the current system configuration will be referred to as **Base Case** going forward. A representative facility is shown in Figure 2.1.

FIGURE 2.1: Illustrative aerial view of a typical wastewater treatment plant.
Source: EcoMENA [13].

The WTTP layout in the Base Case is showcased in Figure 2.2. The site consists mainly of two wastewater treatment facilities, shown as a single unit in the figure. These facilities are designed to handle a continuous inflow of wastewater, although the volume can fluctuate significantly. These fluctuations are mainly due to variations in weather, unpredictable industrial discharges, isolated events such as local festivals or pipe failures, and other occasional emergencies that may require the treatment of unexpectedly large volumes of water.

At the site, electricity is the **only** energy carrier used for water treatment, and therefore the electric demand is strongly linked to the hydraulic load. A simple diagnostic over the 2023 year, shows a Pearson correlation of **0.727** between them, which means the two series move together in a fairly strong, positive linear way. As observed in Figure 2.3, when the hydraulic load is higher, the electricity demand tends to be higher and vice versa.

Part of this electric demand is met via direct supply from the grid, but there is also on-site generation. The facility includes a Combined Heat and Power (CHP) plant that operates

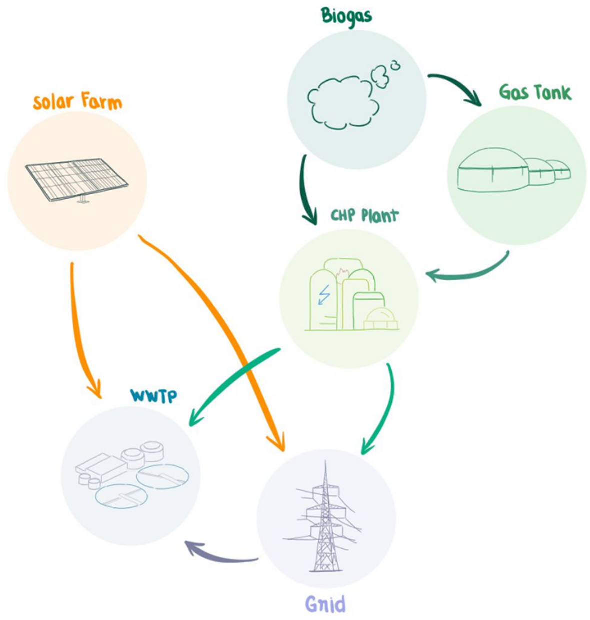
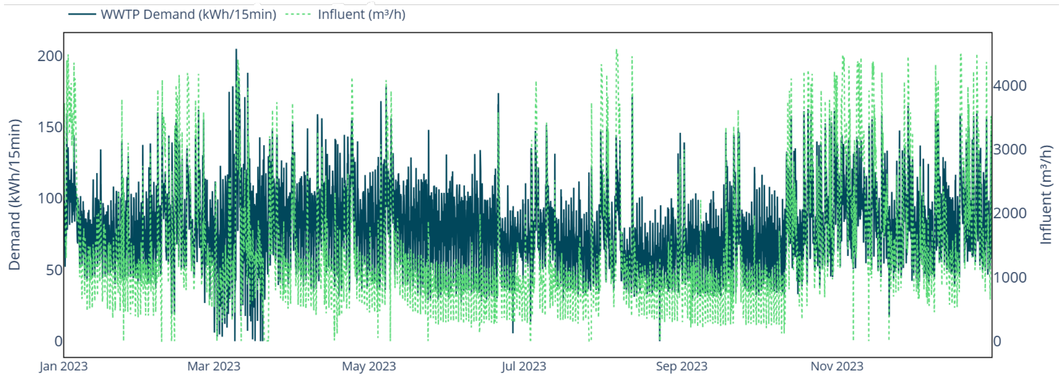
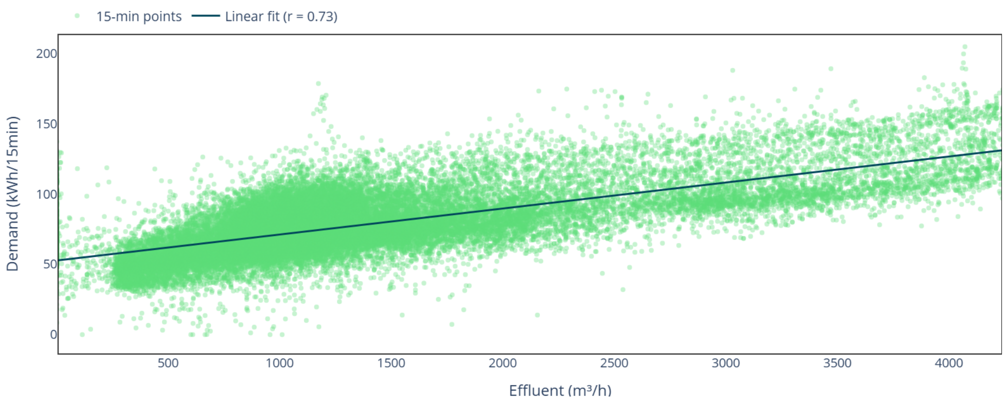


FIGURE 2.2: Current Layout WWTP



(A) WWTP electric demand vs. hydraulic load



(B) WWTP electricity demand vs. effluent flow with linear fit

FIGURE 2.3: Relationship between hydraulic/effluent load and WWTP electricity demand.

on biogas. In principle a CHP delivers both electricity and useful heat, however at this site the heat has no practical use, there is no process demand or nearby consumer, so it is effectively wasted. Operationally, the unit simple behaves as an electricity generator. For clarity in the remainder of the thesis we therefore refer to this unit as the **Gas Plant**. The biogas that fuels this unit is a byproduct of internal processes within the WWTP itself, and it is stored in a dedicated tank with a total capacity of 400 m^3. However, due to operational safety constraints, the usable storage is between 40–75% of its total capacity. In addition to the Gas Plant, the facility operates a Solar Park with **1,944 panels**. A summary of the main specifications of the Base Case layout is provided in Table 2.1.

When it comes to future changes, the demand is expected to increase by approximately 5% over the next five years, in line with the increase in water consumption in the region, but no other major changes are expected.

TABLE 2.1: Key system parameters

Parameter	**Value**	**Comments**
Total demand	2764670 kWh/year	100% electric
Solar production	423902 kWh/year	450 kWp
Gas Plant production	1991336 kWh/year	450 kW
Biogas production	951693 m3/year	Byproduct
Biogas storage tank	400 m^3	Usable range 40–75%

2.2 Grid Interaction Characteristics

Currently, the WWTP interacts with the electricity grid under relatively conventional and static conditions. The plant simply draws electricity as needed to maintain continuous operation, and there is no form of operational optimisation; its assets are managed independently of electricity prices or grid conditions.

The connection to the medium voltage grid is characterised by fixed contractual limits: the plant can import up to **1,290 kW** of power at any given moment (off-take) and can export up to **1,236 kW** (feed-in).

Figure 2.4 shows the WWTP's electricity demand over the year 2023 (15-min resolution) together with the contracted import capacity converted to energy per step. We can see that the electric demand does not show a clear seasonal or daily pattern but remains far below the contracted level at all times. Hence, under current operation the contracted capacity is **non-binding.**

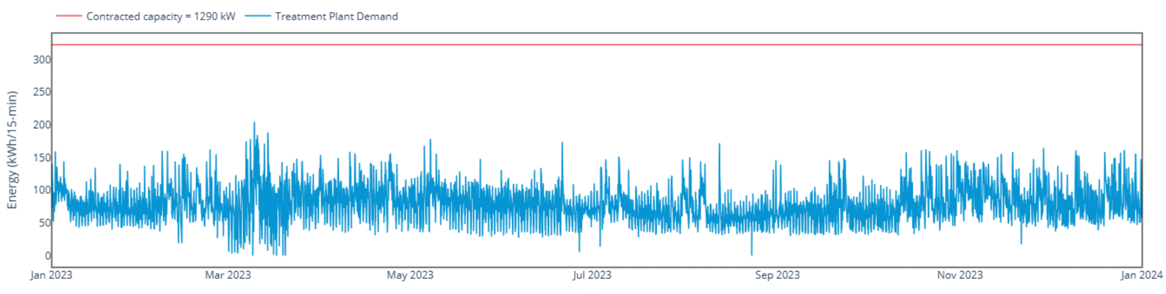


FIGURE 2.4: WWTP Demand vs. Contracted Capacity - Year 2023

Figure 2.5 presents the Base Case grid imports alongside the contracted capacity, and the monthly maximum 15-minute peaks, labelled in kilowatts. Compared with total plant demand, the magnitude and shape of the import series differs slightly, which reflects the contribution of on site generation and its effect in reducing the need for external electricity.

As before, imports fluctuate considerably and no clear seasonal pattern emerges. They remain consistently below the contracted line, which confirms that the contracted cap does not constrain the plant. The monthly peak values that drive the relevant grid tariffs vary notably across the year, with the highest recorded on 9 May.

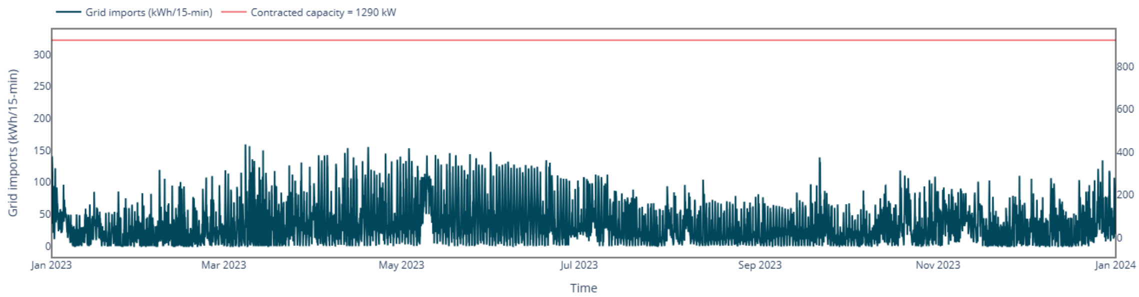


FIGURE 2.5: Grid imports (Base Case) — Year 2023

In Figure 2.6 we zoom in on one week to examine how EPEX prices and grid imports move together. The price series shows spikes and quiet periods, however, imports do not consistently follow these changes. This confirms that grid purchases are driven mainly by process demand, after on site generation is accounted for, rather than by market signals.

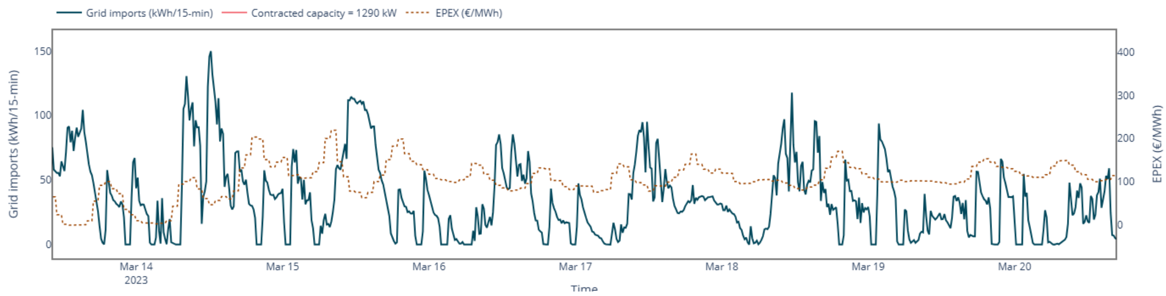


FIGURE 2.6: Grid imports (Base Case) vs. EPEX prices — Zoomed view to inspect short-term co-movement.

Taken together, these visuals indicate that (i) the existing contracted capacity is generously sized and never binding, and (ii) without additional constraints or flexibility the plant's imports are governed by operational needs rather than prices.

Regarding the **grid tariffs**, the plant presents a standard structure for medium to large industries connected to Liander's grid. It consists of a fixed fee, which covers the basic right to connect to the grid, alongside several variable charges. These variable components are calculated monthly and are based on three main factors: the level of contracted grid capacity, the actual volume of electricity transported during the billing period, and the highest measured power peak within that period. Currently, the WWTP does not optimise for these values. According to the characteristics described, the category in which the WWTP falls and the specific prices that apply to it can be seen in Table 2.2.

Segment	**Contracted capacity** (kW)	**Fixed transport charge** (€/month)	**kWh** (€/kWh)	**Contracted kW charge** (€/kW·month)	**Monthly max kW charge** (€/kW·month)	**Reactive energy** (€/kvarh)
MV	>136 to 2 000	36,75	0,016	1,82	2,59	0,00

TABLE 2.2: Grid tariffs of WWTP - Year 2023 - Liander [14]

2.3 Process Constraints and Implications for Flexibility

What distinguishes this facility from other industrial consumers is its **obligation** to treat all incoming wastewater according to established quality standards. Failing to do so could result in serious societal consequences, including localised flooding or the discharge of untreated water into rivers and lakes. Additionally, wastewater arriving at the plant is often already in degradation, emitting odours and allowing bacteria to proliferate, meaning that a fast treatment is necessary. Therefore, delays in treatment not only pose public health risks, but also complicate the treatment process itself. As a result, these plants are designed for continuous **baseload operation** and cannot store wastewater for later batch processing; water storage tanks are used exclusively to buffer short-term inflow surges. Another important characteristic is the high degree of automation of the plant, which eliminates the need for regular on-site staff. As a result, there is no significant electricity demand for lighting, heating, or cooling. In addition, as mentioned before, the heat generated by the CHP unit is wasted Consequently, there is no opportunity to introduce flexibility through adjustments in building management or by using residual heat. These characteristics mean that process-side load shifting or shedding, key elements of many congestion management mechanisms, are largely **infeasible** for this facility.

In essence, we can determine the **site operates continuously under effluent quality regulation with very little scope to reschedule core loads**. As a result, the WWTP's electrical demand is essentially inflexible. We therefore classify the plant as a process industry facility, as defined in Chapter 4, with comparable examples such as cement plants, glass furnaces, and paper mills discussed in that chapter.

Within this important constraint, reducing imports at the **point of common coupling** remains feasible. The point of common coupling is the location in the public network where a customer connects to the distribution system, and it is the reference for metering, voltage characteristics, and harmonic limits [15]. By dispatching the current assets: the gas plant and the on-site photovoltaic production, when available, the site *could* shave grid imports **without altering the treatment process itself**. In other words, the plant demand remains unchanged, but the grid *sees* a different import profile at the PCC, which is consistent with the inflexible nature of the process side. The extent and limits of this import shaving are studied in the following chapters.

Chapter 3

Policy Review

This chapter situates the study within the Dutch policy and market framework for managing a rapidly electrifying, renewable-rich system. It opens with the national climate and energy context, then narrows to the legal and regulatory architecture that governs congestion management. Building on that foundation, it reviews the available instrument explaining how they operate in practice. The chapter closes by matching these mechanisms to the case facility's characteristics, identifying the most suitable option to implement and test in the remainder of the thesis.

3.1 What is Grid Congestion?

A brief clarification of the core concepts of grid congestion will provide a solid basis for the chapters that follow.

Firstly, grid congestion should not be confused with grid imbalances, as they are different but interconnected concepts within power system operations. **Grid imbalances** occur when electricity supply and demand do not match, usually due to unexpected fluctuations, such as the sudden failure of a power plant or unpredicted changes in electricity consumption. To manage these imbalances, the Transmission System Operator of the Netherlands (TSO), TenneT, uses various mechanisms such as the balancing market. It provides real-time solutions through diverse balancing services such as the Frequency Containment Reserve (FCR), automatic Frequency Restoration Reserve (aFRR), and manual Frequency Restoration Reserve (mFRR), ensuring continuous stability in the electricity network.

In contrast, **grid congestion** occurs when a specific point in the electricity grid lacks sufficient physical capacity to transport the required electricity. This congestion can manifest in two forms, **feed-in** (supply) congestion and **off-take** (demand) congestion. Feed-in congestion occurs when an excess of electricity generation, often from VRE sources like wind and solar, exceeds the grid's ability to transport it to areas of demand. Off-take congestion, on the other hand, occurs when electricity consumption surpasses the grid's supply capacity, leading to overloads in transmission lines and substations. Most regions experience both types of congestion at different times of the day or year, depending on fluctuations in energy demand and renewable generation patterns.

Although grid imbalances can contribute to grid congestion, and vice versa, a fundamental distinction is that grid congestion can often be **predicted** and **prevented**. TenneT and the DSOs have the capability to forecast congestion areas based on expected demand profiles and grid capacity limitations for the following day. This predictability allows for preventive measures to be taken in advance, unlike grid imbalances, which often require real-time interventions. Various congestion management mechanisms, designed to address these challenges, will be explored in later chapters.

3.2 Causes of Grid Congestion in the Netherlands

In this section we examine *why* grid congestion has emerged in the Netherlands. We outline the main drivers and how they have developed over time, in order to provide context for the problem and to frame the mitigation approaches.

3.2.1 The decarbonisation efforts

The Netherlands is subject to various climate policies aimed at mitigating the climate crisis. Some key international agreements that influence the Dutch energy transition are the **Paris Agreement** (2015) [1] and the **European Green Deal** (2019). The Paris Agreement seeks to limit global temperature rise to below 2°C above pre-industrial levels, while the Green Deal aims to achieve climate neutrality by 2050 through a 55% reduction in emissions by 2030 compared to 1990 levels [2]. Both agreements emphasise investments in **renewable energy**, **electrification**, and ***green* technologies**. To comply with these commitments, the Netherlands has developed its own national climate policies, the Dutch Climate Act (Klimaatwet), and the Dutch National Climate Agreement (Klimaatakkoord) [3, 4].

Historically, the Netherlands has been a carbon-intensive country, heavily reliant on fossil fuels for energy production. However, following the climate policies, recent years have seen significant efforts to transition towards renewable energy sources [16]. Looking at electricity generation trends in Figure 3.1, fossil fuels, particularly natural gas and coal, previously dominated the Dutch energy mix. Nevertheless, since 2015, coal-fired generation has decreased significantly, falling from the second-largest source to the fourth-largest in 2023, overtaken by wind and solar power. Natural gas still remains the primary energy source, but its overall usage has decreased compared to 2020 levels.

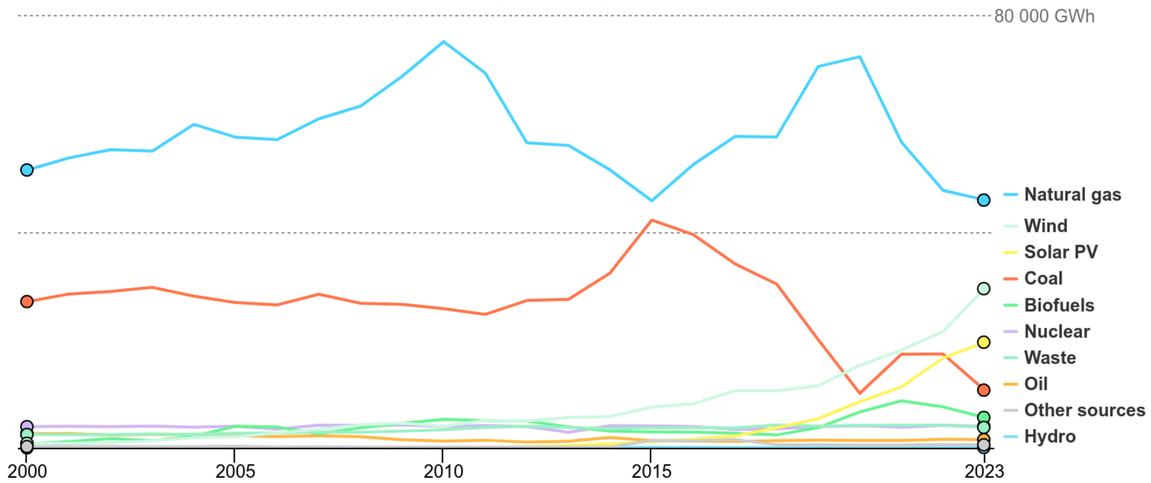


FIGURE 3.1: Electricity generation by source in the Netherlands, 2000–2023. Source: International Energy Agency. Licence: CC BY 4.0 [17].

The growing role of wind and solar energy reflects the Netherlands' commitment to decarbonisation and phasing out thermal power plants. But, despite this progress, natural gas still provided almost forty per cent of electricity in 2023, and fossil fuels taken together, coal, oil, and natural gas, supplied about half of total generation, as shown in Figure 3.2.

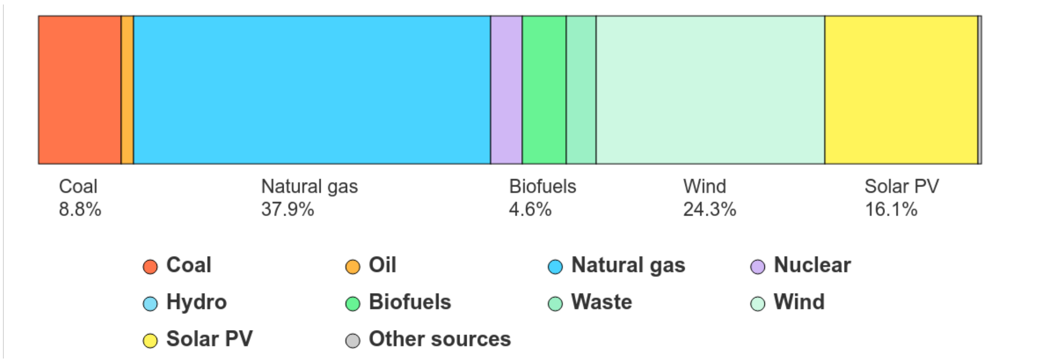
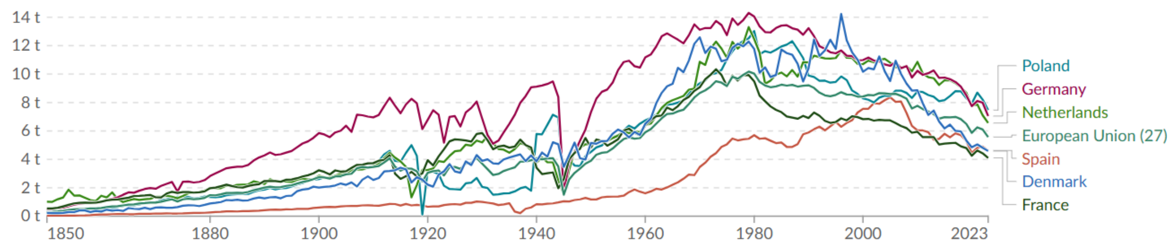
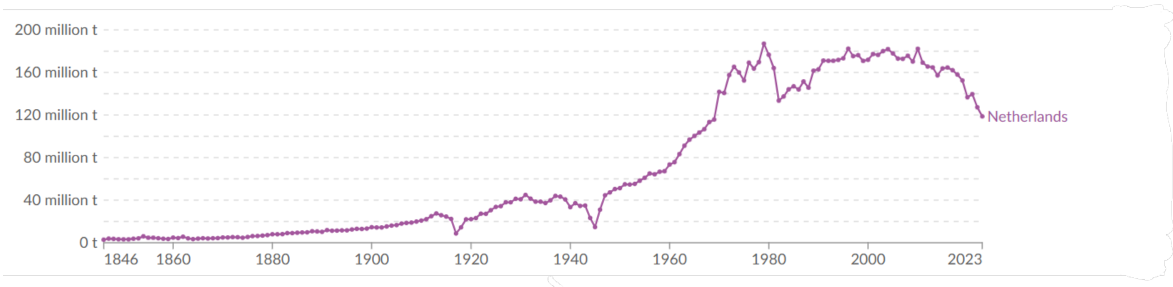


FIGURE 3.2: Mix of generation in the Netherlands, Year 2023. Source: International Energy Agency. Licence: CC BY 4.0 [17].

Furthemore, Figure 3.3 (A) shows that the Netherlands still emits more CO_2 per person than the EU average. Subfigure (B) also makes clear that, although total annual emissions are declining, current levels remain well above what would be required to meet the 2030 target of a 55% reduction relative to 1990.



(A) Per-capita CO_2 emissions for selected European countries, 1850–2023.



(B) Total annual CO_2 emissions in the Netherlands, 1850–2023.

FIGURE 3.3: CO_2 emissions visualisations. Source: Global Carbon Budget (2024). Licensed under CC BY 4.0 [18].

Taken together, these figures tell a consistent story: (i) despite strong growth in wind and solar, the Dutch power mix remains fossil-intensive, natural gas is still the largest source and fossil fuels together supply about half of generation; (ii) total and per-capita CO_2 emissions are declining, but current levels are still well above what would be needed for a 55% cut by 2030.

To close this gap, integration of **variable renewable energy (VRE)** and the **electrification** of heat, transport and industry have become key priorities in both national and international agreements over the past decade. Analyses indicate that meeting climate and energy targets will require the share of electricity in final energy use to almost double, rising from about 33% in 2022 to more than 60% by 2050 [19].

For this reason, as mentioned before, recent years have seen a strong rise in VRE generation, particularly wind and solar PV. In parallel, the Dutch **industrial sector** has increased electrification efforts, replacing thermal systems with electric alternatives. Some examples include the adoption of electric boilers, industrial heat pumps, hydrogen production, and the expansion of electric vehicle infrastructure. If widely implemented, these initiatives could cut industrial carbon emissions by **78%** [20].

Essentially, this **rapid expansion of VRE** and the **growing electricity demand**, have been among the **primary drivers of grid congestion** in the Netherlands [16]. This structural issue stands as a significant challenge within the energy transition process: decarbonisation requires both high levels of electrification and the large-scale integration of VREs, but the intermittency of these energy sources adds significant strain to the electricity grid, increasing system instability and making it harder to match supply and demand. As a result, there is a growing need for flexibility mechanisms to help balance the system more effectively [11]. At the same time, the limited grid capacity to support electrification and renewable deployment directly hinders broader decarbonisation goals. This creates a **self-reinforcing loop**, where the conditions required to solve the problem are restricted by the problem itself, which emphasises the urgency of targeted solutions.

3.2.2 Other factors that worsen grid congestion

The growth of renewable energy and electrification is leading to grid congestion because the existing grid was not originally designed for such fluctuating supply and rapidly increasing demand, and its **expansion** simply cannot keep up with these changes. The Dutch network operators estimate that about €219 billion of investment will be needed between 2024 and 2040, which in turn requires electricity network to rise between 4.8% and 6.7% per year, roughly doubling or tripling by 2040 [21]. Expansion is further slowed by competition for scarce land and a shortage of technically qualified staff. In addition, long permitting and licensing procedures add another layer of delay, especially due to national nitrogen rules. The Netherlands has the highest nitrogen emissions per hectare in Europe, primarily from its intensive agricultural sector [22]. Therefore, projects that add nitrogen emissions, including those for grid expansion, struggle to obtain permits. Liander alone reports that 317 construction projects are delayed or at risk of cancellation [23]. As a result, the timeframe to expanding the grid, ranges on about 5 to 10 years, depending on the specifics of each location, permitting, capacity and deployment, however, these figures are indicative [24].

There is also a **geographical mismatch**: most renewable generation is located in areas with relatively low demand, while many large industrial consumers are based in regions where the grid lacks the capacity to deliver enough electricity efficiently [25]. On top of this, there is a **temporal mismatch**, industrial consumption patterns often do not align with the peaks of variable renewable energy (VRE) generation, especially solar, which adds even more strain to the system [26].

Moreover, the current **grid tariff structure**, which makes up a large part of the total energy costs for industries, does not always encourage flexible consumption. For example, flat-rate tariffs give industries little incentive to adjust their demand in response to grid constraints or changes in renewable output [14].

It is also important to recognise that the way **transmission capacity is distributed** aggravates the problem. A shortage of connection capacity in an area, is not always (only) the direct result of **physical** grid congestion. As previously explained, congestion refers to periods when the grid infrastructure is temporarily overloaded, which leads grid operators to restrict new or expanded connections, however, the issue is also linked to the way connection and transmission contracts (Aansluit en Transportovereenkomst, or ATO) are managed. Under the current system, industries must contract their desired transmission capacity with the grid operator, but in many areas, no additional capacity appears to be available, even when the physical infrastructure could handle more. This is because capacity allocation is generally handled on a *first-come, first-served* basis, in accordance with European principles of non-discrimination. However, this method can leave existing capacity unused for extended periods. For example, if the first company in line requests more capacity than is currently available, the grid operator is unable to allocate any of that capacity to others further down the queue until the first applicant has either accepted a smaller allocation or withdrawn their request. As a result, this allocation process can be slow and inefficient, causing valuable capacity to sit idle unnecessarily [11, 6].

From this analysis, it is clear that grid congestion in the Netherlands results from a combination of pressing issues:

- **Rapid expansion of VREs**
- **Growing electricity demand, especially industrial**
- **Grid development lagging behind, due to several reasons**
- **Geographical and temporal mismatches**
- **Limited flexibility incentives**

Additionally, the **capacity allocation inefficiencies** intensify the consequences of grid congestion for society.

These structural challenges have made grid congestion an persistent problem, stressing the need for innovative solutions that allow for a **more efficient use of the existing grid capacity** while it gradually expands.

3.3 Regional Distribution of Grid Congestion

Grid congestion in the Netherlands varies significantly by region, with certain provinces and municipalities experiencing more severe constraints than others. The most affected areas include **North Holland**, the southern provinces of **Brabant** and **Limburg**, and several smaller towns.

In North Holland, grid congestion has been a long-lasting issue, primarily due to the rapid housing development, the establishment of data centres near major internet connection hubs, and the expansion of large-scale solar photovoltaic (PV) projects attracted by relatively affordable land prices in the northeast. As a result, most of Northern Holland's electricity network cannot accommodate new large-scale energy users, significantly affecting consumers requiring grid connections above 3×80 amperes [6].

In Brabant and Limburg, TenneT announced a suspension of new grid connection applications in 2022. This suspension was largely due to the rapid electrification of industries aiming to reduce CO2 emissions and industries transitioning from gas to electricity to avoid escalating gas prices during the energy crisis in that year, which aggravated grid congestion. This situation stressed the urgent need for policy interventions, leading to the initiation of the **National Grid Congestion Action Programme** (Landelijk Actieprogramma Netcongestie or LAN) a joint programme of governments, grid operators and market players to tackle grid congestion [27].

Figure 3.4 shows regional differences in **off-take** grid congestion based on Netbeheer Nederland's *capaciteitskaart* [28]. **Red** indicates no transport capacity (waiting list), **orange** marks areas under review (waiting list), **yellow** denotes limited capacity available (no waiting list), and **white** indicates capacity available (no waiting list).

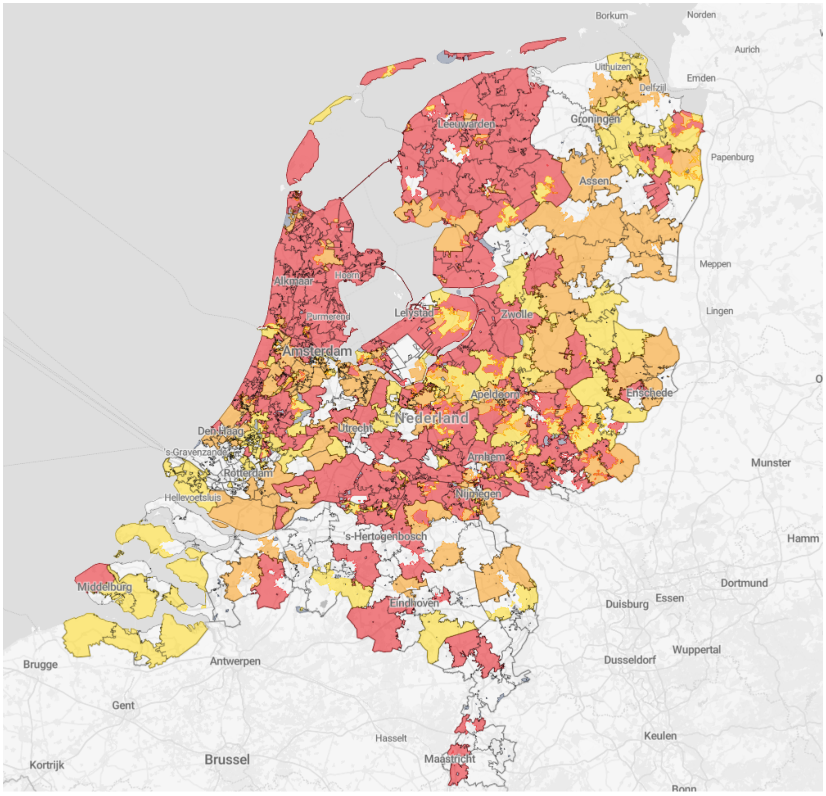


FIGURE 3.4: Distribution of off take grid congestion across regional electricity networks in the Netherlands. Source: [28].

This uneven distribution of grid congestion emphasises the need for targeted and case dependent interventions. For this research, the focus will be on Liander's MV grid, which is one of the most congested areas of the country. Unlike large industrial connections on the HV grid, the many medium-sized industries and installations on the MV grid often do not have the technical expertise or do not see themselves as flexibility providers. As a result, much potential remains untapped, whereas collectively, these could greatly reduce local overload [7].

In further sections, it will be explored how the existing mechanisms in the Netherlands could unlock this capacity.

3.3.1 Issues derived from Grid Congestion in The Netherlands

Before moving forward, it is essential to describe the problems caused by grid congestion in the Netherlands, as well as its wider economic and societal consequences to fully appreciate its seriousness.

Grid congestion has resulted in a range of challenges affecting various sectors. One of the most pressing issues is the **delay in renewable energy integration**. Currently, 3.54 GW of renewable capacity is waiting in line for grid connection. This lack of renewable integration not only slows down the country's sustainable transition but also translates to an estimated annual carbon cost between €329 million and €376 million, as fossil fuels continue to be relied upon for energy generation [29]. Taken together, the setbacks in the sustainable energy transition are estimated to cost the Netherlands roughly €1 to €1.5 billion per year [11].

The **industrial** and **commercial sectors** have also been severely impacted. As of march 2024, over 20,000 companies are currently waiting for grid access, unable to secure or expand electricity connections. This leads to substantial economic losses and hindering national economic growth and investment opportunities. The situation is particularly severe in areas like Amsterdam, where in 2023 alone, 9,396 customers were on a waiting list for grid connections, representing a combined capacity of 5,116 MW [30]. Moreover, as previously mentioned, electricity network rates are expected to increase in order to manage grid congestion, which will also have a direct impact on consumers [24].

Recent studies suggest that resolving congestion on the low and medium voltage grids could generate between **€10 billion and €35 billion annually** across the economy, sustainability, and infrastructure. These figures could increase further if grid congestion worsens in the coming years, leading to reduced investments from businesses. Additionally, for energy-intensive companies in particular, **grid congestion restricts sustainability efforts** and exposes them to rising carbon prices, all while grid costs continue to climb. This could prompt some companies to leave the Netherlands altogether; if just 5% to 10% of them relocate, it could cost the Dutch economy an additional €5 billion to €15 billion per year [31].

Moreover, grid congestion is **delaying new infrastructure development**, such as rail services or housing. For example, in Almere, also located within Liander's grid, the construction of 75000 new houses can be delayed until at least 2030 due to insufficient capacity for new connections [32]. The exact economic impact of these delays is difficult to estimate, but researches suggest it could add anywhere from €0.1 billion to €2.5 billion per year to the already substantial cost [31].

Beyond these economic and development impacts, grid congestion is already **directly affecting residents** in some towns. Research conducted in early 2024 identified six Dutch municipalities where residents face severe grid congestion related issues: Steenenkamer, Stokkum, Markenbinnen, Herten, Dedgum, and Diphoorn. Inhabitants of these areas report power quality issues, malfunctioning inverters, flickering lights, and extended waiting times for new electricity connections [33].

This shows that grid congestion is a structural issue that **affects all electricity grid users**, with extensive economic and environmental consequences. If not addressed, these challenges will continue to hinder renewable energy deployment, restrict industrial expansion, drive up costs, and threaten national climate targets.

Due to their strong commitment to green policies, the Netherlands and Germany are among the countries hardest hit by grid congestion. The solutions they develop and the lessons they learn will be crucial to other nations that are just beginning to see early signs of congestion and to those that want to prevent it altogether [10].

3.4 Dutch regulation and legislation of grid congestion

Now that grid congestion, its causes, locations, and economic and social impacts have been defined, we can look into what mechanisms are currently available to address it within the Dutch territory.

In The Netherlands, the regulation of grid congestion management is overseen by the **Autoriteit Consument & Markt (ACM)**, the Dutch national regulatory authority. Together with grid operators, governmental bodies, such as the Ministry of Climate and Green Growth (Ministerie van Klimaat en Groene Groei or KGG), and with Netberheerd Nederland, the trade organisation of all electricity and gas network operators, they collaborate to develop frameworks, incentives, and tariffs that encourage users to adapt their consumption based on grid conditions [34]. As mentioned in Section 3.3, in December 2022, the Ministry of Economic Affairs and Climate Policy (currently KGG) presented the LAN to the Dutch House of Representatives (Tweede Kamer). This program was drafted by grid operators, the ACM, central government, and market parties. One of the LAN's main objectives is to steer stronger towards a more efficient usage of the grid and clear the queues in congested areas [27].

In the Netherlands, (industrial) consumers can access electricity only once they have secured transmission capacity from their grid operator. This is formalised in a **connection and transmission agreement** (*Aansluit en Transportovereenkomst*, ATO), in which the customer specifies the contracted peak capacity. Under this framework, the Electricity Act (1998) requires grid operators to act in a non discriminatory manner, to apply a *first come, first served* principle, and to honour the terms agreed in each ATO [35]. In addition, all connected users pay regulated network tariffs. For medium and high voltage customers, the principal components are the contracted power (*kW gecontracteerd*) and the highest monthly peak (*kW max*), as mentioned in Section 2.2. Historically these arrangements were sufficient, however, as grid conditions have tightened and congestion has become more frequent, both network operation and customer options have evolved. Consequently, below we set out the measures that customers can use to contribute to congestion management both directly and indirectly.

Alternative Transmission Rights (ATRs) and Non-Firm ATOs (NFAs)

Until recently, **only firm ATOs** were available. Under this type of contract, the connected party is always guaranteed of their contracted transmission capacity. For instance, if an entity has secured a 10MW capacity, that portion of the grid is effectively reserved for them, regardless of being fully used or not.

From January 2024 the landscape began to change. Netbeheer Nederland submitted a proposal to the ACM to amend the electricity codes, introducing a framework for **alternative transmission rights (ATR)** to promote a more efficient use of the grid [36]. ATRs are transmission rights for which capacity availability is **not guaranteed at all times**. This framework includes:

- **Time-Duration-Based Transmission Right, *Tijdsduurgebonden transportrecht* (TDTR)**. With this tariff, consumers are guaranteed access to their contracted transmission capacity for at least 85% of the time annually, equating to a minimum of 7,446 hours per year. During the remaining 15% of the time, grid operators may limit access, especially during peak demand periods to manage congestion. This specific tariff is known as the **ATR85/15 tariff**. Currently, there are no plans to develop another variant where other percentages apply. Under this scheme, the contracted power (*kW gecontracteerd*) component of the grid tariff is waived, reducing the overall transmission costs for participants. Companies could receive an average grid tariff discount of 65% if they optimise consumption. TenneT will confirm TDTR access for the next day by 08:30, half an hour before the FRR (Frequency Restoration Reserve) market closes, allowing consumers to adjust their operations accordingly. The system operator also announces how much transmission capacity is available during the remaining hours of the day. For now, this scheme is limited to the national HV grid and is not available on regional networks: local demand patterns differ too much from one area to another to allow a single *% of hours per year* threshold to reliably prevent overload everywhere. Further application of time-based transmission rights at regional network operators is being investigated.

- **Time-Block-Based Transmission Right, *Tijdsblokgebonden transportrecht* (TBTR).** These rights give users access to transmission capacity in specific, predefined time blocks, outside these block, no capacity is assured. In contrast to TDTR, the usable hours are fixed and known in advance. Network operators apply structurally lower transmission tariffs to reflect this limited access. This arrangement is not offered on TenneT's transmission network because national load patterns vary in ways that do not allow safe definition of fixed time blocks. TBTR is therefore available only on medium voltage distribution grids.
- **Energy-Volume-Based Transmission Right**. This allows users to reserve network capacity according to a specified energy volume (e.g., MWh per year) rather than fixed time blocks. In theory, this would give consumers and producers flexibility to schedule their consumption or injection up to the contracted volume, potentially smoothing peaks and making more efficient use of the grid. However, it remains purely conceptual: the network operators have not submitted any code change proposal for this variant.

The Netbeheer Nederland code amendment introducing these ATRs was approved and, eventually, it entered into force on **1st April 2025**. From that date, TDTR and TBTR schemes are available to customers with connections above 3×80 A. For TDTR, TenneT may choose to offer the product until 1 October 2025, after which it is obliged to do so.

Additionally, **Non firm ATOs (NFA 1.0)**, have been available since 1 February 2024. These allow customers to access residual grid capacity on a day-to-day basis **without a guaranteed right to transmission**. The system operator evaluates the available capacity daily and allocates it accordingly. Customers with these rights must follow hourly power frames communicated at least one day in advance. This type of contract is primarily for users in congested areas who need flexible access to the grid and can accept lower security of transmission, ideal for businesses that prioritise flexibility over reliability. The key benefit is a zero contracted capacity rate (***kW gecontracteerd* = 0**), reducing costs.

It is interesting to note that **NFAs are always granted** by the DSOs. Since this type of contract does not claim transport rights, they are always accepted because they do not consume reserved grid capacity. By contrast, the *first come, first served* rule still applies to ATRs. The position in the queue does not change just because you choose an alternative transmission right instead of a fixed one, as they are treated like any other grid connection request [37].

Therefore, with the new rules there are now **four** types of transport contracts. The existing *standard* connection and transport agreement (**fixed ATO**), which gives a customer 24/7 transport rights to the contracted capacity; a fully variable transport agreement (**NFA 1.0**); a time-block-based contract (**TBTR**) and a time-duration-based contract (**TDTR**).

Capaciteitsbeperkings contract (CBC)

Another mechanism to combat congestion management are capacity limiting contracts or in Dutch, Capaciteitsbeperkings contracten (CBC). It is a **bilateral contractual agreement** between the grid operator and consumers where the consumer agrees to limit power consumption in order to manage grid congestion. There are different types of these contracts, each with specific conditions regarding response times, compensation, and flexibility requirements [38].

- **Call-based CBC**: requires the consumer to adjust electricity consumption or supply when notified by the grid operator, typically one day in advance, before the day-ahead market gate closes. In exchange for adhering to these agreements, the consumer receives financial compensation. For high-energy consumers, participation through a Congestion Service Provider (CSP) is mandatory.
- **Fixed CBC**: allows consumers to self-regulate their power usage at predetermined times without requiring real-time intervention from the grid operator. The conditions and compensation are set in advance and agreed upon directly between the parties.
- **Group CBC:** allows multiple companies to collaborate in reducing energy consumption when necessary. The group collectively decides how to distribute the capacity limitation among participants, and the grid operator provides compensation accordingly. This model provides greater shared flexibility while maintaining grid stability.

Redispatch (voluntary or mandatory bidding contract)

Redispatch is a **market-based tool** for relieving local grid bottlenecks by adjusting production or consumption on **short notice** (from a few hours to minutes before the affected time block). All transactions run through **GOPACS** (Grid Operators Platform for Congestion Solutions), which sits on top of the standard intraday trading venues (ETPA, EPEX SPOT). The platform was launched in early 2019 as a collaborative initiative by TenneT and regional distribution network operators. The biddings can be:

- **Voluntary bids** are placed intraday by any qualified market party, usually via a Congestion Service Provider (CSP). When a grid operator identifies a imminent congestion event, it issues a call for bids in GOPACS. Participants submit their available upward or downward flexibility, price and volume, without any prior obligation. GOPACS then ranks all offers by cost, accepts the cheapest or more convenient ones, and dispatches instructions. Providers are paid an energy fee (the bid price) once they deliver the adjustment [39].
- **Mandatory bidding contracts** require participants to commit their flexible capacity in advance and remain in force until specified in the contract. Under this agreement, a party agrees to respond to every congestion call with a bid at pre-agreed terms. When the operator forecasts a constraint, it triggers the mandatory bid process via GOPACS. Accepted bids receive both a capacity reservation fee and an energy activation fee; failure to deliver can incur penalties [40].

GOPACS itself states that participation is intended for *large consumers and other market parties* willing to adjust their generation or load intraday. Smaller users, households, small businesses, even many medium-sized firms, do not meet the minimum size or lack the contractual setup (via a Congestion Service Provider, detailed metering and trading platform access) to participate directly. In addition, interviews with Dutch market participants found that redispatch is seen as the most *effort-intensive* congestion management tool. Therefore, many potential providers prefer simpler mechanisms [41].

For these reasons, participation on GOPACS remains quite limited to this day. At present, only about 20 market parties are offering flexibility services to grid operators via the GOPACS platform [42].

Cable pooling

Cable pooling is a method for *sharing* a single grid connection among multiple energy installations by aggregating their contracted capacity into one point of access. Each participant (can be a solar park, wind farm, battery or other installation) has its own separate metering and settlement, but they all feed into or draw from the same physical cable. Because different technologies and sites often have complementary output or demand profiles (for example, solar peaks at midday while wind can pick up in the evening), cable pooling lets the total capacity be used more efficiently than if each installation had its own dedicated connection [43].

Direct Line

A direct line is a direct electricity connection between a producer and a customer, completely separate from the public grid. It allows, for example, a large consumer to relieve the grid by buying locally generated energy directly. A direct line needs to be reported to the ACM, but usually no separate permit from the grid operator is required [44].

Gesloten distributiesysteem (GDS)

A closed distribution system or in Dutch, Gesloten distributiesysteem (GDS) is a private electricity network connected to the public grid through exactly one connection. Within that shielded network, connected companies or installations can exchange energy among themselves and agree on maximum loads. The grid operator manages only the single connection point and the goal is to minimise the use of the public grid. A GDS requires an approval from the ACM, but no separate application to the grid operator [45].

Groepstransportovereenkomst (GTO)

A group transmission agreement, in Dutch Groepstransportovereenkomst (GTO), allows several customers to conclude a single transmission contract together with the grid operator. The group determines among themselves how the joint capacity is distributed, allowing individual peaks to be absorbed more efficiently. The large-scale application of GTOs is not expected before the fourth quarter of 2026 [46].

Meerdere Leveranciers Op Één Aansluiting (MLOEA)

MLOEA (Multiple Suppliers On One Connection, in Dutch, *Meerdere Leveranciers Op Één Aansluiting*) makes it possible to connect two or more energy suppliers to one physical connection by using an additional metering point. For example, a company can purchase power from supplier A and supply it back to supplier B, without the need for separate cable infrastructure. This offers more freedom of choice and can lead to lower grid costs and more favourable contract conditions [47].

MLOEA might sound similar to cable pooling, but the key difference is scope: MLOEA is a commercial arrangement (one physical connection split across multiple suppliers/BRPs), while cable pooling is a technical/connection arrangement (multiple generators share the same grid connection capacity). In simple words, MLOEA splits the bill; cable pooling shares the cable.

Gebruik Op Tijd Of Raak Het Kwijt (GOTORK)

Besides the possibilities mentioned above, the grid operators can implement the UIOLI measure, an abbreviation for *use it or lose it*; or in Dutch, *gebruik op tijd of raak het kwijt*, GOTORK. Grid operators will be able to purchase transport capacity that **has been contracted** but is **not used** for a long period and to offer it to other customers. For grid users connected to a voltage level of up to 110 kV, their contracted capacity can be reduced if they have not used more than 50% of their transmission capacity for a year, or if they have not used 1 MW or more for a year. For high-voltage customers from 110 kV, the aforementioned 50% rule also applies for a year, or the non-use of at least 10 MW of transmission capacity for a year [48].

However, some additional conditions still apply. If grid users can demonstrate that they will (re)use the unused portion of their contracted capacity within a reasonable period, the grid operator will not reduce the contracted capacity.

New Grid Tariff Scheme

Additionally, another measure proposed includes **time-based transmission tariffs** for the **high-voltage** grid. Off-peak power usage will be weighted at a reduced factor (0.6 or 0.7), while peak hours will remain at a factor of 1, effective from April 2025. The goal is to give users a clear price signal to shift demand away from system peaks and thereby relieve congestion and defer costly grid reinforcements. For medium-voltage networks, these time-based tariffs do not apply for now [49].

FIGURE 3.5: New TenneT grid tariff structure , source [49]

Other measures

In congested areas, network operators can determine whether a **participation obligation** is used, and in which direction this obligation applies (feed-in or off-take). The threshold for this obligation has been reduced from 60 MW to connections above 1 MW. The obligated companies must demonstrate how much of their contracted capacity can be used flexibly and must offer a price at which they are willing to limit their in-feed or load. If these market-based means are exhausted, the grid operator may switch to **non-market-based** congestion management. This means that participants are asked to adjust power for a fixed fee, set by the Grid Code, without the need for a marketplace mechanism. These instruments are intended as a last resort and are applied according to the stage of congestion management in each region [27].

In parallel, the rigid *first-come, first-served* allocation rule has been amended to allow network operators in congested areas to **prioritise specific transport requests**. Under the new framework, effective from October 2024 and mandatory for all operators, parties serving a major public interest receive precedence in capacity allocation [50]. The priority groups are:

- **Congestion alleviators** (e.g. grid-connected batteries) that free up additional network capacity.
- **Safety services**, including defence, police and emergency healthcare.
- **Basic needs**, such as drinking water, housing, education, heating and gas supply.

Operators may implement this prioritisation framework immediately if their systems are configured accordingly. Although framed as an operational change, this measure is fundamentally intended to mitigate the societal impacts of grid congestion.

Some other measures from the LAN that have not yet been implemented include a **transport tariff** for electricity **feeders**, at present, producers are now exempt from transport costs.

A summary of all these mechanisms can be found in Table 3.1.

3.5 Applicable Congestion Management Mechanisms

Now that the system has been described in detail and the Dutch congestion management options have been set out, we turn to selecting a mechanism that fits the case. As defined in Chapter 2, the WWTP is a process driven site with an inflexible demand profile and very limited scope to shift core loads. The choice must therefore favour instruments that are measured and enforced at the point of common coupling and that remunerate asset based flexibility without altering the treatment. On this basis, we identify the most suitable mechanism to test against the study facility.

Several options can be set aside at once. Measures that require group participation or shared infrastructure, such as Group CBC, cable pooling, direct lines, closed distribution systems, group transmission agreements, and multiple suppliers on one connection, are outside our scope because the analysis concerns a single, stand alone WWTP with no coordination with neighbouring users. Likewise, some alternative transmission rights and non firm access contracts either target high voltage users or introduce levels of uncertainty that are not compatible with continuous operation at this site.

What remains for an individual, process driven connection at medium voltage are plant level instruments that act and settle at the meter. These are the redispatch, the TBTR, and the CBCs. Redispatch is technically possible and the plant can provide downward relief by reducing net imports, however the lead times are short, the operational overhead of bidding and delivery verification is high, and local market liquidity can be limited. Voluntary participation could play a role, but as a primary tool it is unlikely to be practical for this facility for the same reasons. TTBTR offer guaranteed capacity only in predefined windows. Outside those windows there is no guarantee, which creates unacceptable exposure for a continuously operating plant and forces arbitrary block choices that do not align with the variable net load. Fixed CBCs are simpler to administer but require setting restriction periods well in advance. Given the site's variable net import, this risks constraining when not needed and missing relief when congestion actually occurs.

Call based CBCs appear as the right balance. Calls arrive day ahead, which is close enough to real conditions to be targeted, yet early enough to allow scheduling. The mechanism measures performance at the point of common coupling, pays for the asset based reduction the site can credibly deliver, and potentially leaves the treatment process untouched. Because the contract is bilateral, it can embed site specific safeguards. For these reasons we select call based CBC as the mechanism to be implemented and tested in the remainder of the thesis, and we now describe its representation and evaluation.

3.6 Call-based CBCs: In depth

Call-based CBCs, are **bilateral agreements** in which the DSO, may temporarily limit a participant's grid use during forecast congestion windows while leaving access unchanged at other times. The DSO issues a day ahead call for defined hours, the participant keeps its point of common coupling, PCC, within the agreed limit during those hours, and compensation is paid according to the contract. This makes CBCs a practical bridge while reinforcements are built, and a targeted way to free capacity for new or waiting connections without imposing continuous restrictions [38].

In operation the sequence is simple. First, limits and rules of the contract are agreed, direction (off-take or feed-in), the cap level, the seasons or hours when calls may occur, the maximum number of calls or maximum consecutive duration, the metering and validation rules, and the compensation formula. Next, on each day with expected congestion, the DSO sends a **morning notice** for the following day's hours. The participant then prepares assets so that, when the window starts, the measured import or export remains at or below the limit for every settlement interval. After delivery, the DSO validates performance from metered data and settles the agreed fees.

Measurement and settlement are anchored in what the meter at the PCC records during called intervals. This has two implications. First, compliance is technology neutral, the participant may use on site generation, storage, or controllable auxiliaries in any combination, provided the imports stay under the cap. Second, any ambiguity about missing intervals or data sources is resolved in the contract by specifying fallbacks, for example secondary side measurements, sub metering, field tests, or DSO estimation with clear rules [51]. These clauses keep settlement transparent and predictable.

Compensation is usually built from one or more elements that reflect how costs arise. An availability retainer can cover readiness and routine preparation. An activation fee per megawatt per hour pays for the delivered limitation during called intervals. For generators, especially subsidised renewables, an allowance for foregone revenue or lost support, for example SDE or guarantees of origin, is often included. For demand side participants the logic differs, compensation should match incremental costs of holding the cap, extra energy and tariff costs, lost arbitrage, wear and tear, and any operational adjustments needed to keep the process whole.

Feed-in CBCs are comparatively straightforward. Curtailment is a single control action on a known asset, settlement can refer to foregone sales and documented support terms, and the baseline is the unconstrained production profile. Demand CBCs are more involved on both sides. On the participant side, staying under the cap may require coordinating several assets against ramps, state of charge and short term forecasts, while keeping the underlying process unchanged. On the DSO side, there is no simple revenue loss proxy for a reduction in imports, so compensation must be tied to documented incremental cost at the site. For these reasons, demand CBCs benefit from careful tailoring, clear metering rules and modest operating policies that reduce residual risk [52]

A workable call-based CBC for an industrial site therefore has three design anchors. First, **clear windows** and **day ahead notice** so that preparation is feasible and forecast error at the call boundary is contained. Second, clear limits to keep the programme **realistic**. Third, a **compensation structure** that matches the real cost drivers on the participant's side and an unambiguous settlement process. With these elements in place, call-based CBCs can potentially deliver dependable, measurable relief at the times that matter, without forcing changes to core processes.

TABLE 3.1: Overview of congestion–management instruments (layout close to the original figure).

Category	Mechanism	Description
Alternative Transmission Rights (ATRs)	Time–Duration–Based Transmission Right (TDTR / ATR 85/15)	Access to capacity $\geq$ 85% of the year; waived kW–contract component; HV–only.
	Time–Block–Based Transmission Right (TBTR)	Guaranteed access in fixed time slots; MV–only; structurally lower tariff.
	Energy–Volume–Based Transmission Right	Reservation based on MWh/year rather than hours; conceptual only.
Non–Firm ATOs	NFA 1.0	Day–ahead access to residual capacity at zero contracted capacity rate; hourly frames communicated one day ahead.
Capacity–Limiting Contracts (CBCs)	Call–based CBC	Adjust consumption/supply when called (day–ahead notice); financial compensation.
	CBC without call–off	Self–regulated curtailment at predetermined times; pre–set conditions/compensation.
	Group CBC	Multiple companies share one CBC; group decides distribution; collective compensation.
Sharing & Dedicated Connections	Cable pooling	Multiple installations share one physical cable; individual metering; complementary profiles smooth peaks.
	Direct line	Point–to–point connection between producer and consumer; must notify ACM.
	Closed Distribution System (GDS)	Private network behind one public–grid connection; internal exchanges; ACM approval only.
	Group Transmission Agreement (GTO)	Multiple customers share one transmission contract; internal capacity allocation; not yet implemented.
	Multiple Suppliers on One Connection (MLOEA)	Two or more suppliers on one connection via extra meter; no extra cable needed; more choice/lower costs.
Redispatch	Voluntary bids	Intraday calls for flexibility; any qualified party via CSP; ranked by cost; paid energy fee.
	Mandatory bidding contract	Pre–committed bids until local network expansion; capacity and energy fees; penalties for non–delivery.
Post–Contract Utilisation	GOTORK	Automatically reclaims unused contracted capacity after one year of $< 50\%$ use (or $< X$ kW); can be deferred if the user proves future use.
Pricing–Based Incentives	New grid tariff scheme	Time–weighted tariffs on HV; price signals to shift load to relieve congestion.

Notes: HV/MV denote high/medium voltage. CSP = congestion service provider. ACM = Netherlands Authority for Consumers and Markets.

Chapter 4

Scientific Literature Review

4.1 Introduction and scope

In this chapter we review capacity limiting mechanisms for congestion management and situate them within the broader portfolio of tools available to distribution system operators and large users. We first introduce the mechanism landscape and the main design choices, then compare CBC with alternatives, and finally examine modelling approaches, industrial demand response in continuous processes, and implications for MV connected industries with largely inflexible demand.

CBCs belong to a family of capacity limitation mechanisms that restrict network access during congested hours while leaving users unconstrained otherwise. They sit alongside interruptible connections, static capacity subscriptions, and local flexibility markets. The literature suggests that no single instrument dominates, rather, fit for purpose design depends on location, timing, the predictability of congestion, and the controllability of connected loads [12, 53].

4.2 Mechanism landscape and taxonomy by lead time

A helpful organising frame is the lead time at which restrictions are set, that is, near real time, day ahead, or long term. Hennig *et al.* [12] compare three archetypes, their analysis is low voltage focused but conceptually transferable. First, interruptible connections apply near real time caps with a tariff discount. Second, day ahead capacity limitation, DA CL, in which the DSO publishes a reduced limit for the next day, this design is directly relevant to CBCs. Third, static capacity subscription, where the user subscribes to a kilowatt band and exceedances are priced. These instruments differ in user risk and system efficiency. Shorter lead times preserve normal access when uncongested, but increase curtailment risk and exposure to balancing costs. Longer lead times offer predictability, but restrict access even when the grid is actually uncongested [12].

4.3 CBCs, structure, incentives, and enforcement

We now focus on CBCs in their day ahead form. As mentioned before, the DSO announces a limit before the day ahead wholesale market gate closure, the user otherwise enjoys full technical capacity. Hennig *et al.* implement DA CL (CBCs) as a binary step for comparability, for example a five kilowatt ceiling from an eleven point five kilowatt nominal charger, encoded by an activation factor λ such that available power is $\lambda \cdot p$. They show that DA CL performs well when forecasts are accurate, and call for clear fallbacks when forecasts are wrong to avoid over restriction or emergency actions, an important design point for medium voltage deployments. Their case concerns households and electric vehicles with sixty minute steps, however the logic generalises to industrial feeders with hourly calls [12].

The same source notes that capacity limits can be modelled as hard or soft. In Dutch CBC practice, repeated non compliance can result in disconnection from the programme, which highlights that enforcement credibility is integral to the effectiveness of contractual caps [12]. The day ahead notice provides a sweet spot in the trade off between predictability, users can plan, and flexibility, users retain full access when uncongested.

4.3.1 Lessons from capacity subscription tariffs

A closely related instrument is the capacity subscription tariff (CST). Customers select a kilowatt band, for example two, four, or eight kilowatts, pay a fixed annual fee for that band, and face a per kilowatt hour exceedance charge or a band uplift later if they exceed. CSTs are tariffs, not contracts, and the literature discusses them against standard regulatory principles, cost recovery, economic efficiency, cost reflectivity, non discrimination and equity, and understandability and predictability, while emphasising revenue neutrality at the DSO level, the tariff reallocates who pays without changing aggregate revenue. An Alliander study cited in this strand uses stylised annual prices, for example €192, €252, and €480 per band, and a €0.50 per kilowatt hour exceedance fee to illustrate revenue neutral packages [54].

Two design details from CSTs are transferable to CBCs. First, the averaging window. CST compliance typically uses a fifteen minute averaging window, which smooths noise and enhances user controllability. CBCs that cap hourly imports could similarly define measurement windows, for example fifteen minute rolling averages within the hour, to reduce false breaches and encourage actionable responses. Second, the selection period. Monthly band selection offers user flexibility but invites seasonal upsizing and weakens cost reflectivity. Yearly selection better mirrors cost causation that stems from annual peaks. By analogy, CBC call design should consider how frequently the baseline capacity or eligibility conditions can be revised. Hourly CBCs are much finer grained than CSTs, therefore careful rules are needed to discourage strategic self selection while preserving fairness. Finally, who pays differs. CSTs are tariff instruments that reallocate DSO revenue. CBCs are contractual and typically involve DSO compensation for curtailed access, justified when those payments are smaller than the avoided cost of curative actions or accelerated reinforcements. This trade off between DSO payments and deferred congestion costs should be explicit in CBC design [54].

4.4 Positioning of CBCs relative to other tools

This discussion leads to a comparison with other practical instruments. Interruptible connections, in the German BNetzA model, impose strict, enforceable caps on specific flexible devices in near real time, typically remunerated via a tariff discount. They are powerful during emergencies, but provide limited predictability for process industries [12]. Local flexibility markets promise allocative efficiency, yet face thin markets, market power, and transaction cost pitfalls at distribution level. Several proposals require complex locational products, repeated auctions, and sophisticated DSO market coordination. Hennig [53] and Radecke *et al.* caution that DSOs may prefer lower price risk, rules based instruments when congestion is localised and predictable, which is precisely the niche CBCs aim to serve [53, 54]. DSO interventions at low voltage, for example electric vehicle charging set backs and on the day redispatch, show technical feasibility but require prevalent observability and control and may clash with BRP portfolios in segregated markets if large volumes are curtailed without coordination. This is another reason to favour day ahead, contract based limits for MV connected industries [55]. Network reconfiguration and preventive redispatch can relieve congestion but are DSO internal actions with finite headroom. They complement, rather than substitute, demand side contracts in chronically constrained areas [56, 57].

A recent Dutch TSO study that compares redispatch with CBCs helps to place CBCs among other options. When the redispatch market has enough bids (i.e., is liquid), redispatch can fix the same constraint with less total activation, so it often ends up cheaper, it lowers day-ahead (D-1) prices, and it reduces emissions compared with CLC-heavy approaches. CBCs are less targeted, so they usually require more capacity to be limited to get the same physical effect, which can push D-1 prices up and bring more conventional generation online. However, the same study also shows a limit: for off-take congestion, redispatch often does not work because there are too few demand-side bids [41].

4.5 Design choices for CBC implementation

CBC calls with hourly restriction profiles are intuitive and align with day ahead markets, yet fifteen minute metering is standard for industrial billing. Defining the measurement basis, for example an hourly cap evaluated on fifteen minute averages with intra hour tolerance, can reduce inadvertent breaches. The CST literature offers precedents for such averaging choices to enhance understandability and controllability [54]. To mitigate strategic behaviour while keeping programmes attractive, enrolment windows and notice periods, for example quarterly enrolment with monthly opt out and penalties, can emulate CST selection period lessons while respecting the operational reality that CBCs are activated only when the DSO forecasts congestion. Day ahead limitation effectiveness depends on forecast accuracy. The literature stresses clear fallback protocols, for example capped emergency hours, compensation uplifts if realised congestion deviates from forecast, or automatic relaxation rules, to avoid over restriction or unprepared curtailments [12]. Compared with near real time products, day ahead CBCs improve predictability, which is important for low staffed and automated sites, and let participants utilise full capacity when uncongested, which is important to preserving industrial output and safety margins [12]. To avoid BRP conflicts, the CBC activation timing and metering should match market programmes, day ahead nominations and imbalance settlement, which reduces unintended deviations. This alignment remains a common challenge for congestion management schemes that act directly on controlled portfolios [53].

4.6 Compensation, revenue neutrality, and DSO economics

Tariff based instruments are designed to be revenue neutral at the DSO level, reallocating network costs among users, whereas CBCs involve DSO outflows in the form of compensation. The economic rationale is that well targeted CBCs can avoid higher costs, including costly curative redispatch, accelerated reinforcements, or the socio economic cost of connection denials [11]. As CST studies illustrate with designed price bands and exceedance fees, designing the money flows explicitly against revenue neutrality or cost avoidance benchmarks is essential for regulatory acceptance and user trust [54]. Experience from performance based demand response programmes, for example ConnectedSolutions in the United

States Northeast, shows that straightforward eligibility, clear performance metrics, and predictable payouts drive adoption. These lessons are useful when calibrating CBC compensation to be substantial yet simple for industrial sites [57, 56].

Public price data are sparse, but indicative figures from BCG give a sense of scale: activation values in the *thousands* of €/MWh for new peaking assets and around €1,650/MWh for industrial demand-side cases, while DSOs often offer only a *few hundred* €/MWh. The latter may cover foregone export revenue for generators, but typically not the opportunity cost of unserved industrial consumption (reported at €1,200–€4,500/MWh) [11].

4.7 WWTPs characterisation

According to the APICS (American Production and Inventory Control Society) process industries are defined as businesses that add value by *mixing, separating, forming, and/or chemical reactions*; these may run in batch or continuous mode and typically require rigid process control and high capital investment. Fransoo & Rutten adopt this definition and distinguish two extremes: *batch/mix* versus *process/flow*. A *process/flow* manufacturer "produces with minimal interruptions in any one production run or between production runs" of materials with process characteristics (liquids, fibers, powders, gases). Wastewater treatment fits this template: the core sequence is a chemically/biologically driven flow process that must run 24/7 to keep effluent within permit limits, i.e., with minimal allowable interruptions. This places municipal WWTPs firmly on the *process/flow* end of the APICS spectrum and explains their baseline inflexibility [58].

Municipal water and wastewater systems use on the order of a few percent of national electricity (U.S. EPA guidance cites 3–4%), and within WWTPs, electricity commonly accounts for roughly 25–40% of operating expenses, making energy a top controllable cost. Typical electricity intensity spans about 0.5–2.0,kWh per m^3, depending on process configuration and effluent targets. At most plants, the big end-uses are aeration and pumping: together they often comprise about two-thirds to three-quarters of total facility electricity, with aeration frequently the single largest consumer. These facts motivate demand-side flexibility, but the same effluent-quality and public-health constraints that make WWTPs "process/flow" also bound how and when loads can move [59].

4.7.1 Where the flexibility lives in a WWTP

Despite the rigid core process, several subsystems provide actionable flexibility without violating permits when managed correctly:

- **Pumping with storage/equalisation.** Where influent or intermediate basins provide hydraulic buffer, plants can defer part of pumping and downstream processing from peak to off-peak times (subject to overflow constraints and odor/settling limits). Guidance and case work from LBNL highlight using available storage deliberately during demand-response (DR) events.
- **Aeration "oxygen banking."** DO setpoints can be raised ahead of a peak-price/peak-load window to build a short-term oxygen buffer, then safely lowered and blower power curtailed during the event, provided DO and ammonia/nitrite limits remain compliant. Reviews and DR case guidance document this strategy; modern DO control (including MPC/EMPC) further reduces risk and energy while maintaining effluent quality.

- **Solids handling scheduling.** Thickening, dewatering, backwash and ancillary processes can be shifted to off-peak hours with minimal water-quality impact if upstream inventories are adequate.
- **On-site generation and fuel/storage coupling.** Many medium/large WWTPs produce biogas in anaerobic digestion. Running CHP and using gas holders as short-term fuel storage enables peak-shaving and DR participation while keeping critical treatment loads supplied.

4.8 State of the art in congestion management modelling

We now review how congestion management is modelled and implemented, with an emphasis on recent developments. In many countries, pilot programmes and market mechanisms have emerged to leverage flexible demand for congestion relief. In the United Kingdom, several DSOs run local flexibility tenders via platforms such as Piclo where industrial and commercial customers bid to lower consumption when the network is under stress. In the United States, demand response has long been integrated into capacity markets and grid emergency programmes, essentially paying large consumers to drop load during peak events. These efforts are part of a broader smart grid shift, moving from a paradigm of simply expanding supply to one of actively managing demand. Researchers generally categorise congestion management mechanisms into market based interventions, tariff based incentives, and direct control actions [60]. Market based approaches include local flexibility markets or auctions in which the DSO procures congestion relief services from consumers or aggregators in the area of a constraint. Tariff based approaches use pricing signals to influence behaviour, such as dynamic network tariffs that rise during forecast congestion periods or special contract structures that incentivise lower peak usage. Direct control refers to technical solutions where the operator can directly curtail or stagger loads through remote control devices or agreements to interrupt certain equipment. All these methods have the same goal, which is to coordinate flexible loads to avoid simultaneous peak usage that would otherwise overload the grid.

At the distribution network level, a common modelling framework is to perform optimal power flow or power grid simulations that incorporate flexible loads as controllable elements. In such models, the DSO's problem is often formulated as minimising the cost of congestion management, or equivalently minimising overloads, by dispatching flexibility from available participants within network constraints. The flexibility offers can be modelled with their specific characteristics, such as maximum power, duration, and cost bids. Studies have examined different market designs through simulations, for example comparing an auction based local flexibility market with a fixed rebate contract in terms of DSO cost and participant welfare. These models help identify efficient pricing mechanisms and the necessary coordination with the transmission system operator, since distribution level actions can impact the wider system. Another important modelling aspect is uncertainty. Congestion events depend on uncertain factors such as renewable generation spikes or demand peaks. Therefore, state of the art approaches use stochastic optimisation or robust scheduling to ensure reliability, even if some flexible loads under deliver [61]. Despite growing interest in congestion management mechanisms, empirical and modelling research focusing specifically on capacity limiting contracts remains limited, including in the context of individual industrial or infrastructure facilities in the Netherlands. The literature contains few investigations that examine the detailed implementation of such contracts, with most attention focused on broader instruments such as flexibility tariffs or local markets. One exception is the work by Hennig *et al.*, who model congestion management mechanisms that limit flexible loads based on capacity limitations, including interruptible, day ahead limitation, and static subscription schemes in a Dutch grid context. Their analysis remains at a system wide level and does not drill down into facility scale application, and no peer reviewed modelling studies seem to be published that demonstrate CBCs applied to specific installations such as wastewater treatment plants or industrial sites in the Netherlands [12].

4.9 Modelling approaches for CBC evaluation

We now focus on modelling approaches and the treatment of uncertainty that are suitable to test CBC applicability at facility scale and to analyse system level performance. On the network side, the distribution level congestion literature spans rules based limits, market based redispatch, topology reconfiguration, and coordinated control. Reviews emphasise that high distributed energy resource penetration and limited observability complicate real time congestion relief, and that preventive, schedule oriented tools, such as CBCs, can be effective when congestion is predictable in time and location [57]. On the participant side, industrial models must capture process constraints and operational coupling. For continuous process plants, rigorous optimisation, including MILP and model predictive control, is often needed to ensure feasibility and to quantify flexibility envelopes without compromising quality or safety. This is especially true for wastewater treatment plants, where service obligations and limited buffer capacities sharply bound the available actions [62].

Under uncertainty in demand, renewables, and prices, model predictive control provides a natural rolling horizon framework to respect network caps while minimising energy and balancing costs. The congestion management literature shows both centralised and distributed MPC formulations that incorporate network constraints, for example line limits, and align with market timelines, however deterministic MPC can be weak when forecasts contain errors [63]. Recent work introduces chance constrained MPC to enforce line flow or state constraints probabilistically, satisfied with high confidence, which shifts uncertainty from the objective into constraints to balance performance and risk. Although demonstrated in microgrids, the formulation carries over to medium voltage feeders serving industrial nodes, offering a principled way to model forecast risk in CBC compliance planning and to set risk adjusted operating policies, for example reserve margins inside the cap to reduce breach probability [64].

MPC also aligns with day ahead CBC timing. CBCs publish tomorrow's limits before day ahead gate closure. A rolling horizon controller optimises over the day ahead horizon given announced caps and price forecasts, applies only the first step, and reoptimises as caps or forecasts change. Predictive controllers have been formulated around day ahead horizons and multi period costs, including mixed integer variants to capture on or off and shiftable decisions, which is directly analogous to industrial sites that must schedule discrete equipment modes [65]. A core strength of MPC is explicit constraint handling. Hard engineering limits, including power, ramp rates, and state of charge, are enforced directly in the optimisation. This is why MPC has been adopted for real time congestion control via generator or battery redispatch, where line flow limits, actuator bounds, and rate limits are enforced at each step. The same machinery enforces a CBC capacity cap at the grid connection while respecting on site asset limits such as combined heat and power and storage [66]. Economic MPC naturally combines energy charges with peak and demand penalties in one objective. Field tests in buildings explicitly minimise energy cost and demand charge peaks with time varying tariffs, which provides an objective structure transferable to industrial Dutch tariffs and CBC settlement, for example minimise day ahead energy, taxes, and tariffs while obeying the CBC cap and avoiding breach penalties [65].

Two complementary threads strengthen uncertainty aware compliance. First, probabilistic overload assessment uses fast, non scenario methods to estimate probabilistic line loading by combining a deterministic forecast with a learned error distribution, for example via kernel density, which gives DSOs risk metrics at medium voltage level. Embedding these ideas in an MPC testbed helps choose risk aware CBC caps and participant reserve margins [67]. Second, chance constrained MPC requires compliance with a chosen probability, for example ninety five per cent, which provides a natural way to manage breach risk under forecast errors. Recent congestion management work shows CC MPC maintains flows within limits across small and large systems under consumer related uncertainties, which is an analogue to keeping a CBC participant inside its cap despite stochastic disturbances [64]. Where multiple sites share a feeder limit, distributed MPC provides a principled way to coordinate many participants or their aggregators while honouring feeder capacity limits and market roles, for example BRP and DSO, as demonstrated in hierarchical formulations for distribution grids. This makes MPC a credible laboratory to test multi party CBC activation without resorting to fully centralised control [63, 65]. Beyond buildings and microgrids, MPC has been applied to zonal congestion with PTDF based models where external network effects enter as disturbances. Storage and renewable curtailment are scheduled to keep branch flows under limits, showing that predictive control can maintain network feasibility using only local zone models, which is similar to a site responding to a CBC cap at its point of connection [68]. A complementary line of work studies robust planning for reinforcements under deep uncertainty, for example IGDT, which shows that operational tools such as CBCs should sit within a portfolio that also includes strategic upgrades. The two are not substitutes, they are staged complements across different timescales [53].

4.10 Synthesis

After reviewing the literature, a clear picture emerges. CBCs are a practical, near-term tool for local and time-predictable congestion. They keep full access when the grid is uncongested and give day-ahead certainty when caps apply. Their effectiveness depends on three basics: good forecasts with clear fallbacks, credible enforcement, and alignment with market and BRP processes for metering and settlement [12, 53].

A municipal WWTP is a *process/flow* facility in the APICS sense and must run with minimal interruption to meet permits, so its baseline demand is largely inflexible [58]. Even so, studies show usable flexibility beyond biogas/CHP—e.g., pump scheduling with storage, short “oxygen banking” via aeration control, and shifting solids handling—while keeping effluent within limits [62, 59]. In this work we keep the model simple and scalable and focus on the grid-facing cap, but these options indicate extra headroom that could be added later.

For compensation, CBCs are not revenue-neutral tariffs; they involve DSO payments. A fair approach is to compare a *no-CBC* baseline with a *CBC* case and pay pay the gap between the expected total cost without CBC and the expected system cost with CBC. Making these money flows explicit supports regulatory acceptance and user trust [11, 54].

The literature has shown that model predictive control is well suited to CBC contexts with day-ahead caps and forecast uncertainty. MPC optimises on a rolling horizon that aligns with D-1 publication of limits, and it enforces hard constraints at the point of connection and on-site assets (power, ramp rates, state of charge). However, the literature still reports few studies that model and validate CBCs end-to-end for a single facility [65, 63]

Chapter 5

Methodology

This chapter sets out the methodological framework used to quantify the operational and economic effects of capacity limiting contracts at the wastewater treatment plant. The aim is to build a tractable plant level model that captures the physical assets, tariff terms, metering and verification rules, and the contract logic. Section 5.1 sketches the study stages. Section 5.2 defines the system, inputs, horizons, and tariffs. Section 5.3 details Stage A–C experiments; decision rules appear in Section 5.4. Assumptions are collected in Section 5.6; tools and reproducibility in Section 5.7; validation checks in Section 5.8.

5.1 Overview and roadmap

This study considers a process-driven facility whose electricity demand we consider effectively inflexible. As a result, flexibility, and the success of the mechanism under study, must come primarily from on-site assets (the biogas plant and the biogas tank) and from the coordination of imports and exports in response to market prices and network tariffs. Our aim is to design and test a call-based CBC mechanism that (i) reliably honours all curtailment calls, (ii) provides the DSO with the maximum feasible reduction at an efficient price, and (iii) remains economically efficient for the WWTP. Throughout, we focus on *denting* grid imports, that is, delivering real reduction in grid imports, so that the transformer sees a measurable change, and therefore actual congestion relief. All decisions are benchmarked against a **no-CBC** baseline using identical exogenous inputs.

5.1.1 Staged approach

For de modelling we use two approaches. First, the **Perfect Foresight**, where the optimiser knows all inputs over the full horizon, referred from now on as PF. Secondly, the **rolling-horizon MPC**, introduced in Chapter 4 and referred to as RH, the optimiser sees only a limited look-ahead window, implements the first step, then updates forecasts and repeats. The way these approaches are implemented is described in detail in the remainder of the chapter.

We adopt a staged approach that balances four aims: (i) **baseline characterisation and capability mapping** (establish a cost benchmark and reveal the plant's intrinsic flexibility in the absence of calls), (ii) **design** (Perfect Foresight screening across CBC caps and BESS sizes), (iii) **fidelity** (RH realism with finite look-ahead and intertemporal state carry-over), and (iv) **robustness** (RH with forecasts under imperfect information, complemented by exogenous stress tests).

The study therefore proceeds in the following stages:

STAGE A - Baseline characterisation and capability mapping

1. **PF benchmark (no calls).** We first compute the cost-optimal dispatch of the WWTP against EPEX prices and grid tariffs. The Perfect Foresight (PF) model is the *standard* formulation with full knowledge of loads, on-site energy availability, and prices over the entire horizon. The results serve as the baseline against which all subsequent models in **STAGE B** are compared to.

2. **PF yearly cap (no calls).** On this baseline, we test the lowest achievable *yearly* contracted-capacity reduction. This delivers an operational threshold that informs subsequent call-based designs: caps set *above* this threshold will likely be non-binding, whereas caps set *below* it are expected to bind, this is where we want to put the focus on.

3. **PF grid-imports discouraged (no calls).** Finally, for this stage, we test WWTP operation under *typical* congestion periods in which grid intake is completely discouraged (i.e., assigned a prohibitive penalty). On this basis, we quantify the intrinsic flexibility *prior to* any CBC calls. This provides an upper bound on achievable import reductions, a first diagnostic of how much flexibility the current layout can deliver.

STAGE B - Design

1. **PF screening with MC calls.** We then introduce Monte Carlo (MC) call schedules and systematically explore different cap levels and battery energy storage system (BESS) sizes. Similarly as before, it yields an optimistic benchmark for what is achievable when calls are perfectly known and assets can be optimally prepositioned around them. The sweep identifies designs that maximise call feasibility and deliver CBC compensation efficiently.

STAGE C - Fidelity and robustness

1. **RH validation with MC calls.** Next, we validate operational feasibility with the integration of the CBC design chosen from the previous stage with a rolling-horizon optimisation (i.e., a Model Predictive Control formulation). The same CBC calls are introduced but now under the notice typically available in practice: around 08:00 on the day before the call, ahead of the day-ahead market gate closure at 12:00 [69]. Intertemporal states are carried across steps. The RH dispatch is necessarily suboptimal relative to PF but more realistic.

2. **RH with forecasts and MC calls.** Then, we introduce forecasts: within each window one portion *sees* actual data, and the remaining hours use forecasts for demand, solar, and biogas (EPEX prices remain day-ahead known, just like in reality). This variant quantifies robustness under imperfect information. If compliance degrades, we evaluate operational policies targeted at honouring the maximum number of calls.

3. **Stress testing.** With the same RH with forecasts model we stress test the optimal design against three exogenous scenarios: (i) EPEX price spikes (top 10% of prices raised to emulate very high marginal generation costs), (ii) increased demand, and (iii) higher grid tariffs. These tests assess whether feasibility and economics are preserved under adverse but plausible conditions.

5.1.2 Testing periods

The **Stage A** PF experiments are run over a full year to characterise upper-bound potential and to supply thresholds that guide call-based scheme design. **Stages B and C** (PF and RH with MC calls) focus on a stress week, **05–03–2023 00:00 to 11–03–2023 23:45**, identified as the period with the highest electric demand in the data set. Concentrating the tests on this week reduces computational burden considerably but results remain conservative: solutions that are feasible under the most demanding conditions are expected to remain feasible in less demanding weeks. Restricting plots and diagnostics to a week also improves readability.

5.1.3 Objective

Across all models, the optimisation minimises total electricity cost to the WWTP, enforcing the same technical and safety constraints of on-site assets. The cost function explicitly accounts for energy-market terms (import costs and export revenues) and grid-tariff terms (fixed monthly fee, contracted-capacity charge, volumetric transport fee, and monthly peak charge). This is a deliberate new addition, since currently the WWTP does not optimise for grid tariffs. Internalising these charges improves the plant's econonomic efficieny. Other minimal variations to the objective are included, but will be discussed in the next chapters.

Table 5.1 maps each model to its purpose, key assumptions, and outputs (including the benchmark), and Figure **??** illustrates the flow between stages and models and how elements from one, feed the next. These elements are introduced here for orientation and are developed in detail in the following sections.

5.2 System and data

This section sets out what the models *see*: the system representation (nodes and directed energy flows), the temporal resolution and horizons, the assets with their operational limits, and the applicable grid tariff structure. Elements that are *stage specific* are introduced in dedicated subsections: forced congestion windows in **Stage A**, Monte Carlo (MC) call schedules in **Stages B–C**, and forecasted inputs within the rolling horizon in **Stage C**.

5.2.1 System representation

The optimisation model reproduces the real behaviour and interactions of the facility with minor adaptations to improve clarity and computational tractability. The plant is described as a network of nodes linked by energy flows. The WWTP demand node draws electricity from four sources: (i) the solar park, (ii) the gas plant, (iii) the external grid, and (iv) an *unmet demand* slack that activates only when the others are not enough. The solar park is a producer node that may (i) supply the plant, (ii) export to the grid, or (iii) spill into a dedicated *solar sink* representing curtailment. The biogas stream can be routed (i) directly to the CHP, (ii) into the gas tank, or (iii) to a *gas sink* included purely as a modelling safeguard. The gas plant can (i) supply the WWTP or (ii) export to the grid. The grid acts both as source (imports) and sink (exports), each constrained by the corresponding contractual limits. These behaviours are enforced through linear energy-balance equations at every node and time step.

Figure 5.1 sketches the nodes and flows used in the optimisation, and Table 5.2 summarises the entities and their roles.

5.2.2 Data sources and inputs

The models rely on input data and parameters that are exogenous to the optimisation but shape its outcomes. The core operational series are measured **2023** time series at **15-minute resolution** for plant electricity demand, solar production, and biogas production; these are provided directly by the facility. Purchase and sale prices are historic EPEX day-ahead

TABLE 5.1: Model roadmap and taxonomy: purpose, key assumptions, and primary outputs.

Stage	Model	Purpose	Key assumptions	Primary outputs
A	*PF benchmark*	Establish a cost baseline and characterise normal plant operation.	PF; no calls; full-year horizon.	Baseline total cost; reference for later comparisons.
	PF yearly cap	Identify the lowest feasible *yearly* contracted capacity level.	PF; no calls; full-year horizon.	Threshold (kW) for feasible yearly cap; guidance on which caps are likely to bind.
	PF grid imports discouraged	Map intrinsic flexibility under *typical* congestion periods.	PF; no calls; penalty on imports during selected congestion hours; full-year horizon.	Upper bound on achievable import reduction; reliance on on-site assets; diagnostic of flexibility prior to any calls.
B	*PF sweep with MC calls (current layout)*	Design screening without additional flexibility; identify feasible CBC caps under existing assets.	PF; MC call schedules; no BESS; different CBC caps [kW]; stress week.	Acceptance rate by cap; infeasible patterns; incremental cost vs. PF benchmark; normalised price;
	PF sweep with MC calls + BESS	Design screening with a flexible asset; quantify how integrating a battery affects feasibility and cost.	PF; MC calls; add BESS sizes; same caps and stress week.	Acceptance rate vs. (cap, BESS size); cost and normalised price; Pareto frontier (price vs. reduction); selected candidate pair for RH validation.
C	*RH validation with MC calls*	*Fidelity*: verify operational feasibility with finite look-ahead and state carry-over.	RH; same MC calls; practical notice (call known day-before); no forecasts; stress week.	Acceptance rate vs. PF; feasibility; cost gap vs. PF; operational realism.
	RH with forecasts and MC calls	*Robustness*: quantify performance under imperfect information.	RH with forecasts: commitment slice uses actuals; remaining window uses forecasts (EPEX day-ahead known); same MC calls and notice; stress week.	Feasibility under uncertainty; cost sensitivity vs. PF/RH; effect of forecast errors; evaluation of operational policies.
	Stress testing (RH with forecasts)	Test resilience under exogenous shocks.	Same RH-with-forecasts set-up; scenarios: (i) EPEX price spikes, (ii) increased demand, (iii) higher grid tariffs.	Feasibility under shocks; total & incremental cost vs. baseline; impact on CBC compensation.

(2023), resampled to the same temporal resolution and retrieved via a market-data vendor through *Repowered*.

Contractual grid import and export limits, together with the operational limits of the gas plant and the biogas storage tank, are also supplied by the WWTP. Likewise, the biogas to electricity conversion factor, $\phi = 2.10032\,\mathrm{kWh/m^3}$, is provided by the facility and reflects their realised conversion, including efficiencies and losses.

Asset-level **costs** are set on the basis of standard operational accounting. The on-site PV

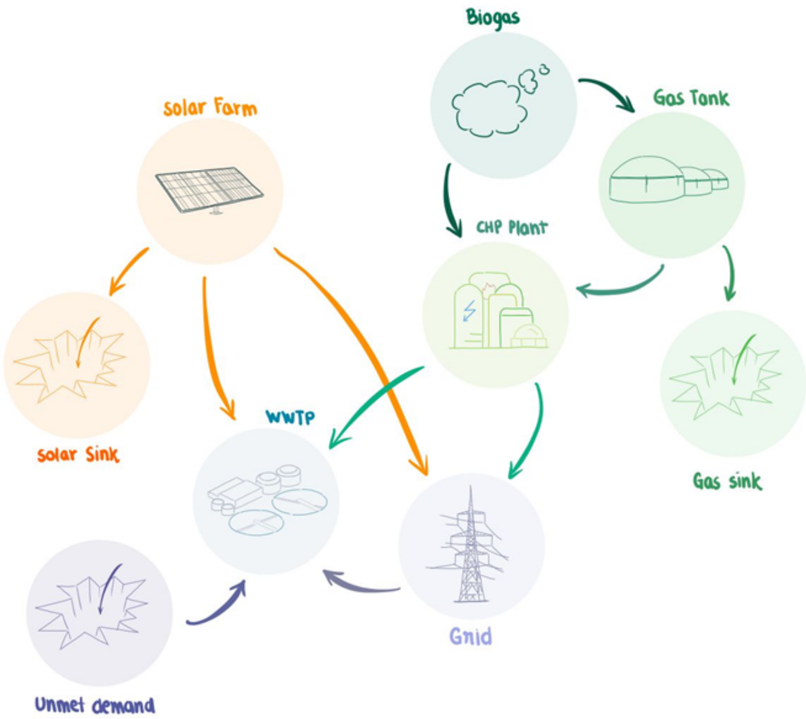


FIGURE 5.1: System representation in model

TABLE 5.2: System entities and their role in the optimisation model.

Category	Asset
Consumer	WWTP
Producers	Solar Park
	Biogas Unit
Storage	Biogas Tank
Conversion	CHP
Sinks	Electricity Grid (export)
	Gas Sink
Sources	Electricity Grid (import)
	Unmet Demand

installation is assigned zero marginal cost. This reflects a focus on marginal costs, the costs that actually change as a result of day-to-day operational decisions. Since the solar installation is fully owned by the WWTP and the investment has already been made, the main ongoing costs (such as fixed O&M) do not depend on whether the solar electricity is used at any particular moment. Therefore, the marginal cost of using solar energy, once the panels are installed is effectively **zero**. The same reasoning applies to the gas plant. Biogas is produced as a by-product of wastewater treatment regardless of its use, so the marginal cost of generating electricity with the CHP is also set to zero. In the same way, use of the *existing* biogas storage tank is treated as cost-free in operation; only if additional capacity were installed would a storage cost be added. Finally, the cost of solar curtailment is also assigned zero value, because any opportunity cost is already internalised by the optimisation in the objective function. As these components are null, they are excluded from the mathematical formulation in the following chapters.

However, virtual sinks (a gas sink, and an *unmet demand* slack) carry very large penalty values so that the optimiser uses them only when no feasible alternative exists. The absolute magnitudes are chosen by us; they are arbitrary in level but deliberately high. In all meaningful solutions these flows should be zero.

The energy tax on imports is taken from the Netherlands Tax Administration [70], and grid tariffs from the 2023 structure used by the facility (Liander) [14]. Further details and the implementation of these charges are provided in Section 5.2.4.

A consolidated summary of all inputs and parameters is presented in Table 5.3.

TABLE 5.3: Inputs and parameters used in the models (symbols as used later).

Item	Symbol	Value / Notes
WWTP electric demand	D_t	kWh/step time series
Solar production	PV_t	kWh/step time series
Biogas production	BG_t	kWh/step time series
EPEX day-ahead price	P_t^{da}	€/kWh
Contracted Grid Capacity (off-take)	κ_{imp}	1290 kW
Contracted Grid Capacity (feed-in)	κ_{exp}	1236 kW
Gas plant rated power	$\text{GasPlant}^{\text{max}}$	450 kW
Gas plant ramp limits	$r_u,\ r_d$	$\pm 5\%/min$
Biogas→electricity conversion factor	ϕ	2.10032 kWh/m^3
Tank capacity	C_{tank}	400 m^3 kWh
Tank SoC bounds	$\text{SoC}^{\text{tank}}_{\text{min / max}}$	$[40\%, 75\%] \times C_{\text{tank}}$
Tank initial SoC	$\text{SoC}^{\text{tank}}_0$	$50\% \times C_{\text{tank}}$
Tank (dis)charge rate caps	$t_u,\ t_d$	50 kWh/step
Energy tax on imports	c^{tax}	0.03942 €/kWh
Volumetric import tariff	c^{vol}	0.0160 €/kWh
Fixed transmission fee	c^{fix}	€ 36.75 per month
Contracted-capacity fee	c^{cap}	€ 1.82 per kW·month
Monthly peak charge	c^{peak}	€ 2.59 per kW·month (on month max)
Unmet demand penalty	M_{unmet}	very large
Biogas sink penalty	M_{bio}	very large
Solar curtailment penalty	M_{pv}	0

Notes: Time step $\Delta t = 15$ min. Where relevant, power limits in kW are converted to per-step energy bounds using Δt (e.g., 1 kW ↔ 0.25 kWh/step).

5.2.3 Model horizons

As mentioned before, the study uses consistent 15-minute time steps across all stages, but the optimisation *horizons* differ by purpose.

Stage A runs over the full year (2023) to characterise baseline operation and the plant's intrinsic capability in the absence of calls.

Stages B and **C** concentrate on a *stress week*, from **05–03–2023 00:00** to **11–03–2023 23:45**, which is identified as the week with the highest electricity demand, this can be easily seen in Figure 5.2.

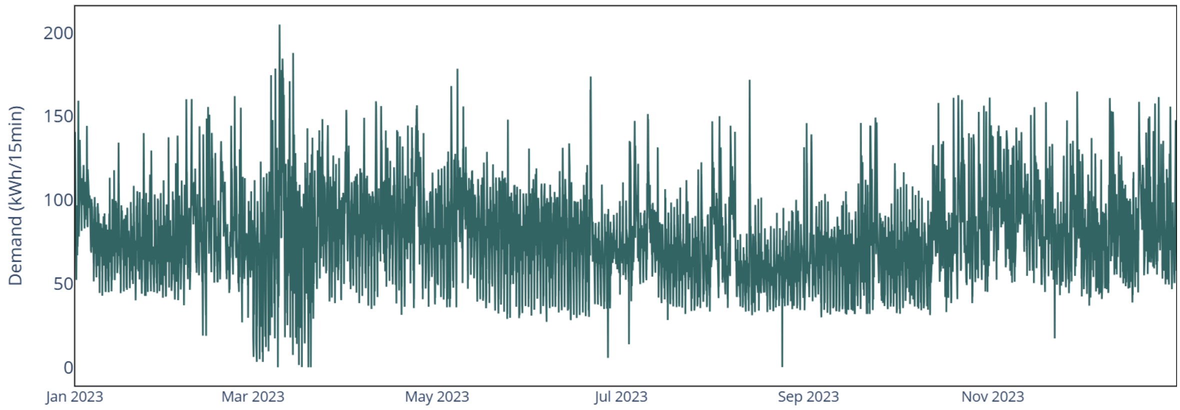


FIGURE 5.2: Yearly WWTP electricity demand (kWh). The stress week used in Stages B–C corresponds to the period with the highest aggregate demand.

In **Stage C** that uses the RH model, the temporal set-up is:

```
step_size     =  4 steps per hour (15 min)
window        =  96 steps (24 h)
recede_steps  =  12 steps (3 h)
```

That is, each optimisation looks 24 h ahead and commits the first 3 h, after which the horizon rolls forward and the problem is re-solved. The rationale of these values is that the 24 h look-ahead aligns with how information arrives and how assets behave. Day-ahead (DA) EPEX prices are known one day in advance, and CBC calls are notified the day before; a full day window therefore lets the optimiser anticipate price patterns, upcoming calls, and biogas/PV availability over an entire cycle.

The 3 h scheduling horizon (recede steps) serves two purposes. First, a 3 h commited window smooths control actions and avoids rapid switching, but still allows frequent reoptimisation as new information arrives. Computationally, solving a 96-step linear programme every 3 h (eight times per day) keeps run-times practical. Secondly, it aligns with the forecast set up in **Stage C**, within each 24 h window the first 3 h use actual data (and are fixed), while the remaining 21 h rely on forecasts. Fixing exactly this *actuals slice* prevents the optimiser from revising actions already executed when forecasts update. Taken together, the 24 h/3 h configuration balances realism, tractability, and the ability to position storage ahead of calls.

5.2.4 Grid prices and tariffs

The plant's electricity bill has two parts: (i) energy bought from or sold to the grid, and (ii) regulated network tariffs. The modelling follows the tariff structure actually faced by the site in 2023 to preserve realism.

At each time step, grid purchases are priced at the EPEX day-ahead price *plus* an energy tax obtained from the Dutch tax authorities [70]. In practice this tax is banded by annual consumption, with higher bands having lower rates (Table 5.4).

As reported in Table 2.1, the total electricity demand is approximately 2.7 GWh per year. Even allowing for on-site solar and biogas generation, this level places the plant in the higher

TABLE 5.4: Netherlands electricity energy tax (2023) by annual consumption band. All amounts in €/kWh. [70]

Consumption band	**0–2,900 kWh**	**2,901–10,000 kWh**	**10,001–50,000 kWh**	**50,001–10 million kWh**	**>10 million kWh**
Tax rate [€/kWh]	0.12599	0.12599	0.10046	0.03942	0.00115

tax bands. To avoid unnecessary complexity from the stepped schedule, the model therefore applies a single constant tax of 0.03942 €/kWh for 2023, which is conservative for this consumption level. Formally,

$$p_t^{\text{buy}} = P_t^{\text{da}} + c^{\text{tax}} \quad \text{with} \quad c^{\text{tax}} = 0.03942\,€/\text{kWh}.$$

Exports are remunerated at the EPEX DA price,

$$p_t^{\text{sell}} = P_t^{\text{da}},$$

so buying is always strictly more expensive than selling. This asymmetry reflects reality and also prevents buy–sell cycles (simultaneous import and export in the same interval), which would be physically meaningless.

In addition to these market terms, the network bill includes (i) a fixed monthly fee, (ii) a contracted capacity charge proportional to the import contract (kW), (iii) a volumetric transport tariff (€/kWh) on imports, and (iv) a monthly peak charge (€/kW) based on the highest instantaneous import within the month. All coefficients used follow the 2023 Liander schedule (Table 2.2). Unlike current practice at the facility, these tariff components are *internalised in the objective function* so that operations are optimised economically with respect to them; this is important because, as seen previously, tariffs greatly affect the total bill.

This treatment is especially relevant when a battery energy storage system (BESS) is present (**Stage B**). A battery tends to charge aggressively in low-price periods and discharge when prices are high. If monthly peak charges were ignored, this behaviour could create expensive import peaks, that could erase gains, and also generate sharp demand spikes that are undesirable from a grid perspective. By including all tariff terms explicitly, the optimiser trades energy purchase against peak exposure, which results in a dispatch that is economical and grid-friendly.

5.2.5 Stage-specific inputs

This subsection collects inputs that appear only for the sake of the experiments and are not common to all stages.

Stage A: PF grid imports discouraged

To map intrinsic flexibility prior to introducing calls, **Stage A** includes one experiment in which grid imports are **discouraged** during *typical* congestion periods by assigning a penalty in those time windows. What constitutes a *typical congestion period* can vary a lot per location, however, certain hours are systematically more exposed [71]. In line with this, the set of hours treated as congested in the experiment is listed in Table 5.5. All other inputs remain exactly as described previously.

TABLE 5.5: Congested weekday windows for time-selective import limits

Month	Congested Windows	
	Morning	Evening
January	07:00–10:00	16:00–20:00
February	07:00–10:00	16:00–20:00
March	07:00–10:00	17:00–20:00
April	08:00–10:00	17:00–19:00
May	—	18:00–21:00
June	—	18:00–21:00
July	—	18:00–21:00
August	—	18:00–21:00
September	—	17:00–20:00
October	07:00–10:00	17:00–20:00
November	07:00–10:00	16:00–20:00
December	07:00–10:00	16:00–20:00

Note: "—" indicates no congestion for that month. Only weekdays are considered.

Stages B–C: MC generation of CBC calls and BESS introduction

From **Stage B** onwards, we introduce call-based CBC constraints over the **stress week**. Unlike the Stage A discouragement, which consistently spans the same consecutive blocks, these calls occurence and intensity varies across run. The capacity-limit level (kW) that these calls enforce can not be defined yet, as its value is informed by the **Stage A** yearly-cap analysis and will be described in Chapter 6.

However, to build the schedules on which capacity reductions apply, we use information from a Dutch DSO (retrieved via *Repowered*) based on forecasted congestion studies that present the more likely hours of occurrence and typical episode lengths. We then generate a battery of **100** call schedules via Monte Carlo sampling so that the sweep in Stages B and C spans a broad range of cap timings and durations.

Concretely, we start from two empirical distributions compiled from the yearly observations in the DSO study: (i) a distribution for *start hours* (how often a congestion episode begins at each clock hour), and (ii) a distribution for *episode lengths* in whole hours (1–6 h). Because congestion is not expected to begin in at dawn, we set the weights for 01:00–05:00 to zero. For each calendar day in the target week, the number of call episodes, N, is drawn from a Poisson distribution with rate λ equal to the total monthly episode count divided by the number of days in that month. Given N, we pick N start hours using the hour-of-day weights and assign each a length drawn from the empirical distribution. A simple overlap check prevents coincident episodes on the same day. Repeating this procedure yields 100 distinct schedules that are stored and used in **Stages B** and **C**.

Each call record contains a **block identifier**, to make it easier to trace specific results, a **start time**, an **end time**, and a **call announcement** time. In line with practice, calls are announced the day before at **08:00**, before the day-ahead market gate closure at **12:00** [51]. This notice is irrelevant under perfect foresight (Stage B) since the optimisation already *knows* the call. However, it is crucial for the rolling-horizon formulation (Stage C), where a call only becomes visible once announced. An example schedule for one Monte Carlo run is shown in Table 5.6.

In addition to testing caps under the *current layout*, Stage B also repeats the sweep with a

TABLE 5.6: Example calling schedule for the stress week

block_id	start_dt	end_dt	call_dt
20230306_0900_MC1	2023-03-06 09:00:00	2023-03-06 11:00:00	2023-03-05 08:00:00
20230307_1700_MC1	2023-03-07 17:00:00	2023-03-07 18:00:00	2023-03-06 08:00:00
20230307_2000_MC1	2023-03-07 20:00:00	2023-03-07 22:00:00	2023-03-06 08:00:00
20230308_1100_MC1	2023-03-08 11:00:00	2023-03-08 12:00:00	2023-03-07 08:00:00
20230310_1900_MC1	2023-03-10 19:00:00	2023-03-10 21:00:00	2023-03-09 08:00:00
20230311_1200_MC1	2023-03-11 12:00:00	2023-03-11 13:00:00	2023-03-10 08:00:00

BESS to study how adding a flexible asset affects feasibility and cost. The exact sizes explored are determined by previous stages results, but the core BESS characteristics are fixed as in Table 5.7. The SoC bounds reflect typical operational practice to protect battery life; equal initial and final SoC avoids artificial gains and losses; efficiency figures are representative of lithium-ion systems; and a degradation charge is applied per unit of energy cycled to discourage unrealistically aggressive cycling.

TABLE 5.7: BESS parameters used in the Stage B design sweep (applied identically across sizes).

Item	Symbol	Value / Notes
Battery SoC bounds	$\mathrm{SoC}^{\mathrm{bat}}_{\mathrm{min}\,/\,\mathrm{max}}$	$[10\%, 90\%] \times C_{\mathrm{bat}}$
Battery initial/final SoC	$\mathrm{SoC}^{\mathrm{bat}}_{0}$	$50\% \times C_{\mathrm{bat}}$
Battery efficiency	η_{bat}	95%
Battery degradation	c^{deg}	0.02 €/kWh

To keep results comparable across sizes, we express degradation as a *per kWh* cost derived from simple lithium-ion values:

Average CAPEX per kWh = C = 300 €/kWh,
Cycle life = N = 7000,
Depth of discharge (DoD) = 0.8.

Charging and discharging one cycle at DoD consumes approximately $2 \times$ DoD kWh of energy per installed kWh. A practical approach to estimate the degradation cost is therefore:

$$c^{\mathrm{deg}} = \frac{C}{2\,N\,\mathrm{DoD}} \quad \left[\frac{€}{\mathrm{kWh}}\right]. \tag{5.1}$$

With the scalars above, (5.1) gives $c^{\mathrm{deg}} \approx 0.02$ €/kWh.

Stage C (RH): forecasts within the optimisation window

Stage C tests the optimal Stage B design under imperfect information. EPEX prices remain day-ahead known across the full 24 h window. For the other inputs (electric demand, solar production, and biogas production) we adopt a mixed actual/forecast view aligned with the rolling horizon in Section 5.2.3: the first 3 h (commitment slice) use *actuals*, while the remaining 21 h use *exogenous forecasts*.

To reflect that 21 h-ahead forecasts are reasonably accurate but not exact, we construct forecasts by perturbing the actual series with two types of error:

(i) a multiplicative deviation by a random percentage drawn uniformly from the range $[-20\%, +20\%]$, and
(ii) a random timing shift (lead or lag) sampled on a discrete grid of 15-minute steps, ranging from a single step up to several hours.

In other words, forecasts may slightly over (or under) estimate magnitudes and may predict features a little earlier or later than they occur. This simple scheme preserves realistic structure (daily patterns and correlations) but introduce a credible forecast error in both amplitude and timing.

Figure 5.3 shows a one day zoom comparing actual measurements with an example of the constructed forecasts. The solid blue line is the actual load; the dashed grey line is a forecast.

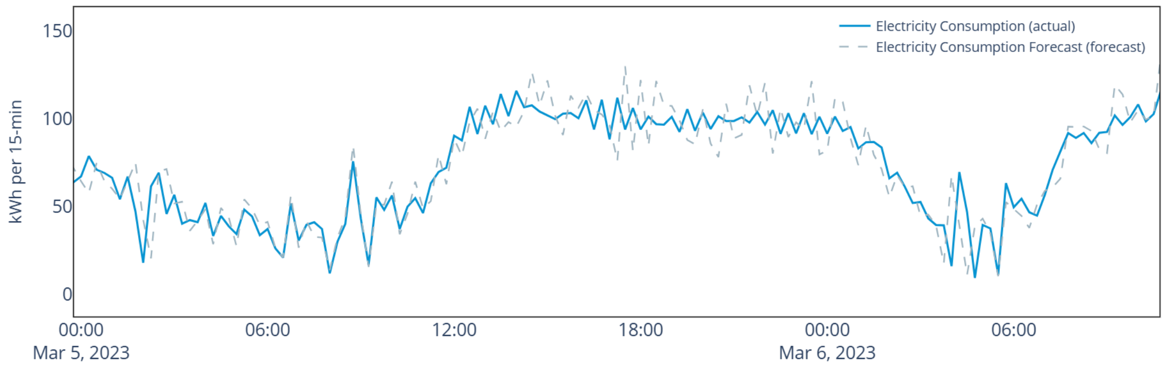


FIGURE 5.3: Electric Demand - Real vs. Forecasted

Therefore, including the inputs described before and the additional stage-dependent elements, the final table that summarises all data inputs and parameters is presented in Table 5.8.

5.3 Experiment design

This chapter details how each modelling stage is executed, what each experiment is designed to reveal, and which outputs relevant. The design follows the roadmap in Section 5.1: (i) a full-year *baseline and capability* characterisation under Perfect Foresight (PF), (ii) a *design* sweep on a stress week that introduces Monte Carlo call schedules under PF, first with the current layout and then with a battery energy storage system, and (iii) a *fidelity and robustness* assessment using a Rolling-Horizon formulation, first without and then with forecasts, complemented by exogenous stress tests. Throughout, results are benchmarked against the PF baseline.

5.3.1 Stage A: Baseline and capability (PF, full year)

This stage establishes a cost baseline, maps the plant's intrinsic flexibility *before* introducing calls, and derives thresholds that guide the call-based design in Stage B. All experiments use the cost-minimising Perfect Foresight (PF) model over the whole of 2023 at 15-minute resolution, with the technical limits and tariff structure described in Section 5.2.

TABLE 5.8: Inputs and parameters used across models (common and stage–specific).

Item	Symbol	Value / Notes
Time base & exogenous series		
WWTP electric demand (actuals / forecast)	D_t, $\hat{D}_t$	kWh/step time series
Solar production (actuals / forecast)	PV_t, $\hat{PV}_t$	kWh/step time series
Biogas production (actuals / forecast)	BG_t, $\hat{BG}_t$	kWh/step time series
EPEX day-ahead price	P_t^{da}	€/step time series (converted)
Grid limits & tariffs		
Contracted import capacity	κ_{imp}	1290 kW
Contracted export capacity	κ_{exp}	1236 kW
CBC caps explored (Stage B)	$\bar{\kappa}$	TBD [kW]
Monte Carlo call schedules (Stage B & C)	$\mathcal{C}^{(s)}$	100 schedules (MC1–MC100)
Energy tax on imports	c^{tax}	0.03935 €/kWh
Volumetric transport tariff	c^{vol}	0.0160 €/kWh
Fixed transmission fee	c^{fix}	€ 36.75 per month
Contracted–capacity fee	c^{cap}	€ 1.82 per kW·month (on κ_{imp})
Monthly peak charge	c^{peak}	€ 2.59 per kW·month (on month max)
On–site assets		
Gas Plant rated power	$P_{\text{GAS}}^{\text{max}}$	450 kW
Gas Plant ramp limits	r_u, r_d	±0.5 %/min (converted to kWh/15min)
Biogas → electricity conversion	ϕ	2.10032 kWh/m^3
Tank capacity (energy basis)	C_{tank}	400 m^3
Tank SoC bounds (min / max)	$\text{SoC}_{\text{min / max}}^{\text{tank}}$	[40%, 75%] × C_{tank}
Tank initial SoC	$\text{SoC}_0^{\text{tank}}$	50% × C_{tank}
Tank (dis)charge rate caps	t_u, t_d	50 kWh/step
Battery (Stage B)		
Explored battery sizes	C_{bat}	TBD [kWh]
Battery power limit	$P_{\text{bat}}^{\text{max}}$	TBD [kW]
Battery SoC bounds	$\text{SoC}_{\text{min / max}}^{\text{bat}}$	[10%, 90%] × C_{bat}
Battery initial/final SoC	$\text{SoC}_0^{\text{bat}}$	50% × C_{bat}
Battery efficiency	η_{bat}	0.95
Battery degradation cost	c_{deg}	0.02 €/kWh
Virtual sinks & penalties		
Unmet demand penalty	M_{unmet}	Very large
Biogas sink penalty	M_{bio}	Very large
Solar curtailment penalty	M_{pv}	0

A1. PF benchmark (no calls)

The **purpose** of this model is to provide the reference operation and cost against which all subsequent designs are compared. This is *not* the BAU case. Here, we economically optimise the responses to market signals and grid tariffs, providing a best-case from a cost perspective, subject only to the technical constraints.

The **set-up** is very simple, we solve the PF model for the full year with no CBC constraints or calls. The optimiser schedules PV, CHP, tank usage, and grid exchanges against EPEX prices plus taxes and network tariffs, enforcing all asset limits, summarised in Table 5.8.

The **primary outputs** are therefore: (i) Total electricity cost and its components (imports, exports, monthly-peak charges, and volumetric tariff), and the resulting grid import profile. These allow comparison with BAU to verify that the economic optimisation indeed improves cost and to observe how those improvements affect grid imports. (ii) The same metrics restricted to the *stress week* used in **Stages B** and **C**, so that results are directly comparable across stages.

A2. PF yearly cap (no calls)

The **purpose** of this experiment is to identify the lowest feasible *annual* contracted-capacity level and clarifies when a fixed cap would bind. As Figure 5.2 shows, a large portion of the contracted import capacity is unused over the year. If caps were tested arbitrarily, many values would be accepted simply because they never challenge plant operation. This is not inherently a problem, but they deliver no reduction at the transformer and therefore do not physically address congestion. Because our aim is to test caps that *do* make a difference at the transformer, this experiment isolates the range of cap levels that are likely to bind over the year.

The **set-up** is as follows: we impose a single import cap that applies uniformly throughout the whole year and progressively tighten it, starting from the current contracted level of **1290 kW** until any further reduction would force withdrawals from the unmet-demand slack (i.e., render the problem physically infeasible). The feasibility is **monotone non-decreasing** in the cap, therefore, if the model is infeasible at some level k_0, it remains infeasible for any tighter cap $k \leq k_0$. We exploit this with a bisection search. We first bracket the threshold with larger steps and then tighten the interval iteratively until finding the lowest feasible cap value with 1kW tolerance. This procedure is simple, computationally efficient, and reliably finds the feasibility threshold.

The **primary outputs** are: (i) the threshold cap (kW) that remains feasible over the full year; and (ii) the incremental cost relative to the A1 PF benchmark. Together these results indicate which cap levels are likely to bind in practice and they inform the cap range explored in Stage B.

Finally, it is important to mention that this test aligns with the **GOTORK** measure discussed in Chapter 3, under which DSOs may remove persistently unused contracted capacity. We comment on that policy implication alongside the results.

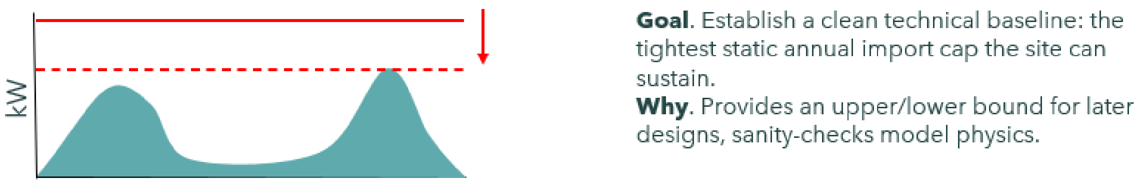


FIGURE 5.4: A2. Yearly fixed cap experiment (schematic).

A3. PF grid-imports discouraged (no calls)

The **purpose** of this experiment is to quantify the plant's intrinsic flexibility *during congestion-prone hours* before any call mechanism is introduced. In plain terms, it asks how the import profile would change if grid intake were strongly discouraged at a fixed set of hours each day. The comparison with **A1. PF benchmark** provides an early view of how much the grid imports can be reshaped with the current layout alone and offers a useful check for the stages that follow: if a given reduction appears feasible here, later call-based designs should be similar.

The **set-up** is straightforward. A set of *typical* congestion hours, previously defined in Table 5.5, and a prohibitive per kWh penalty is applied to imports within those windows. This implements a *soft* cap, the violations remain possible at high cost to preserve feasibility, in contrast to the hard caps used in real CBCs. All other inputs, prices, and technical limits are identical to **A1**.

The **outputs** are primarily diagnostic. The run reveals an *upper bound* on achievable import reductions in the selected hours with the existing assets; it shows how the import profile is reshaped and the extent to which on-site resources (PV, CHP, and the biogas tank) are relied upon to substitute for the grid. These results provide context and validation for Stage B's call-based experiments.

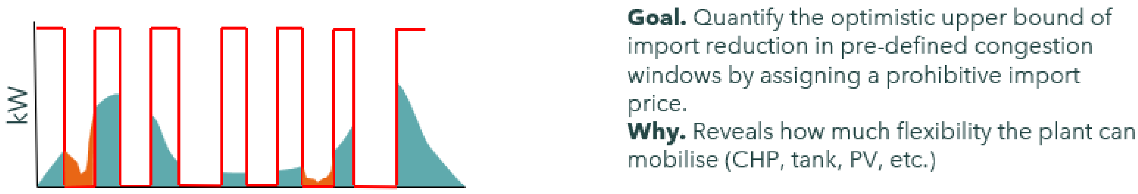


FIGURE 5.5: A3. Grid imports discouraged (schematic).

5.3.2 Stage B: Design sweep on a stress week (PF with MC calls)

This stage screens call-based CBC designs on the *stress week* to identify combinations of cap level and on-site flexibility that (i) accept all calls safely and (ii) deliver reductions at an efficient price. We first test the *current layout* and then add a BESS to study how this flexible asset affects feasibility and costs. All runs use the PF formulation, so calls are perfectly known and assets can be optimally pre-positioned around them, which provides a best-case performance.

Calls are generated by Monte Carlo sampling from empirical distributions, Section 5.2.5 details the construction. We create 100 distinct schedules for the stress week and use the **same** schedules for both experiments in this stage to allow for comparison. The technical limits, tariffs, and cost components are as defined in Section 5.2.

B1. PF sweep with MC calls (current layout)

The **purpose** here is to test candidate CBC caps informed by Stage A, experiment A2, and to determine whether any *safe* fixed cap can be implemented with the existing assets, as well as to identify (if it exists) an optimal CBC limit under the current layout.

The **set-up** is the following: for each cap level and for each Monte Carlo schedule, we solve the PF model and assess feasibility call by call: a call is feasible only if the **unmet demand** slack remains zero throughout the called window. We also record the **cost** of each run (in other words, the energy bill, which also includes grid tariffs).

B2. PF sweep with MC calls (with BESS)

The **purpose** of this experiment is the same as B1 but this time with an added flexibility. A BESS is introduced because it is expected to, for example, shift PV output or inexpensive imports to sustain grid import caps during the called windows. The sizes tested should reflect realistic plant-scale options; that is why the exact values are determined once B1 results are available. Ultimately, we seek the combination of BESS size and CBC limit that yields the maximum reduction at the lowest price.

The **set-up** is identical to B1, with an additional layer in the sweep: for each cap level and each Monte Carlo schedule, we repeat the run for each candidate BESS size, check feasibility, and report costs.

The main **outputs** are the same for both experiments. Among configurations that achieve *100% acceptance*, we compute a *normalised price*: the incremental cost relative to the Stage A PF benchmark per unit of reduction actually delivered during the called windows. This allows for a fair cost-effectiveness comparisons across caps and layouts. We then trace a Pareto frontier of reduction versus normalised price and select, for subsequent validation, the knee point that balances depth of reduction with cost. For designs including a BESS, we also include a post-process capital charge to reflect the investment. In summary, results are reported at three levels: per call (feasibility), per schedule (acceptance rate and weekly cost decomposition), and in aggregate across the 100 schedules (distribution of normalised prices and reductions, together with the Pareto frontier). These outputs identify the most promising cap/BESS combinations to carry forward to Stage C.

5.3.3 Stage C: Rolling-horizon validation and robustness (RH on the stress week)

This stage tests whether the optimal selected design remains feasible and economically credible once decisions are made with finite look-ahead and rolling commitment, and when forecasts replace perfect information. It separates performance losses due to the rolling horizon from those due to forecast error. Where acceptance falls, we consider simple operational policies to recover feasibility; finally, we probe robustness to exogenous inputs outside the WWTP's control.

All experiments use the same *stress week* (05–03–2023 00:00 to 11–03–2023 23:45) and the same 100 Monte Carlo call schedules as in Stage B, enabling one-to-one comparisons. The rolling-horizon (RH) configuration follows Section 5.2.3: a 24 h optimisation window with a 3 h receding commit. Tariffs, technical limits, and cost components are as in Section 5.2.

C1. RH without forecasts (RH–PF horizon)

Its **purpose** is to provide a clean control that isolates the effects of rolling commitment and finite look-ahead from those of forecast error.

For the **set-up**, within each 24 h window the model has perfect knowledge of all inputs across the window; at each receding step the first 3 h are committed, and the horizon advances.

The key **outputs** mirror Stage B, acceptance per call and total cost relative to PF benchmark.

C2. RH with forecasts (RH+Fcst)

This experiment **purpose** is to quantify the impact of imperfect information and more closely reflect operational reality. EPEX prices remain day-ahead known across the window. For demand, PV, and biogas, the first 3 h (the committed slice) use actuals, while the remaining 21 h use exogenous forecasts constructed as described in Section 5.2.5. Calls are announced the day before at 08:00 (ahead of the day-ahead market gate closure) and become visible to the optimiser upon announcement.

The **outputs** are as exactly as above.

C3. Robustness: external factors

The main **purpose** is to assess sensitivity to external shocks outside the WWTP's control and to observe how call feasibility and costs respond under adverse but plausible conditions.

For the **set-up** we use the RH+Fcst configuration, we run three scenarios on the *stress week*. (i) *Extreme price events:* the top 10% of observed prices (around 180 €/MWh) are lifted to 500 €/MWh to emulate potential future years with very high marginal generation costs [72]. (ii) *Higher demand:* a uniform +20% increase is applied to plant demand across the week, which exceeds the medium-term growth discussed in Section 2 and therefore provides a conservative test. (iii) *Higher grid tariffs:* volumetric, contracted-capacity, and monthly-peak charges are each tripled, reflecting anticipated network investment pressures in the coming years [11][7]. For interpretation, costs are compared in *incremental* terms against the corresponding no-CBC baseline under the same shocked inputs.

The **primary outputs** are again, for each scenario, call feasibility and the incremental cost relative to the PF bechhmark. We also assess whether we can conlude if the per-call CBC compensation derived in Stage B remains defensible under these shocks.

5.4 Decision criteria

This chapter defines how each stage is judged against the study aims introduced in Section 5.2. The guiding principles are: (i) feasibility first (100% of calls must be honoured, in order to be considered feasible), (ii) comparisons are made against the PF benchmark baseline under identical inputs, and (iii) where multiple designs are feasible, preference is given to those that deliver reduction at the lowest normalised cost.

5.4.1 Stage A (PF, full year): benchmarking and thresholds

Stage A is primarily diagnostic. The *PF benchmark* establishes the reference **cost** and **import profile** under 2023 prices and tariffs. The *PF yearly cap* experiment yields a feasibility threshold for a fixed, year-round import cap; this single number is used to delimit the range of call-based caps explored later. The *PF grid-imports discouraged* run provides an upper bound on how far the existing layout can reshape imports during congestion-prone hours, and thus serves as a reasonableness check for the call-based designs.

Decision-wise, Stage A therefore does not select a contract; rather, it produces inputs for Stage B: the feasible annual cap threshold, and an envelope for the achievable import reductions against which subsequent results can be validated.

5.4.2 Stage B (PF, stress week): feasibility and cost-effectiveness

In Stage B the design space is tested under Perfect Foresight (PF) on the *stress week*, using the 100 Monte Carlo (MC) call schedules defined in Section 5.2.5. Two layouts are tested (current layout and with BESS), and a single cap level c is applied uniformly to all calls within a run.

Feasibility. For a given pair (c, b), cap c and battery size b (with $b = 0$ for the current layout), feasibility is assessed per schedule by requiring that the unmet-demand slack remains zero at every time step inside each called window. A schedule is *accepted* if all its calls satisfy this condition. The feasibility ratio for (c, b) is the fraction of the 100 schedules that are accepted.

For design (c, b) and each MC schedule s, let $U_s(c, b)$ be the *total* unmet–demand slack accumulated *only within called windows* (kWh). The schedule is *accepted* if $U_s(c, b) = 0$. If $\mathcal{S}$ is the set of 100 schedules, the feasibility ratio is

$$\mathrm{FR}(c, b) \;=\; \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} \mathbf{1}[\,U_s(c, b) = 0,]\,, \tag{5.2}$$

i.e. the fraction of schedules that accept all calls without using slack.

Incremental cost and baseline. Costs are computed against the PF benchmark for the *same stress week and inputs*. The incremental cost of operating under (c, b) is

$$\Delta C \;=\; C_{(c,b)} \;-\; C_{\text{baseline}}, \tag{5.3}$$

where $C_{(c,b)}$ includes market charges and tariffs as in Section 5.2. For designs that include a BESS, we also add a capital charge in post-processing to reflect investment.

We annualise the upfront specific energy cost c_E (EUR/kWh) with the capital recovery factor (CRF) at real WACC i and life n years:

$$\mathrm{CRF}(i, n) \;=\; \frac{i(1+i)^n}{(1+i)^n - 1}, \qquad c_E^{\text{ann}} \;=\; c_E \cdot \mathrm{CRF}(i, n), \tag{5.4}$$

and allocate it to a week as $c_E^{\text{wk}} = c_E^{\text{ann}}/52$. The weekly capital charge for a battery of usable energy E_B (kWh) is then

$$C_{\text{week}}^{\text{BESS}} \;=\; c_E^{\text{wk}}\, E_B. \tag{5.5}$$

With $c_E = 300$ EUR/kWh (for a system with C-rate = 0.5), $i = 7\%$, $n = 12$ years we obtain $c_E^{\text{ann}} \approx 44.07$ EUR/kWh·yr and

$$c_E^{\text{wk}} \;\approx\; \frac{44.07}{52} \;\approx\; 0.847 \text{ EUR/kWh·week}.$$

Normalised price (levelised cost of callable reduction). To properly compare designs, we report a normalised price that relates incremental cost to the reduction actually *made available* by the contract. If the contracted import limit is K (kW), the CBC cap is c (kW), and the calls $i \in \mathcal{D}$ have durations h_i (h). The callable reduction over a schedule is

$$T_{\text{red}} \;=\; \sum_{i \in \mathcal{D}} (K - c)\, h_i \qquad [\text{MWh}], \tag{5.6}$$

and the energy-normalised metric is

$$\mathrm{LCCR}_E = \frac{\Delta C}{T_{\mathrm{red}}} \qquad [\mathrm{EUR/MWh}]. \tag{5.7}$$

Selection rule and Pareto frontier. Among designs with a feasibility ratio of 100%, we construct a Pareto frontier of $\{\text{reduction}, \text{LCCR}\}$ and select the knee point as the preferred trade-off. This selection, together with the full feasibility and cost distributions, is carried forward to Stage C for validation under rolling horizon, forecasts and exogenous events.

5.4.3 Stage C (RH, stress week): realism and robustness

Stage C takes the preferred $(\text{cap}, \text{BESS})$ design selected in Stage B and validates it under rolling, finite-horizon decision-making on the same stress week. All experiments use the RH set-up in Section 5.2.3 and the same 100 Monte Carlo call schedules. The ultimate aim is to quantify how *feasibility* (acceptance of all calls) and *incremental cost* (relative to the PF benchmark) change when moving from perfect foresight to realistic information. In short, across the Stage C experiments we track the variation in feasibility and Δ cost to separate horizon effects from forecast error, and to assess robustness under operational policies and external shocks.

This final set of tests indicates whether the selected design is genuinely robust and could be safely implemented. If call acceptance remains at 100% and the incremental cost relative to the no CBC benchmark stays within the expected range across all scenarios, the design can be considered robust. Conversely, any systematic loss of feasibility or material cost escalation means further tuning is required before recommending that design.

5.4.4 CBC representation in the models

2.2 Contract-consistent hard-cap formulation.
The second variant incorporates the contractual logic of CBCs more explicitly, rather than penalising energy volumes above a target, it imposes a hard cap on import capacity during congestion-prone hours. To reflect the economic reality that capacity reduction is compensated, the model also attaches a price component to the reduction (expressed in €/kW for the purposes of this experiment), allowing it to endogenously choose the depth of capacity reduction that is economically optimal in each constrained hour. This formulation exposes the direct relationship between revenue and capacity reduction and provides a practical benchmark for subsequent call-based analyses.

5.4.5 Objective function

The objective is to minimise total electricity cost over the horizon. Energy purchases are priced at the EPEX day-ahead rate plus the statutory energy tax (as specified by the Dutch tax authority [70]), while exports are credited at the EPEX price. This construction creates a deliberate buy–sell asymmetry: each kilowatt-hour imported costs more than each kilowatt-hour exported earns, because taxes and network tariffs apply to imports but not to exports. The objective also includes the volumetric transport tariff, the contracted-capacity fee, the fixed monthly fee, and the monthly peak charge, in line with the grid-tariff structure applicable to the plant (described in Chapter 2.2). Framed this way, the cost function mirrors the operator's actual bill, ensuring that simulated savings and trade-offs are directly interpretable for decision support.

This cost-minimising objective reflects operational reality, a policy that forces to minimise grid consumption at all times, regardless of external signals, would be unnecessarily restrictive and economically inefficient. As emphasised by an expert in System Operations at Alliander (see Appendix A), import reductions are more efficient if targeted to specific windows of network stress; outside those periods, the DSO is largely indifferent to consumption provided contractual limits are respected. The models are therefore designed to test precisely this balance between targeted constraint compliance and economically efficient operation.

5.5 Mathematical Formulation

This section sets out the mathematical formulations used to analyse CBC implementation. It begins with the **core formulation**, which applies to all experiments and scenarios and corresponds tu experiment *A1. PF Benchmark*. All case specific features, such as capacity caps, calling schedules, or added assets, are introduced later as additive constraints or variables layered on this base formulation.

5.5.1 Routed flows and attribution: modelling choice

Before introducing the mathematical formulation, we first set out several points that are important for the approach.

This chapter adopts an explicit routing representation in which *behind the meter* sources are split into separate flow variables, for example:

`solar`→`plant`, `solar`→`grid`, `gas plant`→`WWTP` or `gas plant`→`grid`

The choice is deliberate for three reasons. First, it mirrors the physical picture set out in Chapter 2, where energy from each asset can follow distinct paths to serve the WWTP, to charge a storage device, or to be exported. Writing the model in this way preserves that picture in the mathematics. Second, the formulation aligns with *Repowered*'s internal practice, where one variable is assigned per flow. Using the same pattern makes easier coupling with existing tools and simplifies later model handover. Third, using routed flows lets us draw stacked *supply to plant* plots. In those plots we can see, at each time step, how solar, CHP, battery and grid add up to meet demand, and any missing part shows up as a clear gap. If we only plotted imports and exports at the point of common coupling, that shortfall would be much harder to visualise.

However, there is an important caution. As mentioned earlier, in this study, serving the plant has zero marginal value for both solar and CHP, and exporting either source earns the same price. The objective and the constraints therefore depend on the *sums* of routed flows rather than on their split. As a result, several equally good internal allocations can exist at a given time step. For example, when both solar and CHP are producing and some energy is exported, a small amount of solar assigned to the plant can be exchanged with CHP assigned to the plant, but total export and all costs remain unchanged at all timestamps. This non uniqueness affects only the **internal** series, not the contract or the economics. The quantities that drive conclusions, grid imports and exports at the PCC and any unmet demand at the plant, are uniquely determined by the cap, the asset limits, and the very high penalty on unmet energy. They are therefore invariant across runs that use the same data and constraints.

In short, feasibility and settlement depend on totals at the PCC, so we treat the routed series as an *attribution layer* for visualisation rather than as measured physical facts. We check that this layer behaves correctly using the validation tests in Section 5.8. If a unique split is needed in future work, simple fixes are available, for example, an explicit merit order that encodes priorities. In this thesis, that extra structure is not required, because the aim is to assess compliance with the cap at the PCC and the associated operational and economic effects. The model is therefore kept clear for plots and compatible with *Repowered*'s tools.

5.5.2 Core formulation

The Pyomo implementation defines: (i) time-indexed decision variables for flows (to the plant, to export, and to sinks), (ii) a state variable for the biogas tank, and (iii) monthly peak variables used to compute peak charges. Power limits given in kW are converted to per-step energy bounds (kWh per 15-min) via a 0.25 multiplier; prices are applied at the same granularity. The objective minimises total cost (imports at EPEX plus taxes and tariffs, minus export revenue, plus fixed/peak/capacity terms and big-M penalties for non-physical sinks), subject to all constraints.

5.5.2.1 Index sets and time discretisation

We discretise time at 15-minute resolution. Let

$$\mathcal{T} \;=\; \{0, 1, \ldots, N-1\}$$

be the set of time-step indices, where each $t \in \mathcal{T}$ corresponds to a calendar timestamp τ_t. We write

$$\tau_t \;=\; \tau_0 + t\,\Delta t, \qquad \Delta t = 15 \text{ minutes.}$$

In a full non-leap year (2023) with quarter-hour steps, $N = 365 \times 96 = 35{,}040$, so $\mathcal{T}$ enumerates all timestamps from $\tau_0 = $ 01-01 00:00 to $\tau_{N-1} = $ 12-31 23:45.

Monthly peak charges are modelled using a set of calendar months

$$\mathcal{M} \;=\; \big\{\text{month}(\tau_t) \;:\; t \in \mathcal{T}\big\},$$

i.e., the unique months covered by the horizon. We also use the mapping $\mu : \mathcal{T} \to \mathcal{M}$, with $\mu(t) = \text{month}(\tau_t)$, to relate each time step to its billing month in the peak-tracking constraints.

5.5.2.2 Decision variables

All flow variables are non-negative energies per 15-minute step [kWh] and state variables are [kWh] unless stated otherwise.

All decision variables present in the core formulation are summarised in Table 5.9.

5.5.2.3 Objective Function

As mentioned before, the optimisation aims to minimise the total electricity cost over the horizon:

$$\min \Big(\underbrace{C^{\text{imp}} - R^{\text{exp}}}_{\text{net energy-market}} \;+\; \underbrace{C^{\text{vol}} + C^{\text{peak}}}_{\text{grid usage charges}} \;+\; \underbrace{C^{\text{pen}}}_{\text{feasibility penalties}} \Big) \tag{5.8}$$

TABLE 5.9: Core decision variables in the optimisation model.

Pyomo name	Index set	Bounds	Interpretation (units)
grid_import [t]	$t \in \mathcal{T}$	[0, 322.5]	Energy imported from grid to WWTP per step (kWh).
grid_export [t]	$t \in \mathcal{T}$	[0, 309]	Energy exported to grid per step (kWh).
unmet_WWTP [t]	$t \in \mathcal{T}$	≥ 0	Unserved WWTP demand (big-M penalised) (kWh).
solar_WWTP [t]	$t \in \mathcal{T}$	≥ 0	PV energy to WWTP (kWh).
solar_grif [t]	$t \in \mathcal{T}$	≥ 0	PV energy to export (kWh).
solar_sink [t]	$t \in \mathcal{T}$	≥ 0	PV curtailment (spill) (kWh).
bio_tank [t]	$t \in \mathcal{T}$	≥ 0	Biogas sent to tank (expressed as kWh_e equiv.) (kWh).
bio_CHP [t]	$t \in \mathcal{T}$	≥ 0	Biogas sent directly to CHP (kWh_e equiv.) (kWh).
bio_sink [t]	$t \in \mathcal{T}$	≥ 0	Biogas to penalised sink (flare) (kWh).
tank_CHP [t]	$t \in \mathcal{T}$	≥ 0	Biogas withdrawn from tank to CHP (kWh_e equiv.) (kWh).
CHP_WWTP [t]	$t \in \mathcal{T}$	≥ 0	CHP electricity to WWTP (kWh).
CHP_grid [t]	$t \in \mathcal{T}$	≥ 0	CHP electricity to export (kWh).
SoC_tank [τ]	$t \in \mathcal{T}$	[160, 300]	Biogas tank state of charge (m3)
monthly_peak [m]	$m \in \mathcal{M}$	≥ 0	Monthly maximum of grid_import; billed per kW

Notes: Time step is 15 minutes. Power limits in kW are converted to per-step energy bounds by multiplying by 0.25.

The first bracket is the net market payment, imports valued at the EPEX day-ahead price plus energy tax minus revenues from exports. The second bracket contains grid charges that depend on actual usage: the volumetric transport fee on each imported kWh and the monthly kW–max charge computed from the month's highest 15-minute import. Finally, C^{pen} aggregates big-M penalties on non-physical sinks and unmet demand to enforce feasibility only as a last resort.

5.5.2.4 Energy Balance Equations

At each time step, PV production is allocated to on-site use, export, or curtailment:

$$\texttt{solar\_WWTP}[t] + \texttt{solar\_grid}[t] + \texttt{solar\_sink}[t] = PV_t(input), \qquad \forall\, t \in \mathcal{T}. \tag{5.9}$$

Produced biogas is sent to storage, used directly by the CHP, or diverted to the penalised sink:

$$\texttt{bio\_tank}[t] + \texttt{bio\_CHP}[t] + \texttt{bio\_sink}[t] = BG_t(input), \qquad \forall\, t \in \mathcal{T}. \tag{5.10}$$

The biogas tank level evolves with inflow to storage and withdrawals to the CHP.

$$\texttt{SoC\_tank}[t] = \texttt{SoC\_tank}[t-1] + \texttt{bio\_tank}[t] - \texttt{tank\_CHP}[t], \qquad \forall\, t \in \mathcal{T}. \tag{5.11}$$

And it starts from a given state-of-charge and a cyclic terminal condition restores the initial level at the end of the horizon:

$$\texttt{SoC\_tank}[0] = \text{SoC}_{\text{init}}. \tag{5.12}$$

$$\texttt{SoC\_tank}[N-1] = \text{SoC}_{\text{init}}. \tag{5.13}$$

The total CHP electricity equals the biogas energy supplied to the CHP (direct plus from tank) converted to electricity:

$$\texttt{CHP\_WWTP}[t] + \texttt{CHP\_grid}[t] = \big(\texttt{bio\_CHP}[t] + \texttt{tank\_CHP}[t]\big) \cdot \phi, \qquad \forall\, t \in \mathcal{T}. \tag{5.14}$$

subject to the maximum power output:

$$\texttt{CHP\_WWTP}[t] + \texttt{CHP\_grid}[t] \leq \text{chp}^{\max}, \qquad \forall\, t \in \mathcal{T}. \tag{5.15}$$

Step changes in total CHP output are limited by ramp-up and ramp-down bounds:

$$\big(\texttt{CHP\_WWTP}[t] + \texttt{CHP\_grid}[t]\big) - \big(\texttt{CHP\_WWTP}[t-1] + \texttt{CHP\_grid}[t-1]\big) \leq r_u, \qquad \forall\, t \in \mathcal{T} \setminus \{0\}, \tag{5.16}$$

$$\big(\texttt{CHP\_WWTP}[t-1] + \texttt{CHP\_grid}[t-1]\big) - \big(\texttt{CHP\_WWTP}[t] + \texttt{CHP\_grid}[t]\big) \leq r_d, \qquad \forall\, t \in \mathcal{T} \setminus \{0\}. \tag{5.17}$$

The flows that can export electricity to the grid are the PV and CHP:

$$\texttt{grid\_export}[t] = \texttt{solar\_grif}[t] + \texttt{CHP\_grid}[t], \qquad \forall\, t \in \mathcal{T}. \tag{5.18}$$

Finally, the WWTP demand is met each step by PV, CHP and grid imports; additionally, the unmed demand slack is added:

$$\texttt{solar\_WWTP}[t] + \texttt{CHP\_WWTP}[t] + \texttt{grid\_import}[t] + \texttt{unmet\_WWTP}[t] = D_t, \qquad \forall\, t \in \mathcal{T}. \tag{5.19}$$

5.5.2.5 Grid-tariff terms

The two components of the grid tariff that are optimised in the model are the volumetric tariff and the monthly max grid import.

The volumetric transport component used in the core model is:

$$C^{\text{vol}} \;=\; c^{\text{vol}} \sum_{t \in \mathcal{T}} \texttt{grid\_import}[t]. \tag{5.20}$$

which charges a linear €/kWh fee on all imported energy over the horizon.

The monthly kWmax is the maximum, instantaneous power over all quarter-hour blocks in a month, of `grid_import`. On the first day of the next month, the counter resets. To reproduce this mechanism we first define $\mathcal{M}$ as the set of calendar months in the horizon and $\mu : \mathcal{T} \to \mathcal{M}$ maps each time step to its month:

$$\mathcal{M} = \{\text{month}(\tau_t) : t \in \mathcal{T}\}, \qquad \mu(t) = \text{month}(\tau_t). \tag{5.21}$$

The definition and charge read:

$$\texttt{monthly\_peak}[m] \geq \texttt{grid\_import}[t], \qquad \forall m \in \mathcal{M}, \forall t \in \mathcal{T} \text{ with } \mu(t) = m, \tag{5.22}$$

$$C^{\text{peak}} = \sum_{m \in \mathcal{M}} c^{\text{peak}} \left(4 \cdot \texttt{monthly\_peak}[m]\right). \tag{5.23}$$

Equation (5.22) ensures that the auxiliary variable `monthly_peak`$[m]$ bounds all per-step imports. Equation (5.23) converts that maximum to kW by the factor 4 and applies the €/kW·month tariff c^{peak}. This construction is exact with respect to how the charge is calculated on the bill and gives the optimiser a clear incentive to reduce short, sharp import spikes.

The peak charge C^{peak} and the volumetric term C^{vol} are billed in parallel and are distinct from the fixed monthly fee C^{fix} and the contracted capacity fee C^{cap} that are added to the bill outside the optimisation.

5.5.2.6 Other costs and penalties

The imports are valued at the day-ahead market price plus the statutory energy tax. P_t is the EPEX price (€/kWh) at step t, and c^{tax} is the per-kWh energy tax applied to imports. This term reflects the cash outflow for every imported kWh.

$$C^{\text{imp}} = \sum_{t \in \mathcal{T}} \left(P_t + c^{\text{tax}}\right) \texttt{grid\_import}[t]. \tag{5.24}$$

Exports are remunerated at the day-ahead price. This term is a cash inflow.

$$R^{\text{exp}} = \sum_{t \in \mathcal{T}} P_t \cdot \texttt{grid\_export}[t]. \tag{5.25}$$

The large penalties prevent the optimiser from using the sinks except in to prevent infeasibilities. `bio_sink` is the penalised biogas disposal, and `unmet_WWTP` represents unmet plant demand. Big-M coefficients ($M_{\text{pv}}, M_{\text{bio}}, M_{\text{unmet}}$) are set high so these terms appear only to preserve feasibility. `solar_sink` is not included in the equations since the penalty of solar curtailment is zero.

$$C^{\text{pen}} = M_{\text{bio}} \sum_{t \in \mathcal{T}} \texttt{bio\_sink}[t] + M_{\text{unmet}} \sum_{t \in \mathcal{T}} \texttt{unmet\_WWTP}[t]. \tag{5.26}$$

5.5.3 Specific Formulation for Stage A

This section introduces the formulation used to study *fixed* capacity-limiting CBCs under *perfect foresight* (PF). The PF model is identical to the core formulation developed above in its representation of assets, balances, intertemporal states, and tariff structure. The model-specific additions are introduced and explained in the subsections that follow.

5.5.3.1 A2. PF Yearly Cap

This model has a key addition with respect to the core model: the grid off-take (import) limit is imposed as a *single, year-round cap* and that cap is chosen based on feasibility.

Concretely, we replace the default bound on grid imports $\kappa_{\text{imp}} = 1290\text{kW}$ with the lowest value that still preserves physical feasibility of the system over the entire horizon. Feasibility is defined with respect to the unmet-demand slack variable. A solution is deemed infeasible for a new candidate cap κ if the optimisation requires any positive unmet demand at any step,

$$\exists\, t \in \mathcal{T} : \quad \texttt{unmet\_WWTP}[t] \; > \; 0, \tag{5.27}$$

i.e., if the plant would need to draw from the unmet-demand node to satisfy the balances. Because tightening κ can only make the problem *harder* (feasibility is monotone non-increasing in κ), we determine the threshold cap via a *Successive Approximation* search. Starting from the nominal contracted level (κ_{imp}), the algorithm lowers the grid import limit in coarse steps until infeasibility is first encountered, thereby bracketing the threshold. It then refines the step size within the feasible–infeasible bracket to approach the smallest feasible cap with a tolerance of 1kW. This procedure yields:

$$\kappa^{\star}_{\text{PF}} = \min\Big\{\kappa \geq 0 \;:\; \exists \text{ feasible solution with } \texttt{unmet\_WWTP}[t] = 0 \;\; \forall\, t \in \mathcal{T}\Big\}. \tag{5.28}$$

which we interpret as the *theoretical minimum uniform contracted import* the site can sustain across the year with perfect anticipation, without any unmet demand.

5.5.3.2 A3. Grid import discouraged

Similar to the previous model, this benchmark also keeps the full core formulation unchanged, but alters the import price during the pre-specified congestion windows, see Table 5.5. Instead of imposing a hard cap on grid off-take, imports are merely discouraged in those windows by assigning them an extremely high penalty. The model therefore may still import if needed, yet it will exhaust all on-site flexibility to avoid doing it when the penalty applies. This provides an optimistic upper bound on the plant's technical ability to reduce imports during congestion prone hours, without yet committing to a specific capacity limit.

For the congestion periods, let $z_t \in \{0,1\}$ be an indicator built from the congestion calendar. We set $z_t = 1$ if time step t falls within a congested window, and $z_t = 0$ otherwise. Formally, for each month m define a set of weekday hour ranges $\mathcal{H}_m \subseteq \{0,\ldots,23\} \times \{0,\ldots,23\}$ (start and end); then

$$z_t \;=\; \begin{cases} 1, & \text{if weekday}(\tau_t) \text{ and } \exists\,(h_0,h_1) \in \mathcal{H}_{\mu(t)} \,:\; h_0 \leq \text{hour}(\tau_t) < h_1, \\ 0, & \text{otherwise.} \end{cases}$$

Here τ_t is the timestamp of step t and $\mu(t)$ is its calendar month.

Additionally, we introduce a large penalty price P^{CBC} (€/kWh). The *effective* import price used in the objective becomes

$$\tilde{P}^{\text{imp}}_t \;=\; \begin{cases} P^{\text{CBC}} + c^{\text{tax}}, & \text{if } z_t = 1, \\ P_t + c^{\text{tax}}, & \text{if } z_t = 0, \end{cases} \qquad \Longleftrightarrow \qquad \tilde{P}^{\text{imp}}_t \;=\; (P_t + c^{\text{tax}}) + z_t\,(P^{\text{CBC}} - P_t).$$

The import segment of the objective is therefore

$$C^{\text{imp}}_{\text{CBC}} = \sum_{t \in \mathcal{T}} \tilde{P}^{\text{imp}}_t \, \texttt{grid\_import}[t].$$

All other objective terms (export revenue, volumetric tariff, monthly peak charge, fixed and contracted-capacity fees, and penalties) remain as in the core model.

5.5.4 Specific Formulation for Stage B

This section presents the additions to the core model needed to run the *design sweep* on the stress week under Perfect Foresight with Monte Carlo call schedules. The physical representation (nodes, balances, tank state, CHP limits) and tariff terms remain as in the core formulation. The only structural changes are: (i) the enforcement of a *hard grid-import cap* during called windows, and (ii) in the BESS variant, the addition of battery variables and constraints. We denote the stress-week time index by $\mathcal{T}_w \subset \mathcal{T}$.

5.5.4.1 B1. PF sweep with MC calls (current layout)

Let $\mathcal{S}$ index the MC schedules ($|\mathcal{S}| = 100$). For a given schedule $s \in \mathcal{S}$, let $\mathcal{C}^{(s)} \subseteq \mathcal{T}_w$ be the set of 15-minute steps that fall inside any called block. The call-based CBC cap level is a parameter c in kW (swept across candidate values); its per-step energy bound is

$$c^E = 0.25\, c \quad [\text{kWh/step}].$$

The contracted import bound κ_{imp} yields the default per-step bound $\kappa^E_{\text{imp}} = 0.25\, \kappa_{\text{imp}}$.

Then, for each schedule s, the grid import variable is capped by c^E during called steps and by the contractual bound otherwise:

$$\texttt{grid\_import}[t] \le c^E, \qquad \forall\, t \in \mathcal{C}^{(s)}, \tag{5.29}$$

$$\texttt{grid\_import}[t] \le \kappa^E_{\text{imp}}, \qquad \forall\, t \in \mathcal{T}_w \setminus \mathcal{C}^{(s)}. \tag{5.30}$$

As mentioned, calls are treated as mandatory. If (5.31) together with the physical constraints cannot serve the WWTP load, the model may use the unmet demand slack but any run with $\sum_{t \in \text{week}} \texttt{unmet\_WWTP}[t] > 0$ is ***flagged infeasible*** for that call schedule and cap level. This allows reporting of feasibility rates across Monte Carlo ensembles.

All other equations are unchanged from the core model. Because PF is assumed, the optimiser *knows* the full $\mathcal{C}^{(s)}$ set across the stress week and can pre-position on-site assets accordingly.

The objective stays equal to (5.8) with the stress-week $\mathcal{T}_w$ in place of $\mathcal{T}$. There are no extra cost terms beyond those already defined.

5.5.4.2 B2. PF sweep with MC calls (BESS integration)

The subsequent subsection extends this formulation by adding a BESS, introducing its variables and intertemporal constraints. All elements introduced previously, hard import caps during called steps, state carryover for intertemporal assets, and the incremental monthly-peak, remain unchanged and are not repeated here. We only introduce the additional variables, parameters, constraints, and one cost term needed to represent the battery and its integration with the existing flows.

The battery is characterised by the values presented in Table 5.8.

The new decision variables are the battery SoC $\texttt{SoC\_bat}[t]$, charge and discharge energies per step $\texttt{bat\_charge}[t], \texttt{bat\_discharge}[t] \in [0, E^{\text{bat}}_{\max}]$, energy supplied from the battery to the WWTP: $\texttt{bat\_WWTP}[t] \geq 0$, and battery exports $\texttt{bat\_grid}[t] \geq 0$. The model also uses nonnegative splits from PV, CHP, and grid into the battery: $\texttt{solar\_bat}[t]$, $\texttt{chp\_bat}[t]$, and $\texttt{grid\_bat}[t]$.

Also, the 15-minute step, charge and discharge energies are bounded by the power rating:

$$0 \leq \texttt{bat\_charge}[t] \leq E^{\text{bat}}_{\max}, \qquad 0 \leq \texttt{bat\_discharge}[t] \leq E^{\text{bat}}_{\max}, \qquad \forall\, t \in \mathcal{T}_{t_0:t_1}.$$

The SoC evolves according to the net charge/discharge:

$$\texttt{SoC\_bat}[t] = \texttt{SoC\_bat}[t{-}1] + \eta^{\text{bat}}\,\texttt{bat\_charge}[t] - \frac{1}{\eta^{\text{bat}}}\,\texttt{bat\_discharge}[t], \qquad \forall\, t \in \mathcal{T}_{t_0:t_1}.$$

Charging energy can originate from PV, CHP, or the grid; discharging energy can serve the WWTP or be exported.

$$\begin{aligned}
\texttt{bat_charge}[t] &= \texttt{solar_bat}[t] + \texttt{chp_bat}[t] + \texttt{grid_bat}[t], \\
\texttt{bat_discharge}[t] &= \texttt{bat_WWTP}[t] + \texttt{bat_grid}[t], \quad \forall t.
\end{aligned}$$

Now for the solar and CHP nodes we add the battery branches, and the grid import splits into WWTP and battery charging:

$$\begin{aligned}
\texttt{solar_WWTP}[t] + \texttt{solar_grid}[t] + \texttt{solar_bat}[t] + \texttt{solar_sink}[t] &= PV_t, \quad \forall t, \\
\texttt{CHP_WWTP}[t] + \texttt{CHP_grid}[t] + \texttt{chp_bat}[t] &= \texttt{bio_CHP_dir}[t] + \texttt{tank_CHP}[t], \quad \forall t, \\
\texttt{grid_WWTP}[t] + \texttt{grid_bat}[t] &= \texttt{grid_import}[t], \quad \forall t.
\end{aligned}$$

Similarily, the WWTP demand balance is extended to include battery discharge:

$$\texttt{solar\_WWTP}[t] + \texttt{CHP\_WWTP}[t] + \texttt{grid\_WWTP}[t] + \texttt{bat\_WWTP}[t] + \texttt{unmet\_WWTP}[t] = D_t, \quad \forall t,$$

and the export identity includes now battery exports:

$$\texttt{grid\_export}[t] = \texttt{solar\_grid}[t] + \texttt{CHP\_grid}[t] + \texttt{bat\_grid}[t], \quad \forall t.$$

All CHP capacity and ramping constraints from the previous experiment remain, with the only adjustment, the CHP output may now also feed the battery through $\texttt{chp\_bat}[t]$, i.e., the total CHP electrical output in each step is:

$$\texttt{CHP\_WWTP}[t] + \texttt{CHP\_grid}[t] + \texttt{chp\_bat}[t]$$

Hard caps during called steps are exactly as in the previous experiment, applied to all steps t that lie inside any call block.

The rest of the components of the objective function, market terms, volumetric tariff, incremental peak charge, and big-M penalties are the same as those used before. The objective is therefore:

$$\begin{aligned}\min \Big(& \sum_t \big(P_t + c^{\text{tax}}\big)\, \texttt{grid\_import}[t] \\ & - \sum_t P_t\, \texttt{grid\_export}[t] \\ & + c^{\text{vol}} \sum_t \texttt{grid\_import}[t] \\ & + C^{\text{peak}}_{\text{incr}} + C^{\text{deg}}_{\text{bat}} + C^{\text{pen}} \Big)\end{aligned}$$

No binary variables are introduced. With $\eta_{\text{bat}} < 1$ and a positive $c^{\text{deg}}_{\text{bat}}$, the optimiser has no incentive to create simultaneous charge/discharge cycles.

5.5.5 Formulation for Stage C (Rolling Horizon)

This section develops the formulation for call-based CBCs under a rolling-horizon (RH) solution scheme. In a call-based CBC, import capacity is constrained only within windows that are explicitly called by the DSO with advance notice. The RH approach advances through the stress week in overlapping look-ahead windows; at each re optimisation the model (i) enforces any CBC windows that have been called and are known at the window start, (ii) preserves intertemporal states (biogas tank state of charge, CHP ramp continuity), and (iii) carries forward the month-to-date peak for tariff calculation. Acceptance of calls is mandatory: if a call is active in the look-ahead, a hard cap on imports applies during that window. We first present the call-based formulation keeping the asset stack identical to the core model (WWTP, PV, biogas tank, CHP, grid). The subsequent subsections add to the formulation with a BESS and shows how its state and power constraints integrate with call enforcement in RH.

5.5.5.1 C1. RH validation with MC calls (no forecasts)

We first present the call-cap enforcement, the rolling monthly-peak mechanism, and the window structure, keeping the asset stack identical to the core model. A subsequent paragraph explains how the BESS integrates in RH (its state chaining and power/SoC constraints are exactly those introduced in Stage B).

Let $\mathcal{T} = \{0, \ldots, N-1\}$ be the 15-minute time grid with timestamps τ_t and $\Delta t = 0.25$ h. The RH solver operates on fixed-length windows

$$\mathcal{T}_{t_0:t_1} \;=\; \{\, t_0,\, t_0{+}1, \ldots, t_1 \,\}, \qquad t_1{-}t_0{+}1 \;=\; 96 \text{ steps (24 h)},$$

and recedes by 12 steps (3 h) after each solve. At the start of each window the intertemporal states are *chained* from the previously committed plan: the tank SoC at $t_0{-}1$ equals the carried value, and CHP ramp constraints are written across $t_0{-}1 \leftrightarrow t_0$ so that continuity is respected. (If a battery is present, its SoC is chained in the same manner; see the end of this subsection.)

A call schedule is a collection of blocks $\mathcal{B} = \{(b, s_b, e_b, c_b)\}$ with block identifier b, start/end datetimes (s_b, e_b), and a notice time c_b (announcement). In a given window $\mathcal{T}_{t_0:t_1}$ only those blocks can be enforced that *overlap* the window and are *known* at the window start:

$$\mathcal{B}_{t_0:t_1} \;=\; \big\{(b,s_b,e_b,c_b)\in\mathcal{B}:\; c_b\le\tau_{t_0},\; e_b>\tau_{t_0},\; s_b<\tau_{t_1}\big\}.$$

Define the active-step set of called periods in the window:

$$\mathcal{A}_{t_0:t_1} \;=\; \big\{\, t\in\mathcal{T}_{t_0:t_1}:\; \exists (b,s_b,e_b,c_b)\in\mathcal{B}_{t_0:t_1} \text{ with } s_b\le\tau_t<e_b \,\big\}.$$

Let $\bar\kappa$ [kW] be the tested CBC cap (chosen from the Stage B design set) and κ_{imp} the contracted import limit. Their per-step energy bounds are

$$\bar\kappa_e \;=\; \bar\kappa\,\Delta t, \qquad E^{\max} \;=\; \kappa_{\text{imp}}\,\Delta t \qquad [\text{kWh/step}].$$

During called steps the model imposes the *hard* import cap

$$\texttt{grid\_import}[t] \;\le\; \min\{\bar\kappa_e,\, E^{\max}\}, \qquad \forall\, t\in\mathcal{A}_{t_0:t_1}, \tag{5.31}$$

and outside called steps the default bound applies:

$$\texttt{grid\_import}[t] \;\le\; E^{\max}, \qquad \forall\, t\in\mathcal{T}_{t_0:t_1}\setminus\mathcal{A}_{t_0:t_1}. \tag{5.32}$$

All other balances and limits are exactly those of the core model.

Regarding the monthly peak carry-over strategy, let p_e^{sofar} [kWh/step] be the month-to-date peak *carried* into the current window (this value resets to a baseline at the start of each calendar month). The RH subproblem introduces a scalar decision $p_e^{\text{new}}\ge 0$ and enforces

$$p_e^{\text{new}} \;\ge\; p_e^{\text{sofar}}, \qquad \texttt{grid\_import}[t] \;\le\; p_e^{\text{new}}, \quad \forall\, t\in\mathcal{T}_{t_0:t_1}. \tag{5.33}$$

The incremental monthly-peak cost accrued in this window is then

$$C^{\text{peak}}_{\text{incr}} \;=\; c^{\text{peak}}\cdot 4\cdot\big(p_e^{\text{new}}-p_e^{\text{sofar}}\big), \tag{5.34}$$

which charges *only* if a new higher peak is set in the window (factor 4 converts kWh/step to kW). After solving, p_e^{new} becomes the carried value for the next window.

Within each window the **objective** is the same as in the core model, with C^{peak} replaced by $C^{\text{peak}}_{\text{incr}}$ so that peak charges are not double-counted across windows. In implementation, only the first 3 h (the committed slice) of each solve are applied and costs are added accordingly.

5.5.5.2 C2. RH with MC calls and forecasts

Here the RH window uses a mixed actual/forecast view consistent with operational notice. Let the committed slice be

$$\mathcal{K}_{t_0} \;=\; \{\, t_0,\, t_0{+}1,\dots,t_0{+}11 \,\} \quad (3\,\text{h}),$$

and define an indicator $\chi_t = 1$ if $t\in\mathcal{K}_{t_0}$ and $\chi_t = 0$ otherwise. The inputs used by the optimisation are

$$\tilde D_t = \chi_t D_t + (1-\chi_t)\hat D_t, \qquad \widetilde{PV}_t = \chi_t PV_t + (1-\chi_t)\widehat{PV}_t, \qquad \widetilde{BG}_t = \chi_t BG_t + (1-\chi_t)\widehat{BG}_t,$$

with EPEX day-ahead prices kept known over the full window. The balances in the window replace (D_t, PV_t, BG_t) by $(\tilde D_t, \widetilde{PV}_t, \widetilde{BG}_t)$; all other constraints, including call enforcement and rolling-peak carry-over, are unchanged.

BESS integration in RH (if present). If a battery is part of the tested design, the battery equations from Stage B (SoC bounds, charge/discharge limits, node splits, and degradation term) are included in each RH window. The SoC is chained across windows by fixing the pre-window state to the carried value and respecting the configured admissible range:

$$\texttt{SoC\_bat}[t_0-1] \;=\; \alpha_{\text{car}}\, C^{\text{bat}}, \qquad \alpha_{\min} C^{\text{bat}} \;\leq\; \texttt{SoC\_bat}[t] \;\leq\; \alpha_{\max} C^{\text{bat}}, \quad \forall\, t \in \{t_0-1, \ldots, t_1\},$$

where $\alpha_{\min}, \alpha_{\max}$ are the SoC bounds and α_{car} is the carried SoC fraction from the previously committed slice. Optional operational policies (e.g., a minimum SoC at call start) can be added as linear constraints activated on $\mathcal{A}_{t_0:t_1}$ when testing policy variants.

5.6 Main assumptions

The results in Chapter 6 rest on the following modelling assumptions:

System operation.

- WWTP process demand is inelastic (no process-side flexibility).
- PV and biogas production are exogenous time series; PV curtailment has zero marginal cost; a biogas sink and unmet-demand slack exist only as penalised safety valves.
- CHP respects a fixed nameplate limit and ramp rates; the biogas tank respects SoC bounds and per-step (dis)charge caps.
- Grid import/export limits equal the 2023 contracts; buy–sell cycles are disincentivised via price asymmetry (import price > export price).

Information and forecasts.

- 15-min resolution; day-ahead EPEX prices are perfectly known.
- Rolling horizon uses a 24 h window with a 3 h committed slice. The first 3 h use *actuals*; the remaining 21 h use forecasts with modest amplitude/timing error. No information beyond the 24 h window is known.

CBC calls.

- Calls are mandatory: if a call is active, the import cap must be honoured at every step; there is no option to skip or partially accept.
- Call notices arrive at 08:00 the day before (pre–DA gate closure).
- Monte Carlo call schedules are sampled from DSO informed study with the probability and duration distributions; the same 100 schedules are used across layouts.

Economics and tariffs.

- Objective minimises the WWTP's electricity bill: EPEX energy, energy tax (single band), volumetric transport fee, fixed monthly fee, contracted-capacity fee, and monthly peak charge (on the month's max 15-min import).
- PV and gas plant are modelled at zero *marginal* cost.

- CBC compensation is evaluated as a *normalised* cost (LCCR); no explicit DSO payment rule is imposed.

Battery (when present).

- Tested as a site asset with a linear degradation charge introduced into the objective, plus a weekly capital charge.
- No revenue from ancillary services or arbitrage outside the modelled tariff/price terms.

Scope limitations.

- No explicit network constraints; losses, outages, and maintenance are ignored.
- Asset parameters (limits, efficiencies, taxes, grid tariffs) are taken as constant over the study period.

5.7 Tools and reproducibility

The WWTP's energy system couples multiple assets (PV, CHP, biogas tank, grid interface) through intertemporal states and limits (tank SoC, CHP ramping, monthly peak). A mixed-integer linear programming (MILP) framework is therefore a natural choice: it can represent discrete operating modes when required while keeping a transparent link between decisions and bill items. In this study, however, all models are solved as linear programmes (LPs) without integer variables, on/off logic was not needed, this was softly enforced within the formulation, when needed. However, formulation remains *MILP-ready* if in future work other discrete controls are required.

Implementation stack

- **Language / modelling:** Python 3.x with `Pyomo` for model algebra and data piping.
- **Solver:** Gurobi 12.0.3 (LP barrier/dual-simplex). Threads set to use available cores unless otherwise stated.
- **Data handling:** `pandas/numpy` for 15-min time series (demand, PV, biogas, prices), tariff tables, and Monte Carlo (MC) call schedules.
- **Plots & tables:** `matplotlib/plotly` for figures.

Solver footprint and performance (illustrative)

- **Machine:** Intel i5-1135G7 (4C/8T), Windows 11, up to 8 threads.
- **Typical LP size:** 35,040 rows (full-year), 672 rows (stress week).
- **Optimiser:** concurrent LP (barrier + dual simplex); presolve $\sim$1.6 s on the stated machine.

Reproducibility

In order to reproduce the methodology described in this chapter, the aspects presented below are crucial:

(1) Data origin.

- **Inputs (15-min):** 2023 WWTP demand, PV and biogas series (plant-provided); EPEX day-ahead prices (historic); grid tariff schedule (Liander 2023); energy-tax band (business consumption band).
- **Derived sets:** stress week (05–11 Mar 2023); congestion windows (Stage A3); MC call schedules (Stage B–C).

(2) Configuration snapshot. To keep experiments consistent and avoid recording errors, all adjusted inputs are centralised in a single `.yml`/`.json` file. This file records: asset limits (CHP rating and ramps; tank capacity and SoC bounds); grid limits (import/export caps); tariff coefficients (fixed, contracted-capacity, volumetric, monthly peak); rolling-horizon settings (window length, commit period); forecast controls (amplitude error, lead/lag); the CBC cap set; and, when applicable, BESS size/power plus degradation and capital-charge parameters.

(3) Solver & environment versions.

Component	Version / Setting
Python	3.x
Pyomo	6.x
Gurobi	12.0.3 (LP)
Threads	up to 8 (set `Threads=1` for bitwise repeatability)

(4) MC Call schedules (generic recipe). The exact *hour of day* weights and episode length statistics used in this study are based on confidential DSO data, so reproducing the *exact* MC schedules created here is not possible. However, the procedure is generic: what is needed is (i) a likelihood of congestion *by hour* (ideally per month/season), and (ii) a distribution of *episode lengths* (consecutive congested hours). From these, stochastically daily call episodes can be drawn and build an ensemble of schedules.

Inputs needed.

- Hour–of–day weights $w_{m,h}$: for month (or season) m and hour $h \in \{0, \dots, 23\}$.
- Episode–length weights $p_L(\ell)$: for lengths $\ell \in \{1, 2, \dots, L_{\max}\}$.
- Daily episode count distribution (e.g., Poisson with rate λ_m per month).
- Optional rules: forbidden hours (e.g., 01:00–05:00), weekday–only, non-overlap, max episodes per day, etc.

What the code does. For each calendar day in a target week: (1) draw the number of episodes from the daily count distribution; (2) sample distinct start hours using the month's hour weights; (3) sample an episode length for each start and trim to the 24 h boundary; (4) enforce non-overlap; (5) set the call announcement time; (6) repeat to build N Monte Carlo schedules. Set a fixed random seed for reproducibility.

(5) Rolling-horizon (RH) loop - pseudocode. To reproduce the RH experiments consistently, the skeleton below mirrors the implementation. The *solve LP* block represents the actual solver call.

```
t = stress_week_start
peak_sofar = carry_month_peak(t)   # month-to-date 15-min import max (kWh/step)
SoC_tank   = initial_tank_soc
SoC_bess   = initial_bess_soc (if any)

while t <= stress_week_end:
    window = [t, t+24h)
    commit = [t, t+3h)
    calls_known = {blocks with call_dt <= t and overlap window}
    inputs = actuals on commit; forecasts on window\commit

    solve LP:
      - balances, ramping, SoC evolution
      - import cap = CBC limit on steps inside calls_known
      - monthly peak: decision p_new >= peak_sofar and import[t] <= p_new
      - objective = energy + tariffs + incremental peak (p_new - peak_sofar)
                    + (degradation + cap charge if BESS)

    apply first 3h decisions; log costs and states
    update SoC_tank / SoC_bess at end of commit
    update peak_sofar := max(peak_sofar, max(import on window))
    t := t + 3h
```

(5) What to save alongside the figures. (Stages B and C) For each run (cap, BESS size), save: (i) config snapshot; (ii) solver log; (iii) weekly cost decomposition; (iv) call-by-call feasibility flags and unmet volumes; (v) normalised cost metrics (LCCR); (vi) seeds and generated schedule identifiers. This guarantees that outputs can be regenerated exactly.

Strengths and limitations The optimisation set-up is deliberately linear and transparent: system physics are encoded as the exact balance and limit equations used in practice; tariff mechanics (volumetric transport and monthly kW-max) are modelled explicitly; and intertemporal states (tank SoC, CHP ramping, rolling monthly peak) are handled natively. This keeps a clean mapping from decisions to bill items and makes diagnostics straightforward. The implementation in Pyomo is well-documented and flexible, allowing us to reproduce the plant with the desired level of detail and to switch between Perfect Foresight and Rolling-Horizon runs with minimal structural changes. Reproducibility is supported by centralised configuration and fixed Monte-Carlo seeds/schedules, which enable like-for-like comparisons across experiments.

The approach is not without limitations. Computational effort grows with horizon length and with the RH set-up, which requires repeated solves; RH runs are materially slower than PF. Commercial solvers such as Gurobi keep runtimes tractable but require a licence; open-source alternatives (e.g., CBC/GLPK) worked in our tests but were roughly twice as slow. Numerically, the model remains an LP, so problem sizes on the order of a full year at 15-minute resolution are still manageable, but care is needed with peak-tracking and large coefficient ranges.

Scalability and extensibility are good: adding assets, caps, or policy variants typically involves introducing new variables and linear constraints without changing the overall structure. Pyomo's explicit style is verbose, which can make the code denser than domain-specific libraries; the trade-off is flexibility while the formulation evolves. Now that the models are stabilised, a future refactor to a higher-level framework (e.g., Linopy) could improve readability and maintainability without altering the underlying optimisation logic.

5.8 Validation

This section specifies the checks used to confirm that the models behave as intended so that their outputs can be trusted. We proceed from the core model **internal validity** checks to targeted **call mechanism** checks and PF–RH consistency checks. Numerical results from these tests are reported in Chapter 6.

5.8.1 Internal validity of the core model (PF benchmark).

Since the PF benchmark is the base of all subsequent formulations, it is crucial to verify that its physics and tariff mechanics operate correctly. Concretely, we check that mass and energy balances are coherent at every time step; that no simultaneous grid import and export occurs; that SoC ceilings and floors are respected; and that unmet demand is zero. We also compare overall costs against the plant's current behaviour: by purchasing and selling energy more efficiently and *internalising* grid-tariff terms in the optimisation, we expect a lower electricity bill in this case.

A first inspection of Figure 5.6 confirms that contractual grid limits are respected and that there is no simultaneous import and export (exports are plotted as negative areas to indicate flows *out* of the WWTP). Seasonal patterns also appear sensible, solar output concentrates in the summer months, indicating that the model reproduces expected behaviour.

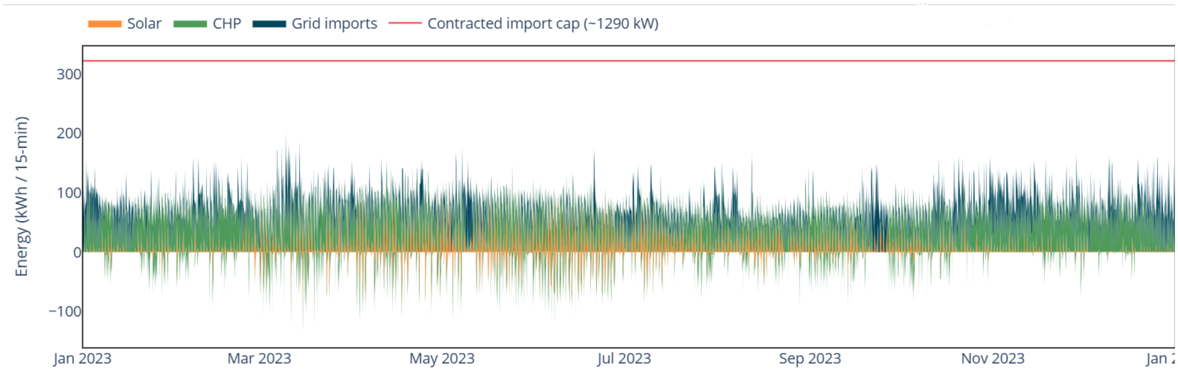


FIGURE 5.6: Yearly Energy flows to WWTP - PF Benchmark

Because the annual plot is visually compressed, we zoom in to the *stress week* to examine dispatch dynamics in more detail. In Figure 6.2 the plant's electric demand is shown as a blue line, and the stacked areas represent the energy flows; the stack exactly matches demand at each time step. Day-ahead EPEX prices are overlaid as a dashed purple line. At this scale it is clear that imports and exports do not occur simultaneously (no negative area is present when the dark-blue import area is non-zero). Also, since the optimisation responds to day-ahead prices, imports concentrate in lower-price periods and exports occur more often during price spikes, which is consistent with the model's objective.

If we also inspect the gas tank state of charge, we see that it stays within limits throughout the whole year. A close-up for the stress week (Figure 6.3) shows the tank cycling roughly twice per day. This is expected given the narrow operating margins and the continuous

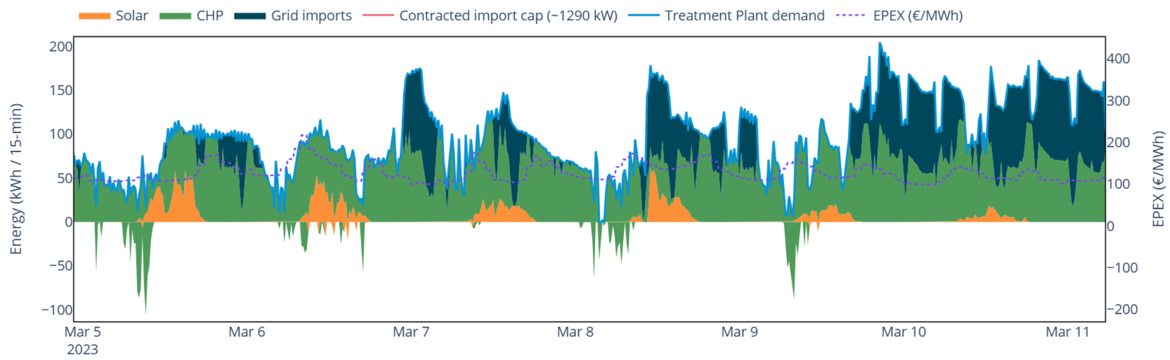


FIGURE 5.7: Yearly Energy flows to WWTP - PF Benchmark - Zoom on *stress week*

biogas inflow from upstream processes, which prevents the tank from idling even when that might be economically preferable.

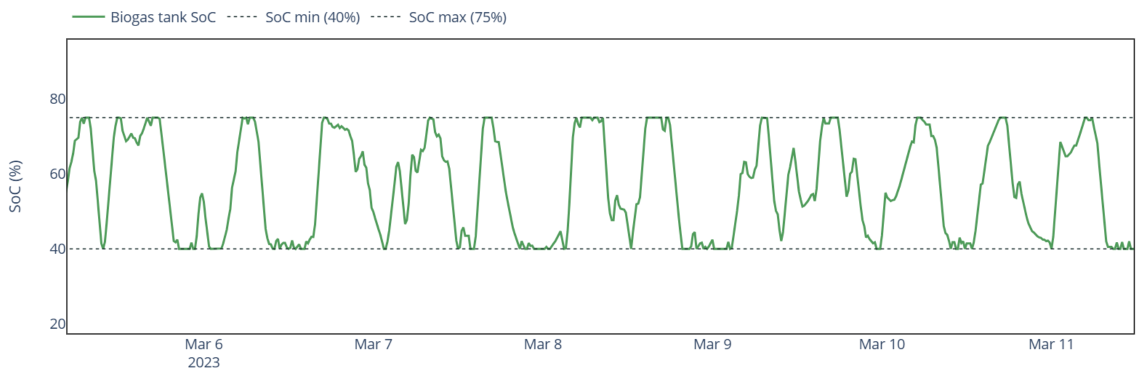


FIGURE 5.8: SoC of gas tank - Zoom on *stress week*

Furthermore, comparing the grid-import profile with the *BAU* case (Figure 5.9) reveals a clear improvement under the PF benchmark. While the two profiles broadly match, the PF solution highly suppresses short, sharp import spikes. Once the kW-max grid tariff term is internalised, such spikes are no longer *worth it*, which is exactly the incentive the tariff is designed to create. Additionally, imports are lower during expensive EPEX hours in the PF case than in *BAU*, reinforcing that the optimisation behaves as designed.

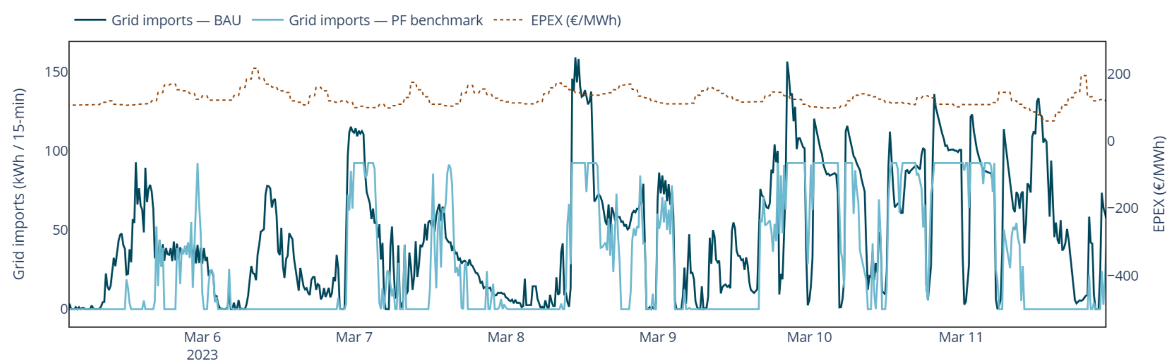


FIGURE 5.9: Grid Import Profile - BAU vs. PF Benchmak - Zoom on *stress week*

Finally, Table 5.10 summarises total electricity costs for the full year and for the stress week. Penalties are zero in the PF benchmark, confirming that all required energy is supplied within the physical limits of the system. Costs are substantially lower than DA prices.

TABLE 5.10: BAU vs. PF benchmark: electricity bill and penalties (year and stress week).

	BAU	PF benchmark
Year (2023)		
Penalties [€]	–	0
Total energy bill [€]	147,305.89	94,783.74
Stress week		
Penalties [€]	–	0
Total energy bill [€]	7,944.21	6,266.81

Taken together, these checks provide a coherent validation: the PF benchmark, and therefore the core structure reused by subsequent models, operates as intended, both physically and economically.

5.8.2 Call-system validity (stress week).

The central question is whether capacity-limiting activations can be applied to this facility. Therefore it is also important to validate the call mechanism itself. For that purpose, we select two easily interpretable, *arbitrary* call windows during the *stress week* whose feasibility can be determined without optimisation (one clearly feasible, one clearly infeasible). We then run the model and confirm that outcomes match expectation. Next, we compare Perfect Foresight with Rolling Horizon on the same set of calls, acceptance rates and total costs from RH should be equal to or worse than PF but never better, because PF is the optimistic upper bound.

Looking once more at Figure 5.9, it shows sustained periods in the stress week where grid imports under the PF benchmark fall to (or very near) zero (e.g., evening of 6th March), and others where imports are higher (around 10–11 March). To construct a clear sanity check, two non-overlapping calls are placed in these contrasting windows and a cap of **50 kW** is enforced in each. Because the calls are well separated in time, the feasibility of one cannot influence the other.

The call schedules tested are:

TABLE 5.11: Example calling schedule for the stress week

block_id	start_dt	end_dt	call_dt
`20230306_1800_test`	`2023-03-06 18:00:00`	`2023-03-06 20:00:00`	`2023-03-05 08:00:00`
`20230310_2100_test`	`2023-03-10 21:00:00`	`2023-03-10 23:00:00`	`2023-03-09 08:00:00`

The outcomes of this test align with the price signal and intuition. The first call, during 6th of March is fully accepted, grid imports remain below the 50kW cap throughout the called window, also coinciding with with high EPEX prices. By contrast, the second call, during the later high demand period, is not honoured: the model requires grid intake above 50kW to meet demand, and the unmet-demand slack becomes positive within that window. These results are illustrated in Figure 5.10.

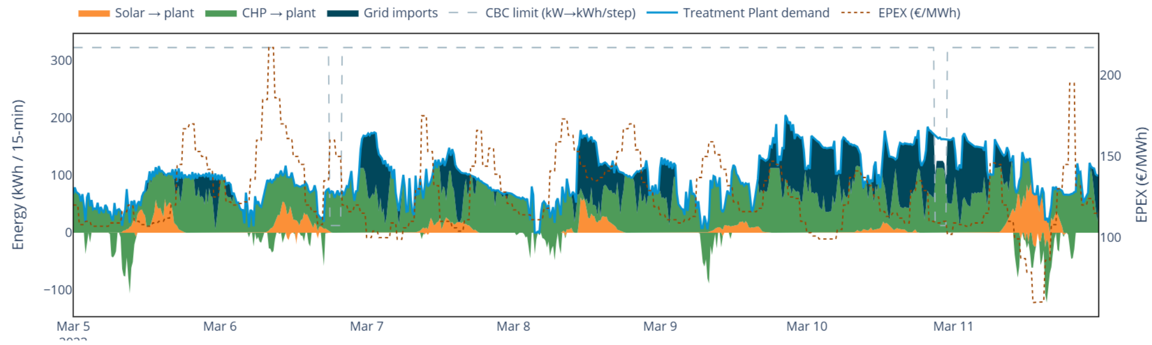


FIGURE 5.10: PF with MC calls - test

We then repeat the test with the same two calls with a **rolling-horizon** solver using a 24h look-ahead. The feasibility pattern and costs match PF for this test: the first call is honoured and the second is not. Rather than indicating a modelling flaw, this indicates that a 24 h horizon is sufficient to pre-position the on-site assets so that PF's decisions are effectively reproduced in this specific case.

To verify this, the horizon is exceptionally shortened to **4h**. The feasibility pattern remains the same (first call accepted, second rejected), but the *volume* of unmet demand during the rejected call increases. This confirms the expected RH effect, with less foresight, the plant cannot prepare to the same extent as under PF or RH with 24 h.

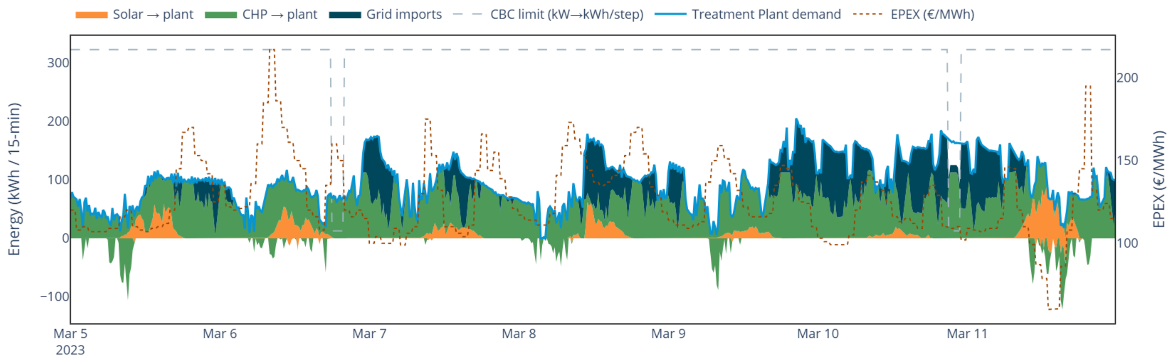


FIGURE 5.11: RH with MC calls - test

The numeric summary appears in Table 5.12: both PF and RH(24 h) accept 1/2 calls with 380 kWh unmet during the rejected episode, while RH(4h) also accepts 1/2 but with a larger shortfall of 500 kWh. These checks corroborate that call enforcement and feasibility logic operate as intended, and that deviations from PF arise when the look-ahead is deliberately curtailed.

TABLE 5.12: Call acceptance and unmet demand across modelling approaches.

	PF MC Calls	**RH MC Calls (24 h window)**	**RH MC Calls (4 h window)**
Feasible Calls	1/2	1/2	1/2
Unmet Demand [kWh]	380	380	500

5.8.3 RH with forecasts and exogenous shocks.

No additional structural validation is required beyond the checks above. The RH with forecasts and the exogenous shock scenarios (price spikes, demand growth, tariff increases) do not alter model mechanics; they only change inputs. Provided the PF/RH validation passes, these variants should behave as expected. Our focus there is on how feasibility and costs vary under imperfect information and external conditions, not on re-validating the modelling framework itself.

Chapter 6

Results

This chapter presents the empirical results obtained from the optimisation models introduced in Chapter 5. We proceed from general validation and baseline characterisation to targeted evaluations of the different CBCs. First, we establish the baseline (no CBC) with operational and cost benchmarks against which all CBC designs are compared. We then report results for *fixed* CBCs under perfect foresight (PF), which provide, best-case benchmarks for import reduction potential, followed by *call-based* CBCs solved in a rolling-horizon (RH) setting, both without and with a battery energy storage system. Finally, we test the preferred designs against exogenous scenario stresses (prices, tariffs, demand) and synthesise the implications for operators and DSOs. Throughout, we report the full bill decomposition, physical KPIs, congestion-relief delivery within called windows and feasibility rates across Monte Carlo (MC) ensembles.

6.1 Stage A — Baseline and capability (PF, full year)

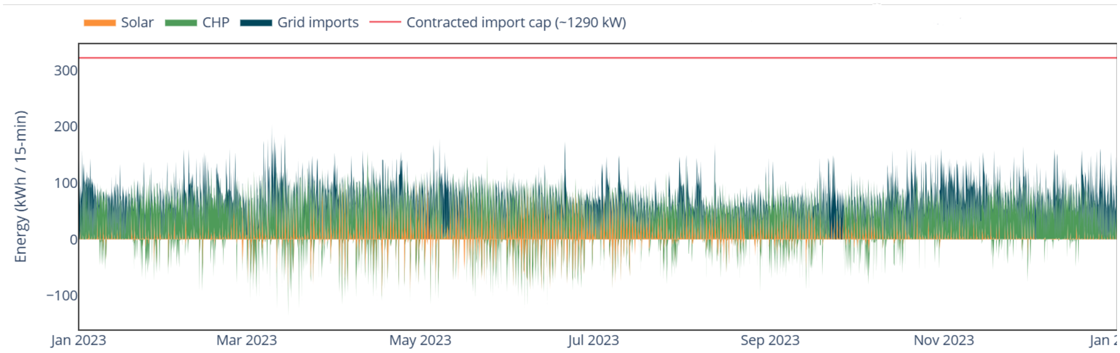
6.1.1 PF benchmark vs. Base Case (full year and stress week)

Firstly, recall from Section 5.8 that internal checks confirm the PF benchmark behaves physically and economically as intended. Building on that, Figure 6.1a shows the full-year dispatch when grid tariffs and DA EPEX prices are internalised in the objective. Grid imports remain well below the contracted off-take limit throughout, and exports never approach the feed-in limit. This is consistent with the plant's demand level and indicates that, without CBCs, contractual caps are not binding in normal operation.

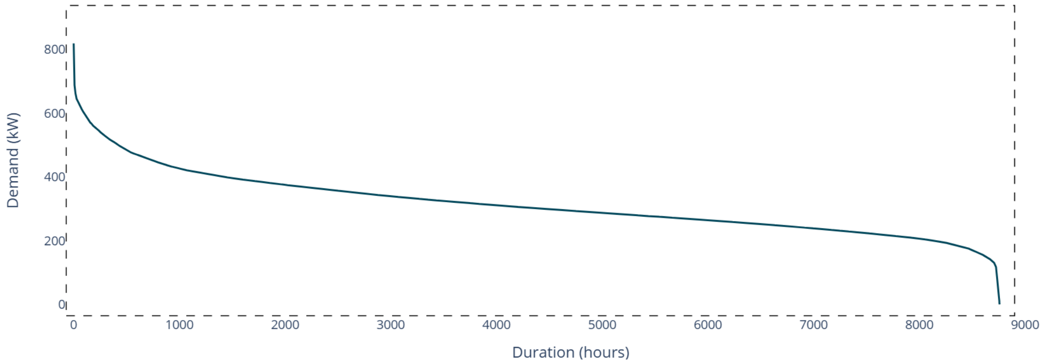
Moreover, the *stress week* zoom in Figure 6.2 clarifies the daily dynamics. Generally, midday imports drop as PV covers part of the load; when PV and CHP exceed on-site demand, the model exports the surpluses (plotted as negative areas). The gas plant adjusts to changes in demand, and the biogas tank enables a modest shifting so that plant output avoids sharp jumps that are not physically feasible. Additionally the import peaks are clipped, due to the implementation of the max peak tariff, particularly in the darker periods when electricity demand is higher and PV production is lower.

Now, we consider how the tariff structure shapes these patterns. Because the objective includes the energy taxed cost of imports and the volumetric transport fee, importing in high price hours becomes comparatively expensive; *additionally*, the monthly kW-max charge penalises tall, isolated import spikes that would set a new monthly peak. Together, these terms encourage flatter import plateaus compared to the BAU and a gas plant profile that ramps within its limits to shave peaks.

Furthermore, the model's treatment of surpluses is economically coherent. PV curtailment



(A) Yearly energy flows to WWTP — PF benchmark.



(B) Load duration curve of WWTP demand.

FIGURE 6.1: Annual operations and demand profile: (a) energy sourcing vs. demand; (b) load duration curve.

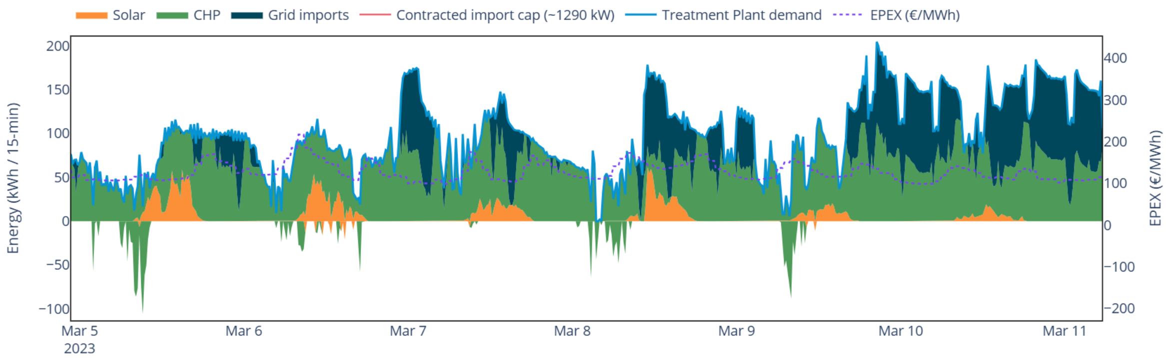


FIGURE 6.2: Yearly Energy flows to WWTP - PF Benchmark - Zoom on *stress week*

carries zero penalty, but biogas flaring is heavily penalised while exports earn revenue; consequently, the optimiser prefers "use or sell" over "spill", which is exactly what is observed in the graph.

In terms of storage behaviour, Figure 6.3 shows a regular tank cycle between 40% and 75% SoC across the *stress week*. Because storage is modelled lossless and costless (and flaring is penalised), the tank naturally fills in and discharges following economically optimal DA scheduling. The gas plants's ramp constraints reinforce this: the SoC swings provide a buffer that allows effective electricity supply from the gas plant without violating ramps. In practice, this means SoC often sits at a bound and then moves quickly to the opposite bound as price and tariff incentives flip across the day.

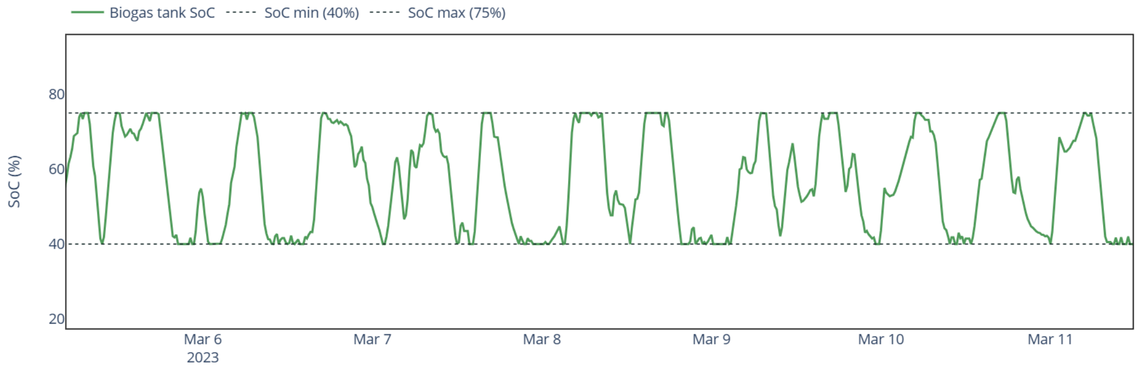


FIGURE 6.3: SoC of gas tank - Zoom on *stress week*

Additionally, this pattern has sizing implications. Frequent saturation at both bounds indicates that the usable band (about 35% of capacity) is fully exploited most days. Given the current parameters, this energy volume is modest relative to the capability of the gas plant, so the tank empties and refills quickly. This is a healthy diagnostic: bounds are respected, flaring is avoided, and the tank provides economically meaningful time-shifting. It also suggests that *additional* flexibility, either a larger tank or complementary storage, could further mitigate imports and monthly peaks.

Finally, this PF dispatch is an appropriate *upper-bound* benchmark against which to evaluate CBC designs. It reflects optimal pre-positioning under perfect information while accounting for the full tariff stack, so any deviation observed under call enforcement or limited foresight can be attributed to those mechanisms rather than to modelling artefacts. Table 6.1 summarises the cost decomposition versus BAU for the year and for the stress week; *importantly*, the stress-week totals are the figures used in **Stage B** to derive CBC compensation on a comparable basis.

TABLE 6.1: Cost breakdown and volumes: Base Case vs. PF benchmark (full year and stress week).

	Full year (2023)		**Stress week (5–11 Mar)**	
Item	**Base Case**	**PF**	**Base Case**	**PF**
Volumes				
Grid imports [kWh]	1,175,940	552,948	30,543	19,307
Grid exports [kWh]	825,925	184,324	13,335	3,655
Market cash flows				
Import cost (variable) [€]	137,788.77	63,172.23	4,936.29	3,062.64
Export revenue (EPEX) [€]	-54,377.80	-16,823.63	-1,516.00	-446.05
Network tariffs				
Volumetric tariff [€]	18,815.05	8,847.16	488.69	308.92
Monthly peak charges [€]	16,382.99	10,973.37	1,648.59	956.75
Fixed transmission fee [€]	441.00	441.00	36.75	36.75
Contracted-capacity fee [€]	28,173.60	28,173.60	2,347.80	2,347.80
Penalties				
Total penalties [€]	—	0.00	—	0.00
TOTAL electricity cost [€]	**147,223.61**	**94,783.74**	**7,942.11**	**6,266.81**

6.1.2 PF yearly cap (no calls)

This experiment identifies the lowest *uniform*, year–round import cap, $\kappa^{\star}_{\mathrm{PF}}$, that remains physically feasible under Perfect Foresight (PF). The resulting value is used to delimit the cap range explored later with call–based CBCs.

Figure 6.4 presents the grid imports and two horizontal lines: the current contracted import of 1290 kW (dashed line) and the *feasible* uniform cap inferred from the PF run (solid line). The PF dispatch, can keep imports low enough that a cap of approximately **500kW** would be feasible, but non-binding for most of the year and would bind only during a small number of intervals. This value therefore provides a robust design threshold for the call–based experiments in **Stage B**.

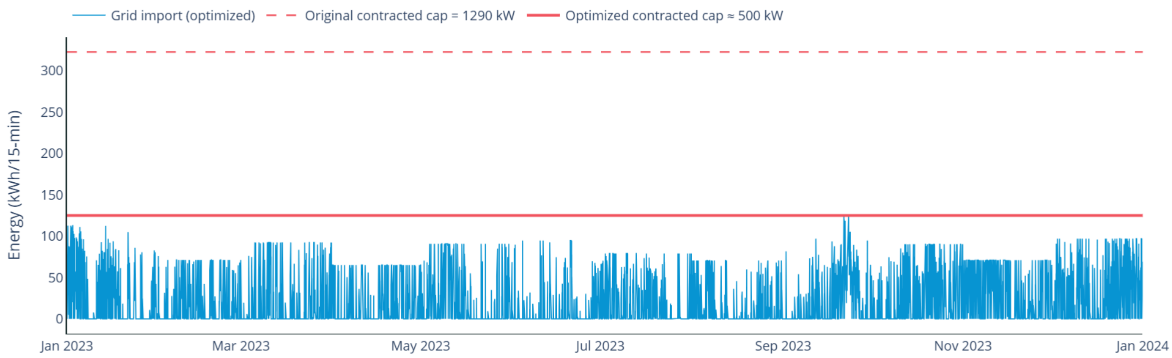


FIGURE 6.4: Yearly grid-cap reduction (Stage A2): PF imports vs. base case and reference caps.

Comparing with the PF Benchmark, imposing a uniform cap at 500 kW leaves the annual import *profile and volume* largely unchanged (Table 6.2). The few binding intervals cluster in September rather than in March, even though March contains the week with the highest electricity demand. This is not contradictory, grid imports are shaped as much by the timing of PV output, the gas plant or the gas tank, as by the absolute demand. In September, earlier PV sunset and evening price/peak patterns make imports relatively more attractive, so the cap binds there first. We nonetheless retain 05–11 March as the common *stress week* for Stages B–C for two reasons. *First*, it was pre-selected on an objective criterion (highest weekly demand), which is a conservative test for tight energy balances: if a design is feasible then, it is likely feasible in less demanding weeks. *Second*, March exhibits grid import patterns that recur across multiple months, providing a more *representative* test bed than September's few isolated binding evenings.

Therefore for the Stages B and C experiments across the 05–11 March week; the cap sweep spans 50 kW to 450 kW, i.e. the entire binding band below 500 kW.

This diagnostic treats the cap as in real life; we do not lower the contracted capacity fee, even if a fixed yearly CBC is in place. If a facility were to implement a year-round cap near 500 kW, the difference in contracted-capacity charges relative to the previous contract provides a simple benchmark for neutral cost recovery. In this case, that value can be seen in Table 6.2.

This experiment can also be related to the GOTORK measure, caps *below* 500 kW would render normal operation infeasible; caps *above* that level are effectively non-binding and would not materially affect transformer loading, though they may help recover persistently unused capacity.

TABLE 6.2: Weekly comparison (5–11 March 2023): Base case vs. PF benchmark vs. PF (uniform annual cap ∼500 kW).

Metric	**Base case**	**PF (benchmark)**	**PF (yearly cap)**
Costs			
Import cost (EPEX) [€]	4,936.29	3,062.64	3,059.55
Export revenue [€]	-1,516	446.05	444.66
Volumetric tariff [€]	488.69	308.92	308.88
Monthly peak charges [€]	1,648.59	956.75	956.75
Fixed transmission fee [€]	36.75	36.75	36.75
Contracted-capacity fee [€]	2,868.06	2,347.80	2,347.80
Contracted-capacity discount [€]	-	-	-1,438.41
TOTAL electricity cost [€]	**7,942,11**	**6,266.81**	**4,826.66**
Energies			
Grid imports (kWh)	30,543	19,305	19,305
Grid exports (kWh)	13,335	3,627	3,627
Solar → WWTP (kWh)	—	5,394	5,394
CHP → WWTP (kWh)	—	41,158	41,158
WWTP supply shares			
Solar share [%]	—	8.2	8.2
CHP share [%]	—	62.5	62.5
Grid share [%]	—	29.3	29.3

In essence, this experiment indicates a feasible uniform cap of about **500kW**, binding only in limited periods. This threshold cleanly defines the *interesting band*, **[50-450] kW** for the call-based designs that follow in mechanism in **Stages B–C.**

6.1.3 PF imports discouraged (no calls)

This diagnostic quantifies the plant's *intrinsic* flexibility during congestion prone hours, without imposing hard caps. We implement a penalty for grid imports within pre-defined congested prone windows (Table 5.5); all physics, tariffs and limits remain exactly as in the PF benchmark. Because violations are still allowed (at very high cost), this result traces an optimistic upper bound on how far imports can be re-timed out of congested periods.

The week zoom in Figure 6.5 illustrates the intended behaviour. Inside the red bands (morning/evening peaks), optimised imports (solid line) collapse to near zero on most days; consumption is shifted towards periods immediately before and after the congested windows. The pronounced rebounds at the edges of some bands are informative, they signal binding physical limits, in other words, flexibility is real but not unlimited. By contrast, the base case (dashed) presents imports across the congestion periods, and larger imports overall.

Over the stress week, total grid imports are 19.388 kWh, of which only 1.803 kWh fall inside congestion windows, about **9.3%** of weekly imports. The penalty therefore succeeds in suppressing the bulk of off-take during red periods, with the small residual driven by the constraints noted above. Weekly costs rise relative to the PF benchmark (Table 6.3): energy and volumetric terms remain close to PF, while the kW-max charge increases due to the edge rebounds that set higher short, sharp peaks. (The artificial penalty is a modelling device and is *excluded* from reported totals.) The energy split (Table 6.3) stays similar to PF but slightly pushes further towards the gas plant as the optimiser substitutes away from imports during congested hours.

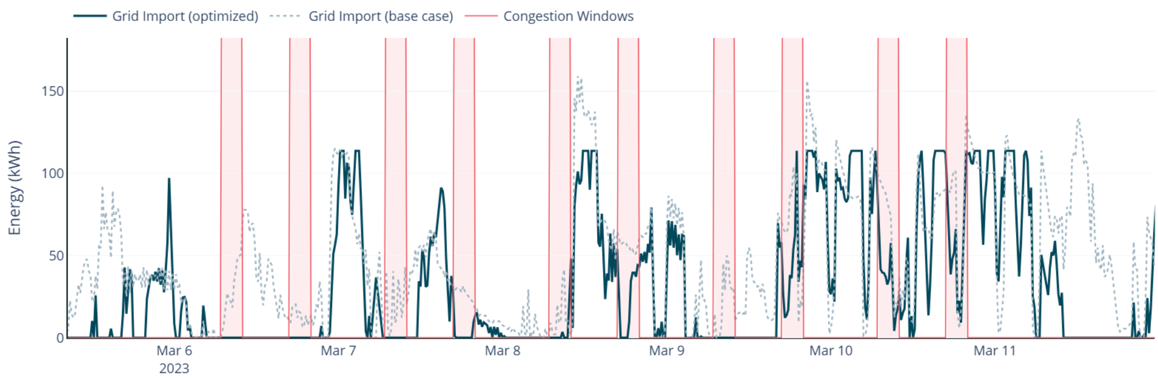


FIGURE 6.5: Grid Imports - PF imports discouraged vs. BAU

TABLE 6.3: Weekly comparison (5–11 March 2023): Base case vs. PF benchmark vs. PF (imports discouraged).

Metric	**Base case**	**PF (benchmark)**	**PF (imports discouraged)**
Costs			
Import cost (EPEX) [€]	3,734.42	3,062.64	3,067.50
Export revenue [€]	-1,516.00	446.05	454.08
Volumetric tariff [€]	488.69	308.92	310.20
Monthly peak charges [€]	1,648.59	956.75	1,178.43
Fixed transmission fee [€]	36.75	36.75	36.75
Contracted-capacity fee [€]	2,347.80	2,347.80	2,347.80
TOTAL electricity cost [€]	**7,942.11**	**6,266.81**	**6,486.61**
Energies			
Grid imports (kWh)	30,543	19,305	19,388
Grid exports (kWh)	13,335	3,627	—
Solar → WWTP (kWh)	—	5,394	5,248
CHP → WWTP (kWh)	—	41,158	41,222
WWTP supply shares			
Solar share [%]	—	8.2	8.0
CHP share [%]	—	62.5	62.6
Grid share [%]	—	29.3	29.4

Taken together, these results provide an upper-bound *flexibility envelope* for the existing layout, prices can materially steer imports out of congested windows, but the residual exposure on those windows reflect genuine operational limits. This envelope is useful in two ways. First, it offers a check for the call-based designs in later stages, if reductions of this order are not achievable under calls, either the cap is too aggressive or additional flexibility is needed. Second, it screens suitability. For instance, a site that shows little avoidance in this test is unlikely to benefit from CBCs. Here, the WWTP shows meaningful, though limited, flexibility; hence proceeding to the call-based experiments is justified.

6.2 Stage B — Design sweep on stress week (PF with MC calls)

In this Stage we tested call–based CBC designs on the stress week (05–11 March) under Perfect Foresight (PF). For each cap $c \in \{50, 100, \ldots, 450\}$ kW, decided in the previous Stage, and for each of 100 Monte Carlo calling schedules we solve the PF model and assess feasibility *per run*: a run is feasible if all the individual calls in that run are, this is, when the *unmet-demand* slack is zero throughout the called window. Weekly costs are computed against the PF benchmark of Stage A.

6.2.1 Current layout: PF sweep with MC calls

The results from Stage A guide the expectations here. First, the yearly feasibility threshold is approximately 500 kW; however for the **stress week** the highest grid imports are around the 400kW level, therefore we expect all caps at 400–450 kW levels to be easy to honour. However, lower, binding caps will be the true test of operational flexibility.

Figure 6.6 shows the fraction of accepted calls across the 100 MC schedules. Acceptance increases monotonically with the cap: starting from ∼62% at 50 kW. There is a pronounced jump from the cap 250kW to 300kW, and 100% feasibility at 400 kW and 450 kW. This matches the Stage A insight, once the cap approaches the 500 kW yearly threshold, calls are almost always feasible with the current assets.

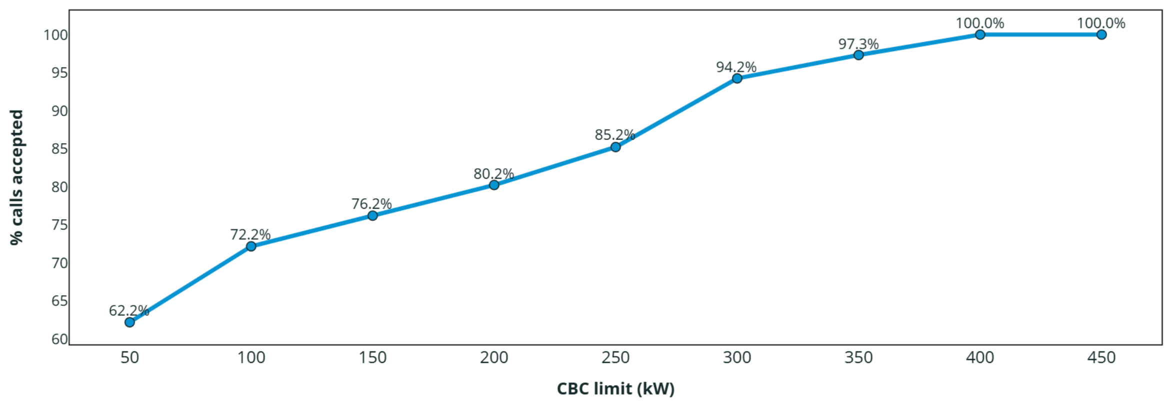


FIGURE 6.6: Percentage of feasible calls vs. CBC limit - Current Layout

Weekly incremental **costs** relative to the PF benchmark are small and decline with the cap, as one would expect (Figure 6.7). Median Δcosts fall from about €3.3/week at 50 kW and 100 kW to €2.1 at 150 kW, €1.8 at 200 kW, and around €1.5–€1.6 for 250 kW to 450 kW. The P10–P90 band narrows steadily as the cap relaxes.

The dotted mean line shows a spike of delta cost at 250 kW where a small amount of schedules need more *effort* to be honoured. The statistics at that cap show a large gap between the *mean* and the *median*, while the P10–P90 band remains low. This indicates that the spike is driven by a small number of outlier schedules, not by a systematic effect across all schedules.

Note that Δ cost is computed only over *feasible* schedules. At very tight caps (≤200 kW), the few runs that achieve 100% acceptance are simply the least demanding in practice, so the average Δ cost stays modest. Around 250 kW the cap sits on a threshold: some schedules that were infeasible at lower caps become feasible, but only by adopting costlier behaviour (e.g., running close to the cap for sustained periods, shifting imports into shoulder hours, and occasionally setting a new monthly peak). These effects inflate the weekly Δ cost. Once the limit relaxes (≥300 kW), the same schedules no longer need such aggressive positioning, and the mean drops back toward the trend.

Overall, the PF optimiser can pre-position gas plant and tank so that the cost increase of honouring calls is modest, even for binding caps.

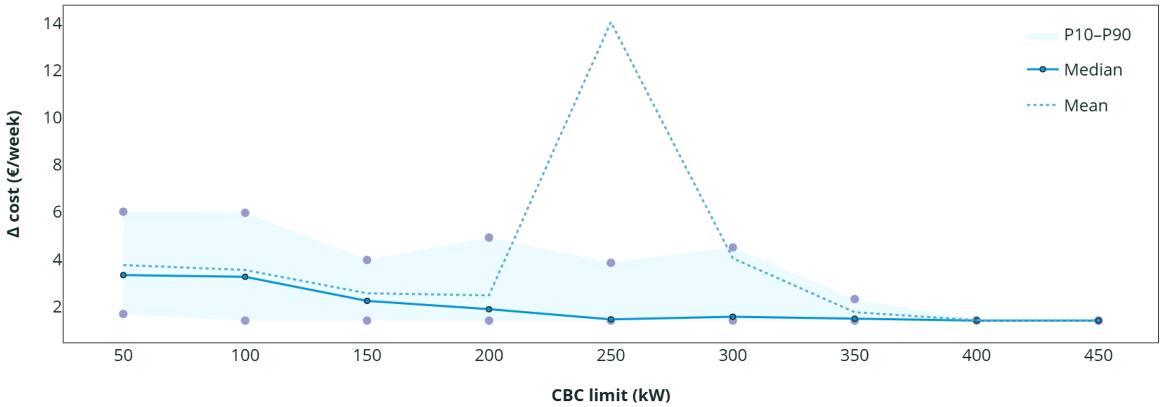


FIGURE 6.7: Cost difference vs. PF Benchmark per CBC Limit (feasible calls)

Figure 6.8 summarises the WWTP supply shares (Grid, Solar, CHP) across all schedules for each cap. As caps tighten, the grid share during called hours compresses and gas plant participation increases, with solar playing a secondary but visible role. At relaxed caps (≥350 kW) the distribution converges to the PF benchmark mix; at tight caps (≤200 kW) the optimiser leans heavily on CHP (subject to power and ramp limits) while keeping imports below the cap.

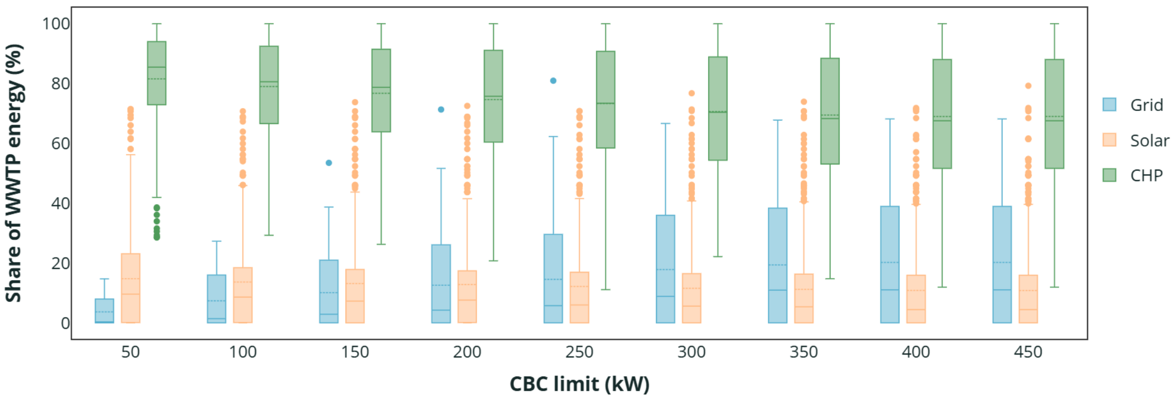


FIGURE 6.8: Resource Share Per Feasible Call

To understand why some calls become **infeasible**, we split schedules by whether the call window coincides with sun (PV present) and compute acceptance rates per cap. Table 6.4 shows that, at tight limits, calls that overlap sun production are easier to honour. For example, at 50 kW acceptance is 65.6 % with sun versus 53.7 % without; by 350 kW the gap has narrowed, and at ≥400 kW both cases reach 100 %. This is consistent with intuition: PV provides additional zero-marginal-cost supply exactly when it is most needed under a tight import cap.

Among the *accepted* calls we also flag those that appear *solar-critical*, i.e., acceptance coincides with non-trivial PV contribution and tight headroom (at the cap for a large share of steps and/or with minimal margin). Table 6.5 shows that this share is around 20–26% for ≤100 kW, then declines to ≤14% by 300 kW, and vanishes for ≥400 kW. In other words, the tighter the cap, the more acceptance hinges on PV timing; as the cap relaxes, reliance on PV coincidence recedes and disappears.

The energy-mix for *infeasible* calls (Figure 6.9) shows a considerable *Unmet* fraction at caps

TABLE 6.4: Acceptance rate by cap and sun presence (stress week, PF with MC calls).

CBC cap [kW]	No sun: accepted [%]	Sun present: accepted [%]
50	53.7	65.6
100	59.1	77.4
150	67.1	79.8
200	67.1	85.5
250	77.9	88.2
300	85.9	97.6
350	92.6	99.2
400	100.0	100.0
450	100.0	100.0

TABLE 6.5: Accepted calls flagged as "solar-critical" (share of accepted, stress week).

CBC cap [kW]	Solar-critical among accepted [%]
50	21.3
100	26.3
150	20.9
200	19.9
250	11.3
300	13.6
350	9.7
400	0.0
450	0.0

≤200 kW (order 10–20%), shrinking steadily and disappearing by ≥400 kW. Two mechanisms explain this:

1. **PV coincidence at tight caps.** When the cap is very low, feasible operation often requires PV to overlap the call window. The acceptance-rate gap between "sun" and "no sun" in Table 6.4, and the higher solar-critical shares in Table 6.5, both point to this dependence.

2. **CHP/tank limits in evening hours.** Without sun, feasibility is constrained by the CHP rating, ramp limits, and the available energy in the biogas tank. If the cap is too low, these constraints force a need of grid imports that is not feasible due to the cap, producing the residual *Unmet* component seen in the infeasible-call mix.

Overall, acceptance improves monotonically with the cap and becomes agnostic to sun presence by 400–450 kW, matching expectations from Stage A (PF benchmark and yearly-cap threshold).

Implications and design takeaways

The sweep aligns with the Stage A diagnostics and clarifies the trade-off:

- **Caps at** 400 kW **to** 450 kW**.** As anticipated from the yearly cap test, these caps achieve 100% acceptance in the PF week and incur negligible incremental cost, but they still represent a substantial reduction versus the current contracted capacity of 1290 kW.

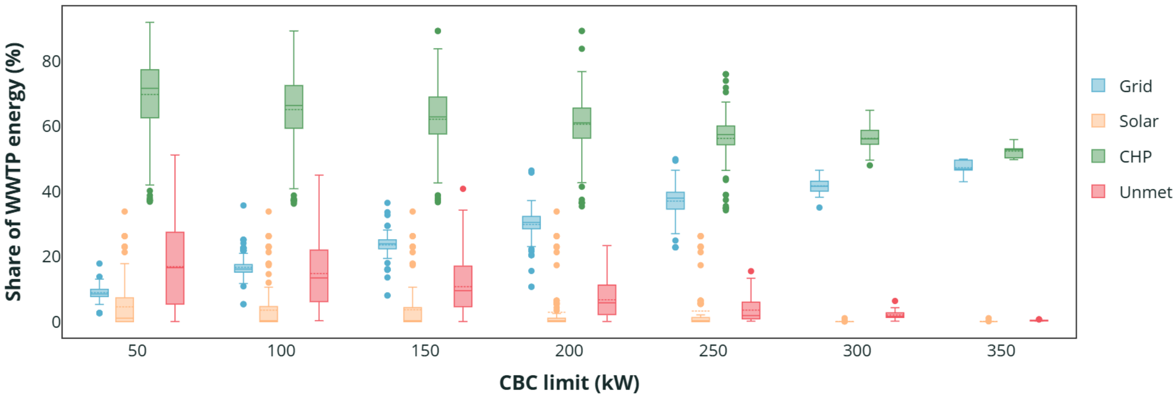


FIGURE 6.9: Resource Share Per Infeasible Call

However, Stage A showed that such high caps are rarely binding over the year. If a DSO nonetheless values the *availability* of this headroom (e.g., as insurance against occasional stress), CBC compensation would be attractive for the WWTP, but the congestion relief impact will be small.

- **Binding caps (≤300 kW).** These are the only levels that consistently reshape imports in the stress week, but they are *not* reliably feasible with the current layout: acceptance ranges 62–94% and failures are concentrated in evening/no-sun calls where CHP/tank limits bind. Under the "honour every call" principle, the current layout is therefore ***not* safe enough to guarantee delivery for materially binding caps.**

BESS integration rationale

The Stage A diagnostics and the Stage B PF sweep point to the same bottlenecks. First, tight caps are hardest to honour in **evening calls without PV**, where feasibility is limited by the **CHP rating**, **ramp limits**, and the **available energy** in the biogas tank. Second, several accepted runs at intermediate caps (e.g. ∼250 kW) achieve feasibility only by **"hugging the cap"** for long stretches and by creating **rebound peaks** just outside constrained windows, both behaviours inflate weekly Δ cost. Third, acceptance at low caps is often **solar–coincident**: Table 6.4 shows acceptance rates rising sharply when the sun is present, and Table 6.5 indicates that a non-trivial share of accepted calls at low caps appears *solar-critical*. Finally, the yearly-cap experiment suggests that very high caps (400–450 kW) are trivially feasible most of the time, but genuinely *binding* caps (≤300 kW) expose the limits of the current layout.

A **BESS** directly addresses the observed constraints and can improve feasibility:

1. **Decouple from PV timing.** Shift midday PV to evening calls, reducing reliance on "sun-coincidence" for acceptance. This should lower the fraction of *solar-critical* accepted calls at tight caps and raise acceptance in *no-sun* windows.
2. **Fast, controllable power for residual headroom.** During a call, the residual net load is $r_t = \max\{0, D_t - PV_t - CHP_t^{\text{feas}} - \bar{\kappa}\}$. A BESS could supply r_t with high ramp capability, relieving the CHP from infeasible step changes and compensating when the tank cannot discharge fast or deep enough.
3. **Peak shaving and rebound damping.** With the monthly kW-max charge internalised, a BESS helps avoid setting a new monthly peak and *dampens rebounds* before/after

called windows, reducing the Δ cost inflation observed at intermediate caps.

4. **Tariff-aware arbitrage that remains economical.** With DA prices known, the optimiser charges in low-price/low-exposure hours and discharges during expensive or capped hours, only when benefits exceed round-trip losses and the degradation charge, preserving economic efficiency while raising feasibility.

5. **Robustness under rolling horizon and forecasts.** In Stage C, finite look-ahead and forecast error erode the PF advantage. A BESS provides a buffer that carries flexibility across windows (state chaining), improving both acceptance and cost stability relative to the current layout.

6. **Complement to the gas tank.** The existing storage band (40–75%) is narrow and frequently saturated. A BESS adds *independent* storage not tied to biogas inflow constraints, increasing the ability to reshuffle energy without penalties or ramp breaches.

These points align with prior findings: the yearly-cap threshold and the PF sweep confirm that 400–450 kW caps are easy to honour, but add little insight where caps are *truly binding*. A BESS is most valuable precisely in the ≤300 kW band where infeasibilities and cost spikes appear; the acceptance asymmetry between *sun* and *no-sun* calls and the measurable share of *solar-critical* accepted calls both point to time-shifting needs a BESS is designed to meet; and the cost spike at 250 kW should be mitigated by BESS smoothing.

Implication for Stage B/B2. We therefore sweep candidate (P, E) pairs and report the Pareto frontier between (i) acceptance of *binding* caps and (ii) normalised cost (LCCR), targeting a knee point that meaningfully raises feasibility for caps in the ≤300 kW range while keeping costs defensible under the tariff structure.

Battery sizes and rating.
Because (i) calls in our MC schedules last 1–6 h and the calls start binding mainly in the 200–300 kW band (leaving residual net loads of a few hundred kW), (ii) evening no-sun feasibility is limited by the 450 kW gas plant rating, ramping and tank state, and (iii) added flexibility must not create new monthly kW peaks during charging, we choose to test a 1 MW BESS at three energies: 1,000, 2,000 and 3,000 kWh. A 1 MW rating provides sufficient *power* headroom to cover typical residuals and damp post-call rebounds without oversizing; the three energy points span ∼1–3 h at full power (and longer at sub-MW residuals), which matches observed call durations. In combination with the 450 kW gas plant, this yields up to ∼1.45 MW of on-site supply during tight windows, directly targeting the evening constraints identified in Stage B.

6.2.2 PF sweep with MC calls (with BESS)

Next we repeat the previous sweep with the three battery sizes (1, 2 and 3 MWh, all with 1 MW power) decided and overlay a weekly capital charge in the cost metrics for the investments. The goal is to see how added flexibility changes (i) call acceptance at each CBC cap and (ii) the price of callable reduction.

Figure 6.10 shows the percentage of calls accepted as a function of the cap. Two patterns stand out. *First*, batteries primarily help at *tight* caps: acceptance jumps sharply with a 1 MWh unit already at 50 kW to 200 kW, and we achieve 100% from 200 kW to 300 kW upwards. *Second*, returns to scale diminish: moving from 1 to 2 or 3 MWh yields only marginal improvements because the binding constraints (CHP rating and short call lengths) are already buffered by the smaller unit. This is consistent with Stage A: by 400 kW to 450 kW the baseline model was already at 100% feasibility, so BESS adds little there.

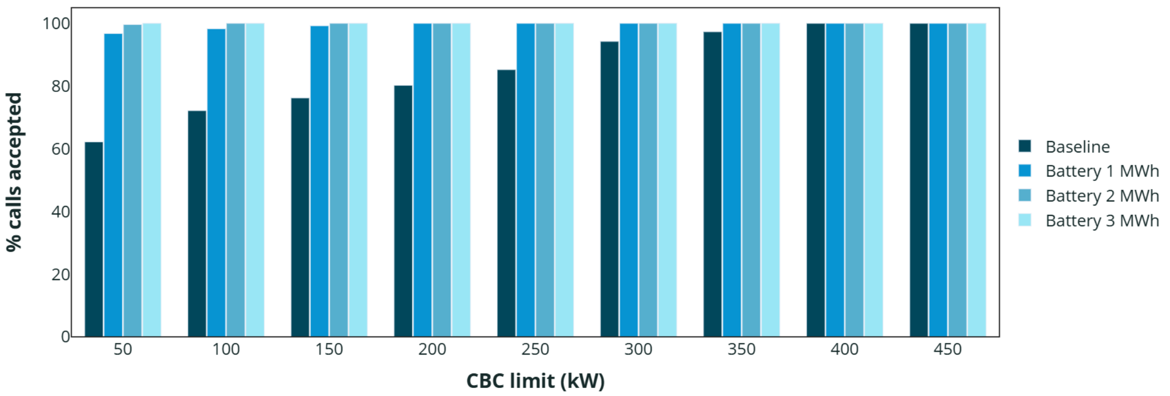


FIGURE 6.10: Percentage of calls accepted vs. CBC Limit vs. BESS size

Figure 6.11 shows the dispersion of the normalised cost (€/kWh delivered) for pairs that are feasible in all 100 schedules. Three points stand out. (i) Larger batteries raise the central price level because of the weekly capital charge; across caps, the 1 MWh unit has the lowest median €/kWh. (ii) Variability is lower at the tightest caps and widens as the cap relaxes: when the cap is very low, feasible schedules force a similar "at–the–cap" operating mode and the €/kWh converges; with higher caps, utilisation becomes schedule-dependent (some calls barely use the BESS, others use it heavily), so dispersion increases. (iii) Outliers are most pronounced for the 3,MWh case: when calls are short or easy, the extra energy capacity is under-utilised and CAPEX dominates the €/kWh metric.

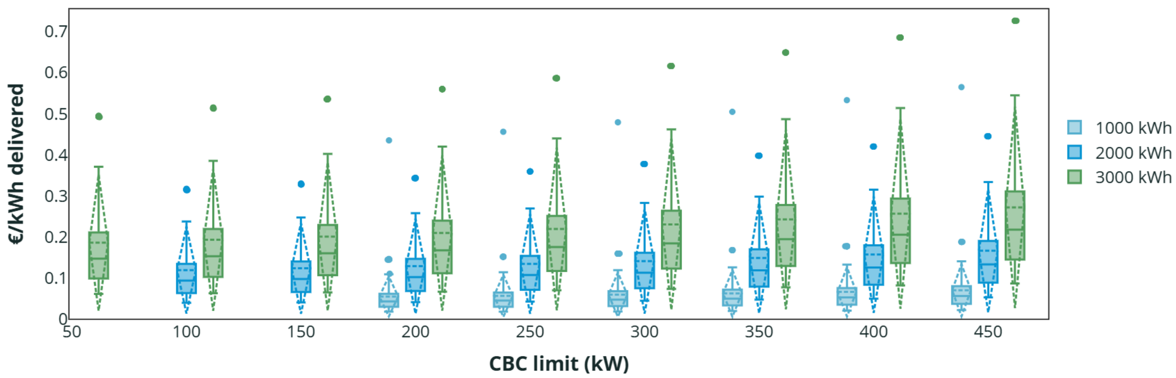


FIGURE 6.11: Price dispersion across feasible pairs (BESS, cap)

To compare designs on a common scale we use the levelised cost of callable reduction (LCCR) and restrict attention to designs that are feasible in *all* schedules. Figure 6.12 plots callable reduction on the x–axis against mean LCCR on the y–axis, and highlights a Pareto frontier: moving along the frontier increases reduction but at rising €/MWh. The knee region provides the best trade-off: sizeable callable reduction at a modest LCCR, beyond which additional reduction comes at sharply higher cost. We select a configuration in this knee as the reference for Stage C : **Battery size = 2 MWh, CBC Limit = 100kW**.

However, two important notes. First, several neighbouring (cap, BESS) pairs around the knee deliver very similar callable reduction *and* levelised costs; the differences are within the dispersion seen in Fig. 6.11. Hence, multiple options are equally defensible and the final pick can be guided by practicalities (space, cycling limits, procurement constraints, expected call frequency). Second, we do not identify a single optimum via a marginal-benefit curve

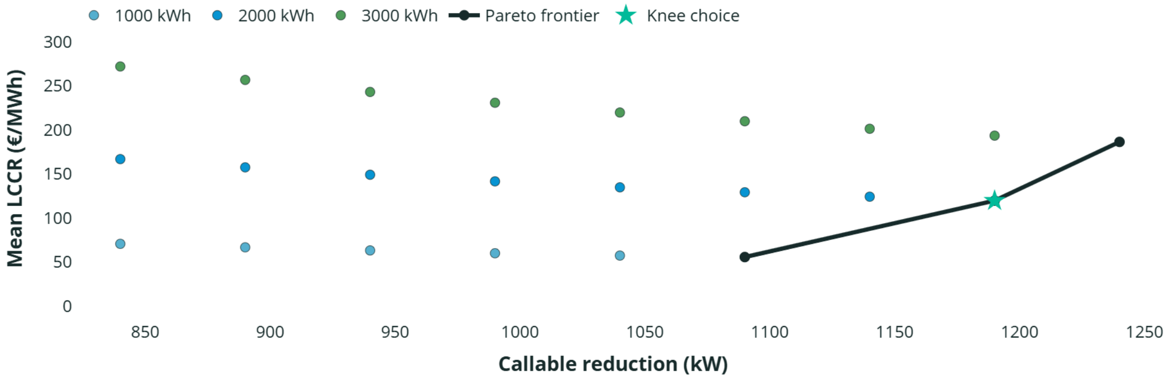


FIGURE 6.12: Reduction vs. LCCR

because the DSO's willingness to pay (value of reduction by time and duration) is not specified. While we could plot a marginal *cost* curve, selecting an optimum requires overlaying the DSO's benefit curve; in its absence, a Pareto frontier is a robust summary, the DSO can later overlay their tariff/WTP to arrive to an agreement, just like in real life, CBCs are negotiated by the two parties.

In essence, batteries unlock high acceptance exactly where the current layout struggles, while keeping the incremental price of reduction reasonable at modest sizes. Larger units (2–3 MWh) produce diminishing operational gains but higher €/kWh due to CAPEX; they would only be justified if the DSO values deeper callable reduction at a commensurately higher price. Finally, recall Stage A's finding that yearly operation rarely approaches the 500 kW threshold: caps in the 400 kW to 450 kW range will often be non-binding in typical weeks. If the DSO nevertheless seeks a firm contractual reduction, small BESS sizes make the binding band ($\leq$300 kW) operationally safe at acceptable cost; otherwise, non-binding caps should not expect frequent activation nor material system impact.

Now, we carry forward the selected $\{\text{cap}, \text{BESS}\}$ to Stage C.

6.3 Stage C — Fidelity and robustness (RH on stress week)

6.3.1 RH without forecasts

Firstly, we isolate the effect of a finite look-ahead and commitment, holding all other modelling choices fixed, and compare against perfect foresight. In this variant the optimiser cannot see beyond a 24 h rolling window, but all data inside the window are known perfectly. As already previewed in Section 5.8, the result for the pair $(\text{BESS}, \text{cap}) = (2000\,\text{kWh}, 100\,\text{kW})$ shows that a 24 h look-ahead is wide enough to pre-position both the biogas tank and the battery SoC, because calls are revealed the day before. This is consistent with the tank plots, which show multiple daily charge–discharge cycles, and with the battery's power–energy ratio, a 2 MWh, 1 MW unit can, in principle, traverse its full energy range in about ≈ 2 h.

Economically, when we remove the week foresight, we reduce intertemporal arbitrage. The observed average weekly cost difference is:

$$\Delta^{\text{RH vs PF}} = 5{,}762.4 - 5{,}718.47 = \mathbf{43.93}\,€,$$

excluding CAPEX, which would be identical in both cases. The difference is small because calls are known one day ahead and both the tank and the battery can be repositioned quickly within each 24 h window.

Although feasibility remains at 100%, this is still optimistic relative to practice. In reality, facilities do have look-ahead but forecasts are imperfect. Simply shortening the window would be artificial; instead we retain the 24 h horizon and introduce forecast error outside the 3 h commit block.

6.3.2 RH with forecasts (imperfect information)

We now quantify forecast-error effects relative to both PF and RH without forecasts. The window and commit settings are unchanged, but information outside the commit block is now imperfect.

This produces visible deviations. Specifically, **eight** schedules become infeasible, dropping the feasibility rate from 100% to **92%**. The mechanism is intuitive, at times the model *over-estimates* on-site generation, behaves as if sufficient local production were available to honour the call (e.g., exports energy that will later be needed), and only when actuals arrive does it discover a shortfall with just 3 h left in the commit block. With gas plant ramp limits, tank bounds and a partially committed battery, the call cannot always be recovered. Similarly, when forecasts *understate* local production, the optimiser may pre-buy or over-charge at sub-optimal prices and then be forced to sell back unfavourably.

Costs react accordingly. Across feasible runs the weekly average rises to **5,856.47€**, a larger deviation from both PF and RH without forecasts. In short, these outcomes are closer to real operation, but they show that under the current configuration feasibility cannot be guaranteed with imperfect information.

Because the shortfall is low, we first test a low-cost operational rule rather than resizing assets. The biogas tank is already tightly constrained and its inflow is not controllable; by contrast, the battery can be explicitly be *directed* towards call honouring. We therefore **force the BESS to reach at least 50% SoC 2 hours before the call start**. Given the call is known the previous day, this target can be met within the 3 h commit block; and then the BESS retains some time to adjust further with perfect data.

This simple rule lifts feasibility: from the eight previously infeasible schedules, **five** become feasible, leaving **three** still infeasible. The weekly average cost also decreases to **5,835.33** €, slightly below the previous RH experiment. Pre-charging reduces late charging during expensive hours and softens reliance on emergency imports, thereby lowering both energy costs and exposure to the monthly-peak charge.

A closer look reveals a common feature of the remaining three infeasible cases: **all the calls exceed five consecutive hours**. Our call generator allows durations from one to six hours, but due to the probabilities, such long calls are rare. Upsizing assets would most likely make these episodes feasible, but doing it for a rarely event is cost-inefficient. Because call-based contracts are negotiable with the DSO, a proportionate mitigation is to limit the *maximum* call length on safety grounds.

When we restrict those specific calls to exactly 4 h and keep the pre-charge rule, feasibility returns to **100%**. The weekly energy cost is now **5,837.96** €, still above PF but now fully feasible under RH with forecasts.

A summary of these experiments can be found in 6.6.

We now assemble the weekly economics for this system, RH with forecasts, 2 MWh/1 MW BESS, 50 % pre-charge rule, and a 4 h maximum call length.

TABLE 6.6: Stage C summary (stress week 05–11 March; 100 Monte Carlo schedules). Energy costs exclude battery CAPEX.

Scenario	Additional rule	Feasible (% of 100)	Mean weekly energy cost [€]
PF + BESS (reference)	None	100%	**5,718.47**
RH, no forecasts	None	100%	**5,762.41**
RH, with forecasts	None	**92%**	**5,856.48**
RH, with forecasts + pre-charge	BESS $\geq$50% SoC at $t-2$ h	**97%**	**5,835.33**
RH, with forecasts + pre-charge + duration cap	BESS $\geq$50% SoC at $t-2$ h	**100%**	**5,837.96**
PF benchmark (no BESS)	None	100%	**6,266.61**

Note: "No forecasts" means perfect information; "with forecasts" introduces imperfect information.

The modelled *energy* bill is

$$C_{\mathrm{RH}} = 5{,}837.96\,€. \tag{6.1}$$

Adding the battery capital charge per week for the 2 MWh unit yields the total plant cost:

$$C_{\mathrm{BESS,week}} = 1{,}694.83\,€, \qquad C_{\mathrm{tot}} = C_{\mathrm{RH}} + C_{\mathrm{BESS,week}} = 7{,}531.83\,€. \tag{6.2}$$

We compare this to the PF benchmark *without* a battery, so that the DSO also internalises the BESS cost. We already revealed that the PF benchamrk weekly cost is $C_{\mathrm{PF}} = 6{,}266.61$ €. The incremental cost the WWTP experiences is then

$$\Delta = C_{\mathrm{tot}} - C_{\mathrm{PF}} = 1{,}265.83\,€. \tag{6.3}$$

Assuming $h = 8$ call-hours in the week, a callable reduction of $\Delta P = 1.19$ MW, and a CBC payment of $p = 120$ €/MWh, the weekly CBC revenue is

$$R(p) = p \times \Delta P \times h = 120 \times 1.19 \times 8 = 1{,}147.301437\,€. \tag{6.4}$$

This does not fully close the gap; the uncovered portion is

$$S = \Delta - R(120) \approx 118\,€, \tag{6.5}$$

which aligns with the observed difference between PF + BESS (which was the model we used to construct the CBC revenues) and RH+forecasts+pre-charge energy costs.

The break-even CBC price per MWh is the value $p^\star$ such that $R(p^\star) = \Delta$:

$$p^\star = \frac{\Delta}{\Delta P \times h} \approx \frac{1{,}265.83}{1.19 \times 8} \approx 133\,€/\mathrm{MWh}. \tag{6.6}$$

Hence, a modest increase from 120 €/MWh would reach break-even. Adopting $p = 150$ €/MWh gives

$$R(150) = 150 \times 1.19 \times 8 = 1{,}428\,€, \qquad \Pi = R(150) - \Delta \approx 162\,€. \tag{6.7}$$

Given that this is one of the highest-stress weeks of the year, a price of **150** €/MWh with $\Delta P \approx 1.2$ MW and 2 MWh of storage provides a prudent margin; in quieter weeks, the same call-hours would typically yield stronger economics.

Under forecasts, the combination cap = 100 kW, BESS = 2 MWh plus a simple 50% pre-charge rule attains 100% feasibility if calls are limited to $\leq$4 h. It delivers a material reduction ($\sim$1.2 MW) and, at 150€/MW h, a workable business case in the stress week. We next stress-test this configuration against exogenous shocks beyond the WWTP's control.

Grid perspective

Now that the chosen configuration achieves 100% feasibility under forecasts, it is useful to look at what the grid actually *sees*, in other words, what happens with the energy drawn from the grid. Table 6.7 compares weekly imports at the PCC for the stress week. As expected, the unoptimised Base Case draws the most energy from the grid. As seen previously, optimising on-site assets against DA prices and network tariffs (A1, PF Benchmark) already cuts imports substantially. Adding the call-based cap and a 2 MWh BESS (C2) cuts a further 0.8 MWh on average across the 100 call schedules. This confirms a **real reduction** from the calls over the stress week. Note that these figures sum the **entire** week. If we look only at congested hours, when relief actually matters, the reduction versus both cases would be larger. Enforcing the caps also causes some *rebound effect* just before and after the calls, as imports rise outside the constrained windows. That spillover dilutes the weekly total, so the true benefit during the calls is even greater than these numbers suggest.

Case	Grid imports (kWh)
Base case (no optimisation)	30 543
A1. PF benchmark (optimised, no CBC)	19 306
C2. RH + forecasts (cap 100 kW; BESS 2 MWh)	18 498*

TABLE 6.7: Grid imports at the PCC during the stress week. * Mean across 100 Monte-Carlo call schedules

6.3.3 Stress tests — exogenous shocks

We finally stress–test the selected configuration against exogenous shocks. Price related shocks (energy market and tariff changes) are not expected to reduce feasibility because they do not alter the physical limits; their purpose is to reveal cost sensitivity. By contrast, a demand surge *could* in principle affect feasibility. All three tests reported below retain **100% feasibility**; we therefore focus on the cost impacts. Results in this section exclude battery CAPEX.

EPEX day-ahead price spike

Feasibility: 100%.
Mean weekly energy cost (excl. BESS CAPEX): 3,129.34 .

Perhaps counter-intuitively, the net weekly energy cost falls well under the spiked prices. With high buy/sell prices, the site's operating pattern (cap-limited imports during calls, and the availability of CHP, PV, the tank and the pre-positioned BESS) shifts the net position towards avoiding expensive imports in peak hours and, at times, earning higher export revenues in shoulders. The combination of capped imports and flexible on-site supply therefore turns the spike into a net reduction in the weekly bill. If battery CAPEX for the 2 MWh unit is added ($\approx$ 1,694.83 per week), the total remains below the normal price PF benchmark without BESS, so the business case is favourable even *before* any CBC compensation is counted.

Demand increase

Feasibility: 100%.
Mean weekly energy cost (excl. BESS CAPEX): 7,636.70 .

Higher demand raises the energy bill as expected. Using the same compensation logic as elsewhere (payment per deployed MW-hour), the indicative weekly margin is

$$\Pi \;=\; p_{\mathrm{CBC}}\,\Delta P\,H \;-\; \bigl[C_{\mathrm{stress}} + C_{\mathrm{BESS,week}} - C_{\mathrm{PF,noBESS}}\bigr],$$

where $p_{\mathrm{CBC}} = 150$ /MW h, $\Delta P \approx 1.2$ MW, H is weekly call hours, $C_{\mathrm{stress}} = 7{,}636.70$ (energy only), $C_{\mathrm{BESS,week}} \approx 1{,}694.83$, and $C_{\mathrm{PF,noBESS}} = 6{,}266.61$. With these inputs the result is *negative* for the tested deployment hours: the CBC revenue does not fully offset the higher energy cost. If such demand bursts are rare, losses would be averaged out over the year; if frequent, they would erode the business case, i.e. there would be little incentive for the plant to participate on a purely financial basis at 150 /MW h.

Grid-tariff increase

Feasibility: 100%.
Mean weekly energy cost (excl. BESS CAPEX): 12,808.16 .

The large increase ihas a simple explanation: higher fixed/volumetric/peak components inflate the bill regardless of CBC participation. Because tariffs are an *exogenous* element, it is not meaningful to compare this figure to a baseline computed under different tariffs. To avoid making the DSO "pay for tariffs", the correct evaluation is to compare *like with like*: benchmark and CBC scenarios must use the *same* tariff set, and the CBC compensation should be assessed against the *incremental* cost relative to that consistent baseline. In short, tariff shocks should be normalised out of the CBC valuation.

Business-case arithmetic (symbols and interpretation)

Let p_{CBC} be the payment in €/MWh, ΔP the callable reduction (MW), and H the weekly call hours. Define C_{ref} as the chosen baseline cost (e.g. PF benchmark without BESS when the DSO internalises storage costs), C_{run} the weekly energy cost in the stress test (excluding CAPEX), and $C_{\mathrm{BESS,week}}$ the weeklyised battery CAPEX:

$$\text{Revenue} \;=\; p_{\mathrm{CBC}}\,\Delta P\,H, \qquad \Delta C \;=\; C_{\mathrm{run}} + C_{\mathrm{BESS,week}} - C_{\mathrm{ref}}, \qquad \Pi \;=\; \text{Revenue} - \Delta C.$$

Under the *EPEX spike* we observe $\Delta C < 0$ (the run is cheaper than the baseline even before CBC revenues), so $\Pi > 0$ at any reasonable H. Under the *Demand increase*, $\Delta C > 0$ and the tested p_{CBC} and H yield $\Pi < 0$ (loss-making); either a higher p_{CBC} or more frequent deployment would be required to break even. Under the *Grid-tariff increase*, ΔC must be computed against a baseline with the *same* tariff regime; only then does Π reflect the CBC effect rather than tariff drift.

Chapter 7

Discussion

This chapter reflects on the methodological choices and findings of the study, discusses their implications for implementation, and offers policy oriented recommendations. The focus is on how Contract-Based Congestion (CBC) can be credibly deployed at a process-driven wastewater treatment plant (WWTP) when flexibility must be sourced primarily from on-site energy assets rather than from the process itself. Throughout, we emphasise what the results mean in practice for both the Distribution System Operator (DSO) and the facility, and I highlight where the approach is strong, where it is limited, and how it can scale.

7.1 Methodological reflection

The study was designed in staged way to separate three questions: (i) what the plant can do in principle (Stage A, full-year Perfect Foresight, PF), (ii) which caps and on-site configurations are attractive on a demanding week (Stage B, PF with Monte Carlo calls), and (iii) whether the resulting designs remain feasible and economically credible once decisions are made with RH, imperfect information and external shocks (Stage C, Rolling Horizon, RH).

This structure had several benefits. First, it avoided mixing technical feasibility with short term forecasting limitations: Stage A revealed intrinsic capability; Stage B provided an optimistic screening under known calls; Stage C reintroduced realism step by step (horizon, then forecasts). Second, by using the same stress week and the same 100 call schedules across PF and RH runs, the analysis enabled one-to-one comparisons that are rare in the literature. Third, by explicitly internalising grid tariffs (volumetric, monthly-peak, fixed, and contracted-capacity) in the objective, the study maintained a direct mapping from dispatch decisions to bill items, which made the results decision-relevant for the facility.

Additionally, the call mechanism is treated as binding once called. There is no option to skip a call. This is deliberate: it clarifies the feasibility frontier and sets a high bar for credible design. In practice, DSOs may tolerate limited non-delivery and issue penalties, or allow small volumes to spill. We do not adopt that here. Introducing fines would erode the business case, and habitual shortfalls, beyond any "reasonable" tolerance, risk programme removal or, in extremis, curtailment/disconnection, since persistent non-compliance can endanger other users (see Chapter 2). On balance, treating calls as strictly binding is both conservative and realistic, and it aligns the model with the contractual intent of CBCs.

The framework was intentionally linear and transparent. This preserved interpretability and computational tractability while still capturing the intertemporal couplings (biogas-tank state of charge, CHP ramps, monthly-peak carry-over) that matter for congestion management. This kept solve times practical, especially for the RH experiments.

One design choice deserves explicit comment. Stages B–C focus on a single *stress week* selected by annual demand. This lowers computational time and yields conservative feasibility conclusions: designs that honour calls in the most demanding week should remain feasible in milder weeks. At the same time, a week cannot represent the entire year's joint distribution of PV, biogas, and load. The study mitigates this by using Monte Carlo (MC) call schedules constructed from empirical patterns, but we acknowledge that the week focus limits exposure to seasonal edge cases (I return to this in Section 7.3).

Finally, forecast treatment in Stage C adopted a realistic split: actuals on the committed slice, forecasts on the remainder of the 24 h window, with errors in both magnitude and timing. This is closer to how facilities like the one studied actually plan, calls known the day before; prices day-ahead; production and demand with uncertainty, and it reveals where feasibility risks arise when the plant is steered by imperfect data.

7.2 Strengths of the approach

A key strength is *traceability*. Because the model internalises the tariff terms actually faced by the facility, cost changes are easy to attribute: reductions in monthly peak, shifts in volumetric charges, and energy arbitrage against EPEX prices all appear explicitly. This is especially important when introducing a battery energy storage system (BESS). Without tariff signals in the objective, a BESS might create import spikes that erode economic gains; here, the optimiser trades energy arbitrage against peak exposure, which produces grid-friendly schedules by design.

A second strength is *comparability*. The study uses fixed Monte Carlo schedules and seeds across PF and RH variants. This control isolates the effect of horizon and forecasts on feasibility and cost dispersion. The same discipline supported the Stage C stress tests, which probed robustness to price spikes, higher demand, and higher grid tariffs without altering the call ensemble.

A third strength is *operational realism* at low modelling cost. The RH set-up (24 h look-ahead, 3 h commit) aligns with how information arrives and with asset dynamics. It preserves intertemporal state carry-over and monthly-peak continuity, which matter for congestion. Importantly, the model remains linear and fast to solve even under RH, so the analysis can be scaled to additional weeks or scenarios.

7.3 Limitations and what they imply

No modelling study is without limitations. We group the most relevant ones and explain their implications.

Week-level focus. Concentrating Stages B–C on one stress week, by design, reduces computational burden and sharpens interpretation. The implicit assumption is that feasibility on the hardest week is sufficient for the rest of the year. While reasonable, it may underrepresent rare but consequential seasonal combinations (e.g., late-autumn low PV coinciding with unusually long evening calls). The mitigation is twofold: (i) the Stage A yearly analysis already showed where uniform caps would bind, and (ii) stress tests in Stage C explored exogenous extremes. Nonetheless, extending the RH+forecast analysis to a small set of additional weeks that differ in solar timing (e.g., September) would strengthen external validity.

Forecast realism. The forecast-error model captures magnitude and timing errors, but forecast quality is site and system-specific. If future operations see more volatile biogas inflows or if solar is systematically mistimed by weather models, the observed 100% → 92% feasibility drop (recovered to 100% with a 50% pre-charge rule and a 4 h cap on call length) could be larger. Conversely, if the operator uses high-quality updates close to real time, feasibility would improve. The key takeaway is not the exact percentage but that forecast misspecification is the dominant source of residual infeasibility once the horizon is finite.

Call honouring. The study assumes that every call must be honoured in full. In practice, DSOs could allow modest non-delivery (spillage) within a defined tolerance and pay for *availability* in addition to *activation*, features that reduce practical risk for participants. We do not include availability payments in the business-case arithmetic; our revenue baseline reflects activation only and therefore represents a conservative lower bound. Modelling calls as strictly binding thus sets a high but useful bar: it clarifies the true feasibility frontier and highlights the minimum operational policies, such as enforcing a 50 % BESS pre-charge at $t-2$ h, that restore feasibility at low cost. In reality, the presence of availability payments (even in weeks with few or no activations) would further improve the facility's economics beyond the conservative figures reported here.

Asset granularity. The WWTP's industrial processes are assumed not to provide schedulable flexibility; all flexibility comes from energy assets (PV, CHP, tank, BESS). This matches the problem statement, but if future audits identified small, time-shiftable loads (e.g., non-critical pumping), those could be layered as additional degrees of freedom, potentially reducing required BESS energy or call prices.

Market scope (day-ahead only). The optimisation and settlement emulate the day-ahead market and Dutch network tariffs; we do *not* model imbalance prices, BRP corrections, or intraday re-dispatch of the site. As a result, any balancing exposure, positive or negative, triggered by CBC actions remains unquantified here. In practice, this could slightly raise (or lower) the true incremental cost of delivery depending on the site's portfolio and BRP arrangements; adding imbalance modelling would tighten the economics.

Mechanism scope (no comparative simulation). While the qualitative positioning of alternatives (redispatch, TBTR) is discussed, the model does not simulate those mechanisms side-by-side with CBCs under identical inputs and constraints. The transfer of insights (e.g., feasible MWh windows, pre-charge rules) is informative but not a proper numerical comparison.

External validity across facilities. Results are derived for one process-driven WWTP with a specific asset mix and tariffs. Although the modelling approach is portable, we have not validated it on other archetypes (e.g., district energy plants, water pumping stations, campuses with BESS). Short pilots at a few additional sites would test generalisability and help calibrate contract parameters to different asset envelopes.

Uncertainty in willingness-to-pay (WTP). DSOs' WTP for callable reduction is not public. The study uses a normalised metric (LCCR) and a Pareto view (reduction vs. compensation) rather than tracing a full marginal-benefit curve. This is appropriate: the Pareto frontier is decision-useful regardless of the eventual tariff; once a target WTP is known, one can pick the point that sits at or below it. We return to price realism below.

7.4 Research Questions answered

SQ1 and SQ2. ***What congestion management programmes are currently available in the Netherlands, and how do they operate and Which of these programmes are technically and operationally suitable for a facility with a fixed, process driven demand profile such as a wastewater treatment plant, considering notice, metering, and enforcement?***

After reviewing the Dutch mechanisms for congestion management, and before any testing, **call-based CBCs** stood out as the most promising fit for this facility. With the current layout (PV + gas unit only), the site proved to be able to hold a 500 kW cap against a 1,290 kW connection. That is a substantial reduction on paper, but offers little practical relief because it rarely binds; and any tighter cap becomes infeasible given the plant's inflexible process load. In other words, with existing assets alone, we cannot guarantee both call compliance and full treatment if we want actual relief at the transformer, so additional flexibility is needed. Once a BESS is added, feasibility improves considerably: with day-ahead notice the plant can pre-position assets and honour caps down to 50 kW while keeping operations. This confirms that **call-based CBCs** are a practical and straightforward choice for this facility.

We can use these CBC results to position the other potential candidates discussed in Chapter 3, **redispatch** and **TBTR**. For **redispatch** (downward) the WWTP could place modest, safe bids if the battery is kept properly pre-charged. This is stricter than under CBCs, since holding more SoC leaves less room for day-ahead trading and pushes costs up. In practice, redispatch looks feasible at volumes similar to our CBC delivery, as long as calls aren't stacked in long back-to-back sequences. For **TBTR**, the present 2 MWh battery can bridge short windows where the capacity is not guaranteed, but the rigidity of fixed time blocks and the risk that capacity outside those blocks is effectively zero make this risky for a process-driven site. Making TBTR robust would require a much larger BESS, more usable gas storage, or process-side flexibility. Any of these paths would raise costs and complexity significantly.

Consequently, call-based CBCs remain the most practical, plant-level mechanism for this WWTP; redispatch is possible at small, cautious volumes; and TBTR is confirmed a poor fit under the current asset mix.

SQ3. ***How can the technical and economic impact of participation be modelled and assessed for such a facility using measured data and Dutch tariffs?***

We build a plant-level optimisation that uses the site's measured 15-minute data (WWTP demand, PV output, biogas production) and enforces all metering at the point of common coupling. The model respects real asset limits, biogas-tank state of charge and bounds, CHP ramping and nameplate limits, and BESS power/energy, and applies a **hard CBC cap** to grid imports during called windows.

Economics are assessed by reconstructing the electricity bill as in practice: **day-ahead** energy **costs** for **imports** minus **export revenues**, plus **Dutch network tariffs**. This lets us compare scenarios on an incremental basis versus a baseline without CBCs. For CBC cases, we generate a **Monte Carlo (MC)** ensemble of call schedules (day/timing/duration) from empirical patterns and evaluate: (a) technical feasibility (no cap breach, no unmet load), (b) grid relief (lower grid imports), and (c) bill deltas.

We then run two complementary executions of the same model. **Perfect Foresight (PF)** solves the full horizon with all inputs known; it provides an upper-bound capability and a clean screening of cap/BESS combinations across many random call realisations. **Rolling Horizon (RH)** mirrors operations: the optimiser looks 24 h ahead but *commits only the first few hours*, then rolls forward. Critically, RH introduces *forecasts* for demand, PV, and biogas over the look-ahead (with day-ahead prices as in reality) while preserving intertemporal carry-over (battery SoC, tank level, monthly peak).

Why this matters: once forecasts are introduced in RH, results **diverge from PF**. Acceptance rates typically dip and cost dispersion widens because the controller no longer "knows the future". PF-only studies therefore **overstate deliverable flexibility**: they are technically informative but not directly replicable in real life. RH with imperfect information is the step that turns a promising design into a **credible** one by testing whether the site can still honour caps when it must plan with partial, noisy information.

In short, the assessment (i) couples a physics-faithful plant model with the Dutch bill, (ii) uses MC calls to map feasible, cost-effective designs under PF, and (iii) *validates* those designs under RH with forecasts to ensure they remain workable in practice. This gives decision-relevant outputs feasibility, grid relief, and incremental bill impacts, without unnecessary complexity, and can be tightened further by plugging in richer forecasts or additional weeks if needed.

SQ4. ***What compensation is required to make participation viable from the perspective of the facility?***

We price CBC participation on an incremental basis: the **DSO should cover just the extra costs the site incurs to hold the cap** (energy and tariff effects during called hours, missed market profit, and asset wear), not the entire electricity bill. Under the preferred design in this thesis (100 kW cap, 1 MW/2 MWh BESS, stress week with 8 call hours and 1.19 MW delivered reduction), an activation price of €150/MWh is sufficient to keep the site whole and slightly positive. This figure already nets out the on-site bill benefits the battery creates (lower monthly peaks, better import/export timing), so the required CBC payment is materially lower than the figures quoted for industry. As reviewed in Chapters 3 and 4, published price figures are in the thousands of €/MWh for new peaking assets and around 1,650 €/MWh for industrial **demand-side** cases, while observing that grid operators often offer only a *few* **hundred €/MWh**, which could cover foregone **export** revenue but rarely the opportunity cost of unserved industrial consumption (reported in the 1,200–4,500 €/MWh range), however, our results prove otherwise.

In practice, compensation often combines a small availability retainer (for readiness) with an activation fee per MWh of delivered limitation. The activation fee scales with how often and how long the DSO calls; fewer hours generally imply a higher €/MWh, more hours a lower €/MWh, for the same annual outcome. For generators, contracts often include foregone-revenue terms; for demand-side sites like this WWTP, the right benchmark is the incremental cost of honouring the cap, since no production is shed.

The broader message is encouraging: even a process-driven facility that wouldn't normally see itself as *flexible* can provide dependable relief with asset-based flexibility at workable prices. With an activation price around the level mentioned above, and an availability component to smooth quieter weeks, the business case is credible for the plant while remaining attractive for the DSO relative to costlier alternatives.

SQ5. ***How do external conditions, and investments in flexible assets, influence the integration of the congestion management programme and its economic attractiveness?***

External conditions mainly change the capital, not the physics, provided the site can position its assets. In our stress tests, the *EPEX price spike* scenario preserved feasibility at 100%. Interestingly, the average weekly energy cost *decreased* (excluding battery CAPEX).

This counter-intuitive result is consistent with the model's tariff-aware logic. It shifted imports out of expensive hours and monetised exports when prices were high.. The lesson is that price volatility alone does not harm feasibility and may even **help** the economics when flexible assets can respond.

The **uniform demand increase** also remained feasible, but raised weekly cost enough that a flat €150/MWh activation price turned the week slightly negative. Two caveats apply. First, if higher demand is a brief burst, losses in one week are likely balanced by gains in other weeks when demand is lower; second, if demand growth persists, contract parameters need retuning: the cap level, deployment expectations, or the activation price must be adjusted to maintain a fair balance.

Tripling key grid tariffs similarly preserved feasibility but raised absolute costs substantially. Interpreting this correctly requires care: tariff increases are *system-wide* and would affect the baseline **with** and **without** CBC. It is therefore inappropriate to charge the DSO for the absolute tariff uplift. The right comparison keeps tariffs consistent across baseline and with-CBC runs, and evaluates the *incremental* cost of honouring calls.

Investment in a **BESS** shows that is a strong lever for both integration and economics. Operationally, a modest battery sized to typical call length lets the plant pre-charge, bridge calls, and keep treatment whole; simple rules (e.g., a BESS forced precharge before a call) can restore perfect compliance under forecast error. Financially, the battery has a dual use: outside CBC hours it naturally reduces the site's own bill (monthly peaks, import timing, export gains), and during calls it supplies dependable reduction at moderate activation prices. Even if the DSO helps fund the battery indirectly, using part of its capacity for day-ahead trading lowers the required CBC compensation because some benefits are already earned on the site's tariff bill. Risk is therefore low: most CBC designs include a small availability payment that cushions quiet weeks, and the BESS delivers value even with **zero calls**. In short, with sensible programme safeguards (day-ahead notice, realistic caps, max call duration) and a right-sized BESS, the plant integrates smoothly, remains robust to external shocks, and offers the DSO reliable relief at an attractive cost.

Having addressed each sub-question individually, we now answer the thesis's main research question in detail:

What is the potential for an industrial facility with inflexible electricity demand to support congestion management on the Dutch medium voltage grid, and under what conditions can participation be economically attractive for both the facility and the DSO?

Potential. A process-driven industrial site with an inflexible load can make a material contribution to congestion management on the Dutch MV grid, providing flexibility from energy assets without need of altering the core treatment processes. In our WWTP case, coordinating PV, CHP and a modest battery turns a rigid plant into a predictable, callable resource. With day-ahead notice the site can pre-position state of charge and reliably hold tight, binding caps during calls without touching treatment. Under rolling-horizon control with realistic forecasts, simple operating rules restore perfect compliance. In short: **asset-based flexibility at the PCC can dent transformer flows when and where it matters.**

When is it attractive for both sides? Economically, the business case is sound when compensation matches incremental costs (energy, tariffs, wear) and credits the battery's on-site co-benefits (lower monthly peaks, better timing of imports/exports). In our stress week, an activation payment around €150/MWh (and assuming a modest availability fixed fee) was sufficient to keep the facility whole while delivering dependable relief. For the DSO, paying for targeted, day-ahead reductions at this order of magnitude compares favourably with redispatch, denial of connections, or accelerating reinforcements. The risk is low for both parties: calls are predictable day-ahead; the battery pays for itself through tariff savings even in quiet weeks; and availability payments cushion variability in activations.

Conditions for success. The ingredients are straightforward: (i) day-ahead notice aligned with market timelines, (ii) clear metering rules (hourly cap evaluated on 15-min data with a small tolerance to avoid "spike" breaches), (iii) feasible caps and durations (tight enough to help the grid, not so low or long that they regularly fail), (iv) simple operating policies, and (v) transparent settlement based on incremental cost against a consistent baseline. With these in place, external shocks (price spikes, tariff changes) mainly move money, not physics; only large demand shifts require retuning cap levels or €/MWh.

7.5 Practical implications

Three operational insights stand out.

First, *tighter but binding* caps are the right target if the DSO seeks to physically alter transformer flows. Stage A's yearly-cap analysis showed that 400–450 kW caps would be non-binding much of the year; this is consistent with Stage B's PF feasibility: caps at 400–450 kW reach 100% acceptance precisely because the plant rarely needs to get that low. These caps deliver large nominal reductions relative to the 1,290 kW connection but may rarely bind; therefore, they are less impactful for congestion relief during stressed hours. Caps in the 150–300 kW range are where the binding action is, and the feasibility levers are timing (PV coincidence), CHP/tank dynamics, and BESS energy.

Second, *simple policies beat complex redesign*. The 50% BESS pre-charge at $t-2$ h (within the 3 h commit) is a low-cost rule that materially improves robustness under forecast error. It exploits the fact that calls are known the day before, that the battery has a 1 MW power limit and 2 MWh energy capacity, and that most calls are shorter than four hours. The rule simply prevents entering long calls from a depleted state.

Third, *size to the call profile, not to the annual peak*. Oversizing a battery for the rarest, longest events raises the normalised €/kWh delivered because CAPEX dominates when call energy is small. The box-plot dispersion analysis in Stage B with BESS confirms that per-MWh cost increases with energy capacity when activation energy is limited. This argues for (i) using programme design to cap maximum call length on safety grounds, and (ii) selecting a battery that fills the typical timing gap rather than the absolute annual peak.

7.6 Scalability and generalisability

We now consider whether the approach demonstrated here can be scaled to other facilities and contexts. Scalability has two dimensions. First, computational scalability, that is, whether the modelling set up and the rolling horizon controller remain tractable as the scope grows. Second, external validity, that is, whether facilities with different asset mixes and operational constraints can deliver reliable reductions under call based capacity limitation with similar design choices.

On computation, the linear programme with a rolling horizon and explicit tariff terms solves quickly and scales linearly with the number of time steps and assets. Adding additional weeks, alternative call schedules, or a small number of extra devices increases runtime in a predictable way without changing the underlying mathematics. Moving from a single facility to a small portfolio is straightforward when interactions are only through a common feeder cap. In that case a shared capacity constraint can be added and the objective left separable, or a distributed formulation can be used if data separation is required. These features suggest that, from a tooling perspective, the framework is ready to be applied to other facilities with limited adaptation.

On external validity, the central finding is that contract based capacity limitation is safer or seen more positively, where flexibility comes from energy assets rather than from core processes. The method is therefore directly transferable to process driven sites that already combine dispatchable generation or controllable imports with a buffer and, if needed, a modest battery. Where such assets are missing, feasibility will depend on the willingness to adjust production, in which case compensation must cover opportunity cost and the price per unit of delivered reduction will be higher.

Facility archetypes and expected fit

Table 7.1 summarises the qualitative fit of call based CBCs across a set of representative archetypes. The mapping reflects the logic used in this study.

TABLE 7.1: Qualitative scalability of the approach across facility archetypes.

Archetype	Typical controllable energy assets	Expected fit with call based CBC	Key adaptations needed
Wastewater treatment, anaerobic digestion sites, district energy plants	CHP with fuel buffer, on site PV, small battery, thermal or fuel storage	High. Same logic as the case study, caps can be met with asset co-ordination and simple pre charge policies	Calibrate storage sizes and ramp limits, align cap levels with safety limits and maintenance windows
Water pumping stations and flood control assets	Pumps with limited scheduling freedom, possible battery at the substation, limited local generation	Moderate to high. Short, targeted caps feasible when hydraulic headroom exists	Define no risk envelopes for water levels, restrict call duration, add availability payment for readiness outside storm events
Food and beverage with digestion or large refrigeration, cold stores and warehouses	Refrigeration load with thermal inertia, possible CHP or boiler, batteries becoming common	Moderate to high if thermal buffers are used as the primary flexibility source	Translate caps into process safe set points, codify maximum off set periods and recovery trajectories
Data centres and campuses with site batteries	Uninterruptible power supplies, battery systems, sometimes on site generation	High for short windows. Caps map to battery discharge while preserving service level agreements	Firm limits on call length, explicit state of charge floors, clear notice and test calls to validate procedures
Commercial campuses with PV and battery, municipal clusters	PV and battery behind one or several meters, limited process coupling	Moderate. Feasibility driven by battery energy and alignment of caps with solar timing	Prefer time block based caps, size battery to typical call energy, consider group CBC for diversity effects
Continuous high temperature industries without energy buffers (glass, cement, aluminium)	Limited on site storage, process is the main energy sink	Low under strict caps without process change. Feasibility requires production adjustment	If pursued, use availability plus activation with high compensation, or use time block transmission rights instead of hard caps

What holds beyond this case, what not

Three elements travel well across facilities. First, the day ahead rhythm. Publishing caps before the day ahead gate closure allows sites to plan, to pre charge batteries, and to schedule generation. Second, metering clarity. Defining the averaging base and any short tolerance (e.g., allowing one 15-min interval to exceed the hourly cap by up to 5–10% without counting as a breach) removes ambiguity and reduces inadvertent breaches. Third, the staged modelling approach. Separating capability under perfect foresight from rolling horizon realism reveals which constraints are technical and which are informational, which supports credible contract design.

Two elements require particular care when transferring the approach. The first is the definition of safe operating envelopes. Each facility has non negotiable and can vary greatly from industry to industry. Caps and call lengths must be chosen within those envelopes and validated with test calls. The second is forecast quality. Where state carry over is important, for example tank levels or battery state of charge, forecast error at the call boundary is the main source of residual risk. Simple operating rules, for example minimum state of charge at a defined lead time, can mitigate this risk without making the programme complex.

Programme design cues for scale

For distribution system operators seeking to scale CBCs across diverse portfolios, three design cues emerge from this work. First, target windows rather than broad days, keep calls aligned with the predicted congestion window and cap the maximum consecutive duration to preserve reliability. Second, use availability retainers in addition to activation payments where readiness has a real cost, especially at sites where calls will be infrequent or where pre charging imposes a standing requirement.

Overall assessment

In summary, the approach scales computationally with little friction and generalises best to sites where flexibility can be delivered by energy assets while core processes remain untouched. Similar process driven utilities and plants with fuel or thermal buffers are strong candidates. Commercial campuses and data centres can participate reliably for shorter windows if batteries are present. Industries without energy buffers and with high opportunity cost of curtailment are less suitable for strict caps, and will require different contractual forms or higher compensation. The modelling framework and the operational policies used here provide a clear path to test those cases before committing to programme rollout.

7.7 Recommendations for industrial operators

We recommend a staged approach, similar to the one developed in this study, to build a comprehensive understanding of what a site can deliver under contract based capacity limitation. Begin by establishing a perfect foresight baseline to reveal intrinsic technical capability without the noise of forecast error. Then move to a rolling horizon set up that mirrors how information arrives in practice. This progression separates what is technically possible from what is operationally reliable, and it highlights the simple operating rules, for example a minimum state of charge at a defined lead time, that keep feasibility high once uncertainty is introduced.

Building on that foundation, it is important to define a clear safe operating envelope before any contract is signed. Translate non negotiable limits for quality and safety, the maximum consecutive call duration the site can sustain, and the required recovery times into plain operating rules. Encode these as standard procedures for the plant and as explicit contract terms. When the cap and the window sit inside this envelope, call acceptance becomes routine rather than a case by case decision.

With scope and rules in place, attention should turn to day ahead operations. Align forecasting, scheduling, and metering with the contract rhythm. Prepare the energy management system to ingest caps the day before, to pre charge storage in time, and to schedule generation consistently with market nominations. Keep verification simple and auditable by using the agreed averaging base and tolerance. Maintain a light state estimation for key buffers so that the plant enters called windows with known reserves.

Asset decisions benefit from the same pragmatism. Size new flexibility assets to the call profile rather than to the annual extreme. Batteries that bridge the typical timing gap deliver more value than batteries sized for rare events, especially when activations are infrequent. Additionally, credit on site co benefits, such as lower monthly peaks and improved self consumption, in the investment case and in internal budgeting. These gains exist regardless of the contract and effectively reduce the compensation level the site must seek.

Contracting should follow a test and learn path. Request commissioning test calls and periodic drills to validate procedures and forecasts, and ask for availability retainers where readiness has a standing cost. Include asset specific safety rules such as state of charge floors and maximum consecutive hours so that delivery expectations remain realistic. After each event, share a short log of what happened and what was learned with the operations team and with the distribution system operator, then refine rules accordingly.

Finally, organise for delivery. Assign clear responsibilities for day ahead planning, intraday monitoring, and post event review, and keep a short playbook that lists pre call checks, guardrails during calls, and recovery actions after calls. This governance discipline is often as important as technology. Taken together, these steps turn contract based congestion from a one off trial into a dependable service, with predictable effort for the site and reliable relief for the grid.

7.8 Policy recommendations

CBCs will be most effective if programme design principles recognise how industrial flexibility actually materialises at the meter.

First, *targeted windows, clear notice, and feasible durations*. The study assumed calls are known the day before and showed that a 4 h cap on call length can eliminate the last percentile of infeasibility. DSOs should prefer finely targeted windows aligned with congestion risk, preserve day-ahead announcement times, and allow asset-specific safety limits on consecutive call length. These features increase reliability without materially reducing available reduction.

Second, *compensation structure that matches cost drivers*. A two-part structure: availability retainer plus activation payment, remains appropriate. Availability covers fixed readiness and supports modest operational rules like pre-charging; activation should reflect the incremental energy and tariff cost during the called windows. Using incremental baselines (same tariffs, same exogenous shocks) avoids shifting system-wide changes onto the CBC price.

Third, *test calls and learning periods*. Programmes should include test calls at commissioning and periodic train calls. The study's RH+forecast results underscore that forecasting and state carry-over create the residual risk; test calls reveal whether operational rules are sufficient, or whether minor retuning is needed.

Fourth, *enable asset-side investment by recognising co-benefits*. When flexible assets such as BESS improve the facility's own bill (e.g., lower monthly peaks), a portion of the investment is naturally justified on-site. DSOs benefit if contracts acknowledge those co-benefits and price only the incremental part necessary to guarantee reduction. This narrows the perceived gap between consultant-level figures for generic demand curtailment and the prices that are workable in practice for asset-based flexibility.

7.9 A critical reflection on call-based CBCs

After this discussion we can now reflect clearly on the call-based CBCs strengths, limits, and practical implications.

Call-based CBCs are a pragmatic fit for MV industrials: they act at the PCC, can preserve core processes, and use day-ahead targeting so reductions land in the hours that matter. They are bilateral and configurable, which lets DSOs respect operational limits (e.g., max consecutive hours, higher demand periods) and lets sites codify simple rules that raise reliability. Under thiese rules, settlement can be simple and auditable (availability + €/MWh delivered), which builds trust and speeds uptake.

However there are some points were CBCs need care. Three design frictions recur. First, forecast dependency. CBCs on their own can't fully eliminate congestion. They shape demand day-ahead, but because forecasts can miss, they must be paired with nearer-real-time tools (e.g., redispatch) to catch residual overloads. Second, metering detail matters: hourly caps on 15-min meters without tolerance invite inadvertent breaches; these must be specified cleanly. Third, system effects: reductions during calls can create rebound imports just outside the window; DSOs should therefore target calls narrowly and diversify timing across participants to maximise net feeder relief.

The mechanism would benefit from light standardisation and clearer risk-sharing: (i) publish uniform measurement and verification rules (averaging base, tolerance, breach treatment); (ii) include operational safeguards in every contract (max consecutive call length, seasonal/time-of-day scoping); (iii) formalise test calls at onboarding and annually; (iv) align CBC activation and metering with nominations imbalance to avoid unintended deviations; (v) use two-part pay (availability + indexed €/MWh) tied to incremental costs; and (vi) where multiple sites share a feeder, coordinate calls to avoid activating them all at once and prioritise portfolio scheduling. Finally, if CBC contracts last several years and include the battery's other savings for the site, the cash flow becomes predictable. That makes it easier for DSOs or even third parties to help finance the battery, cutting the site's upfront costs.

In essence, CBCs are not a substitute for grid reinforcement, but they are a credible, near-term tool that, when designed with these fixes, deliver dependable, measurable relief at competitive cost.

Chapter 8

Conclusion

The central message of this study is that credible and reliable delivery under contract based congestion at a wastewater treatment plant does not require intrusive process changes or prohibitive compensation. By coordinating existing energy assets and, where appropriate, adding a modest battery, the site can honour tight, binding caps with high reliability under realistic forecasting. The staged methodology makes the sources of residual risk explicit. Once perfect foresight is relaxed, feasibility depends on the asset state at call start and on forecast quality. Both are manageable with simple operating rules, for example a minimum state of charge before the window and sensible limits on consecutive call length.

These findings scale beyond the single facility. The approach generalises well to process driven utilities and campuses that already combine dispatchable generation or controllable imports with a buffer, for example sites with combined heat and power, thermal storage, or batteries. Shorter windows can also be served by data centres and commercial campuses where batteries are present. In contrast, continuous high temperature industries without energy buffers will find strict caps less suitable unless production can be adjusted, in which case higher compensation or alternative contract forms are needed. Computationally, the linear rolling horizon set up remains tractable as scope grows, which supports replication across additional weeks, scenarios, and sites.

For distribution system operators, the the outlook is positive and realistic. Programmes designed around targeted windows, day ahead notice, clear metering rules, and feasible durations can yield dependable relief at a much competitive cost that stated in the literature. For industrial operators, the path is equally practical. Adopt a staged assessment to separate technical capability from operational reliability. Define a safe operating envelope and encode it in procedures and in contract terms. Align day ahead forecasting, scheduling, and metering with the contract rhythm, and size new assets to the typical call profile rather than to annual extremes. Run practice calls and use a short checklist so delivery becomes routine. Together, these steps move the conversation away from headline figures and toward implementable contracts that reduce congestion where and when it matters, while keeping risk and effort predictable for both the facility and the grid.

8.1 Limitations and future work

The main limitations concern scope, representation, and pricing. The analysis centres on a single site and assumes that core processes are fixed, so all flexibility comes from energy assets at the meter. The network is modelled through a contract based capacity cap rather than explicit feeder power flows, which is appropriate for contract testing but abstracts from local electrical interactions. Also, very relevant is that rolling horizon results are validated on one stress week with Monte Carlo call patterns, a demanding setting that clarifies feasibility under pressure but it does not span a full year's distribution of weather, demand, and biogas. Moreover, forecast errors are simplified, calls are treated as fully binding with no availability payments or permitted non delivery, and costs are evaluated under a snapshot of Dutch tariffs. For tractability the optimisation omits device details such as generator start costs. Finally, the distribution system operator's willingness to pay is not observed, therefore outcomes are presented as trade offs rather than a single price point.

Future work should widen the scope and add realism in a measured way. A natural first step is to extend the rolling horizon analysis to several weeks across seasons, include intraday forecast updates, and explore stochastic or chance constrained variants that target a chosen compliance probability. Also an interesting focus will be in expanding the asset and tariff model by adding small, safe process flexibilities where they exist, modelling generator start up costs, and repeating the experiments under new Dutch tariff options. Finally, moving from one site to small portfolios on a shared feeder, either with a common capacity constraint or a light feeder sensitivity model, would link delivered reduction more directly to local relief.

Finally, running short pilots and test calls at other sites, reusing this method would be crucial to further validate it on new data, and make it a practical screening tool for programme design and onboarding.

8.2 Closing statement

Overall, this thesis demonstrates that *congestion management does not need to be complex nor expensive*: with transparent, contract-based limits, (modest) storage, and settlement-aligned operation, a process driven facility can provide reliable congestion relief at prices that suit distribution level needs, without putting at risk its core mission. The framework and findings are applicable to other MV industrial types and provide a concrete bridge between policy intent, DSO operations, and plant-level decision-making.

Appendix A

Appendix A

A.1 Contract-consistent hard-cap CBC

The appendix introduces a contract-consistent hard-cap CBC model in which the depth of the import cap during congested hours is chosen endogenously by the optimiser, driven by an explicit capacity remuneration price. In the main thesis we price outcomes ex post (we first schedule to a fixed cap and then compute the bill/compensation). Here, the price signal is brought inside the optimisation: for each congested hour, the model selects an active cap, k, and earns a payment π per kW of reduction relative to the nominal limit, but only for the part above a free threshold θ. Outside congestion, the cap is fixed at the contracted level.

The same congestion hourly indicator as above is used, $z_h \in \{0, 1\}$, inferred from Table 5.5, $z_h = 1$ if hour h is congested, and $z_h = 0$ otherwise. $E^{\max} = \kappa_{\text{imp}} \Delta t$ denotes the nominal per-step energy bound corresponding to the contracted off-take κ_{imp} (kW) and the step size $\Delta t = 0.25$ h. During congested periods ($z_h = 1$), the model introduces an hourly decision variable κ_h (kWh/step) that defines the active import cap for hour h and applies it to every step t in that hour:

$$\texttt{grid\_import}[t] \;\leq\; \kappa_{h(t)}, \qquad \forall\, t \in \mathcal{T}. \tag{A.1}$$

Outside congestion, the cap is locked at the nominal level:

$$\kappa_h \;\geq\; E^{\max} \quad \text{if} \quad z_h = 0. \tag{A.2}$$

Given the variable bounds $0 \leq \kappa_h \leq E^{\max}$, (A.2) pins κ_h to $E^{\max}$ whenever $z_h = 0$.

Additionally, we define the hourly reduction in per-step energy (relative to the nominal cap) as

$$d_h \;=\; E^{\max} - \kappa_h \quad (\text{kWh/step}), \tag{A.3}$$

and θ (kWh/step) denotes a threshold below which reductions are not remunerated (derived here from the threshold found in the *Yearly fixed CBC* experiment (5.5.3.1). Only the portion of the reduction over θ is credited during congested hours. We introduce a nonnegative auxiliary r_h (kWh/step) for the *rewarded* reduction and enforce

$$r_h \;=\; \max\{\, 0,\; d_h - \theta \,\}\; z_h, \tag{A.4}$$

Then we add the capacity remuneration rate, π (€/kW). Since r_h is in kWh per 15-minute step, its kW equivalent is $4\, r_h$. The model credits the following revenue across congested hours:

$$R^{\text{cap}} \;=\; \sum_{h \in \mathcal{H}} z_h \left(4\, r_h\right) \pi. \tag{A.5}$$

The total objective remains that of the core model, with R^{cap} subtracted so that economically optimal cap depths during congestion are chosen internally:

$$\min\left(C^{\text{imp}} - R^{\text{exp}}\right) + \left(C^{\text{vol}} + C^{\text{peak}}\right) + C^{\text{pen}} \;-\; R^{\text{cap}}. \tag{A.6}$$

This formulation is useful to explore dynamic caps or price-based procurement: by sweeping π and θ one can map how much reduction the site offers at each hour and at what cost, including rebound effects and intertemporal constraints. It also aligns naturally with day-ahead CBCs where the DSO publishes a uniform price for tomorrow's congested hours.

In our case, with highly variable demand and limited foresight, the added endogeneity offers less additional insight than under smoother profiles: the optimal k, can swing with short-term load volatility rather than reflect a stable "economic cap". Still, the construct is valuable as a design lab, extendable to rolling-horizon runs, richer price indices, or multi-hour price blocks, and provides a direct bridge to studying price-indexed, dynamic limitation contracts alongside the fixed-cap CBCs analysed in the main text.

A.2 Expert Interviews

A.2.1 TenneT policy advisor on flexibility tools and congestion

Interviewee: Policy Advisor, Electricity Market Design, TenneT
Format: Curated transcript excerpts (with permission)
Purpose: Context on industrial flexibility tooling, congestion modelling, and compensation under capacity-limiting arrangements.
Note: Minor wording has been cleaned for readability; only the most relevant passages are reproduced.

Why TenneT supported tooling, and what the TSO does (and does not) do

> "We are the system operator. We provide market access and connections, but we don't 'need' a battery. Revenues must come from the market (balancing, congestion services), not from TenneT investing in assets."

Storage tool & ValueFlex (industrial flexibility) — scope and ownership

> "The storage tool showed developers there's an electricity market, *you* must test feasibility. The industrial tool (*ValueFlex*) was developed **with and owned by Guidehouse**; serious users needed their support. It wasn't a simple self-serve interface."

Target audience and adoption barriers

> "It focused on smaller industries. Many energy managers have 1–3 year horizons and many other duties; payback expectations can be a year. That short-term focus makes the group hard to reach."

Why the project stopped; current status

> "We ended the project, it's **not our job to make commercial business cases**. Flexibility is now widely recognised as necessary. *Guidehouse still has the tool archived.*"

System view: electrification and flexibility

> "If we want a renewable system, most energy use should be electricity. That requires industrial transformation and *buffers* so demand can move to low-cost power. Flexibility isn't free, it comes with costs."

On modelling congestion (relevance for MV grids)

> "To say something meaningful about *congestion* you need load-flow modelling. Markets can be approximated as copper-plate, but once you include grid constraints, especially at MV/LV, it's much more complex."

CBCs: compensation versus 'revenue'

> "You don't get 'revenue' from a CBC; you get *compensation* because you're asked to do something you weren't planning to do. It covers the inconvenience and cost of limiting or shifting consumption."

TSO tariff note (HV context)

> "From 1 January, consumers on TenneT's grid that paid a kW-max component now pay a *weighted* kW-max based on a colour scheme."

A.2.2 DSO system operations expert on congestion management

Interviewee: System operations expert at a Dutch DSO (Alliander)
Format: Curated transcript excerpts (with permission)
Purpose: Practitioner perspective on CBC design and use, demand vs. in-feed application, operational relevance, incentives, and modelling caveats.
Note: Minor wording has been cleaned for readability; only the most relevant passages are reproduced.

CBCs as a financial choice for industry (compensation vs. revenue)

> "For a normal company it becomes a financial discussion: if I limit power on certain assets, I produce fewer units and lose sales. A CBC should *compensate* that. It's not 'free money'; it's compensation for doing something you weren't planning to do."

Two CBC "flavours" and where they fit

> "We use two flavours: fixed CBC where limits are known long in advance, and call-based CBC. Demand users often prefer fixed because they can plan around it. We use many more call-based on the *in-feed* side (solar, wind): power is core revenue there, so curtailment is operationally simpler than stopping a factory."

When a single site's peak matters: coincidence with substation peaks

> "It's not just the company's own peak, it's whether that peak coincides with the station's peak. Midday load in summer may be fine; evening winter peaks are more problematic. The value is in reducing demand **when the feeder is actually stressed.**"

Lowering contracted capacity vs. creating 'new space'

> "Reducing a contracted level doesn't automatically create capacity. DSOs oversubscribe because not everyone uses their max at once. 'Use it or lose it' policies help, and knowing a firm upper bound can aid planning, but it doesn't magically free capacity on day one."

What DSOs will (and won't) pay for

> "This is communal money. We're not going to spend millions to shave a small peak at one site, that's the cost of upgrading a whole station. Fixed CBC is relatively cheap to hold; call-based CBC pricing is case-by-case, often linked to market levels (e.g., DA averages plus a premium per MW)."

Scale today: need vs. offered flexibility

> "When we ask for flexibility, the *need* is in the gigawatts; the offered volume is in the tens of megawatts, well under one percent."

Product effectiveness and time horizon

> "Redispatch (e.g. GOPACS) isn't very successful for us. CBCs are effective, they change behaviour, but they're not cheap. We'll need flex products for the next 10–20 years, but long-term grid reinforcement is ultimately the cheaper, structural fix."

Why companies participate

> "Money, *and* connection progress. New connections may accept CBCs to get on the grid sooner. Existing large consumers have less incentive if power is a small share of product cost."

Data transparency and locality

> "Effective congestion management needs local coupling to the grid. Detailed usage data are commercially sensitive, and sharing highly granular location/limit data raises security concerns. That's why full transparency is hard in practice."

Bibliography

[1] United Nations Framework Convention on Climate Change, "The Paris Agreement," *UNFCCC Official Documents*, pp. 1–3, Dec. 2015. [Online]. Available: https://unfccc.int/sites/default/files/english_paris_agreement.pdf.

[2] European Commission. "The European Green Deal. "[Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52019DC0640.

[3] Government of the Netherlands. "Dutch Climate Act (Klimaatwet)." Accessed: 2025-04-16. [Online]. Available: https://wetten.overheid.nl/BWBR0042394/2020-01-01.

[4] Netherlands Enterprise Agency (RVO). "National Climate Agreement (Klimaatakkoord)." Accessed: 2025-04-16. [Online]. Available: https://english.rvo.nl/sites/default/files/2020/07/National%20Climate%20Agreement%20The%20Netherlands%20-%20English.pdf.

[5] Alliander N.V. "Alliander Results 2024." Accessed: 2025-07-16. [Online]. Available: https://www.alliander.com/content/uploads/dotcom/Alliander-Results-2024-EN-versie.pdf.

[6] Zsuzsanna Pató. "Gridlock in the Netherlands," Regulatory Assistance Project (RAP).

[7] PwC Strategy&. "Unlocking Industrial Demand Side Response." Accessed: 2025-05-6, TenneT, PwC.

[8] Bo de Wildt and Ross Quirke and Jos Sijm. "System Integration Outlook 2022: Towards Robust and Efficient Energy Systems." Accessed: 2025-04-03, TNO. [Online]. Available: https://publications.tno.nl/publication/34639482/Gliqt5/TNO-2022-P10368.pdf.

[9] J. Sijm, G. Morales-España, and R. Hernández-Serna, "The Role of Demand Response in the Power System of the Netherlands, 2030–2050," TNO, Tech. Rep. TNO 2022 P10131, May 2022.

[10] Wood Mackenzie. "The Netherlands' gridlock: a cautionary tale for the US." Accessed: 2025-04-11. [Online]. Available: https://www.woodmac.com/news/opinion/netherlands-gridlock/.

[11] T. Venema, T. van Swieten, S. van den Boogaard, R. Bieze, and M. Middelbos. "Solving the gridlock: Six interventions to accelerate congestion relief on the Dutch electricity grid," Boston Consulting Group.

[12] R. J. Hennig, L. J. de Vries, and S. H. Tindemans, "Risk vs. restriction—an investigation of capacity-limitation based congestion management in electric distribution grids," *Energy Policy*, vol. 186, p. 113 976, 2024, Open access under CC BY license. DOI: 10.1016/j.enpol.2023.113976. [Online]. Available: https://doi.org/10.1016/j.enpol.2023.113976.

[13] EcoMENA. "Wastewater treatment," Accessed: Aug. 16, 2025. [Online]. Available: https://www.ecomena.org/wastewater-treatment/.

[14] Liander N.V. "Tarieven voor aansluiting en transport elektriciteit 2023." Accessed: 21 July 2025. [Online]. Available: https://www.liander.nl/-/media/files/tarieven/grootzakelijk/tarieven-2023/tarieven-voor-aansluiting-en-transport-elektriciteit-grootverbruik-per-112023v100.pdf.

[15] *Ieee std 519: Recommended practice and requirements for harmonic control in electric power systems*, Official IEEE standard page, Piscataway, NJ: IEEE Standards Association, 2014. Accessed: Sep. 16, 2025. [Online]. Available: https://standards.ieee.org/ieee/519/3710/.

[16] International Energy Agency, "The Netherlands 2024: Energy Policy Review," International Energy Agency (IEA), Tech. Rep., 2024, Accessed: 2025-05-21. [Online]. Available: https://iea.blob.core.windows.net/assets/2b729152-456e-43ed-bd9b-ecff5ed86c13/TheNetherlands2024.pdf.

[17] International Energy Agency. "The netherlands: Electricity." Content used under the Creative Commons CC BY 4.0 licence, Accessed: Sep. 17, 2025. [Online]. Available: https://www.iea.org/countries/the-netherlands/electricity.

[18] H. Ritchie, M. Roser, and E. Mathieu. "Co_2 and greenhouse gas emissions." Data source: Global Carbon Budget (2024). Licensed under CC BY 4.0, Accessed: Sep. 17, 2025. [Online]. Available: https://ourworldindata.org/co2-emissions.

[19] European Commission. "2040 Climate Target." Accessed: 2025-05-21. [Online]. Available: https://climate.ec.europa.eu/eu-action/climate-strategies-targets/2040-climate-target_en#documents.

[20] J. Rosenow, S. Oxenaar, and E. Pusceddu, "Some Like It Hot: Moving Industrial Electrification from Potential to Practice," Regulatory Assistance Project (RAP), Tech. Rep., Dec. 2024.

[21] Strategy& and PwC, "Financiële Impact Energietransitie voor Netbeheerders (FIEN+)," Netbeheer Nederland, Tech. Rep., Mar. 2025, Eindrapport, in opdracht van Netbeheer Nederland. [Online]. Available: https://www.netbeheernederland.nl/sites/default/files/2024-12/241216_fien_eindrapport.pdf.

[22] National Institute for Public Health and the Environment (RIVM). "Nitrogen," RIVM, Accessed: Sep. 23, 2025. [Online]. Available: https://www.rivm.nl/en/nitrogen.

[23] DutchNews.nl. "Hundreds of building projects at risk amid nitrogen deadlock," DutchNews.nl, Accessed: Sep. 23, 2025. [Online]. Available: https://www.dutchnews.nl/2025/04/hundreds-of-building-projects-at-risk-amid-nitrogen-deadlock/.

[24] L. van Dijk, A. J. Swart, and P. van Ees, "Netcongestie: voor wat hoort wat," ABN AMRO, Tech. Rep., Jan. 2025. [Online]. Available: https://www.ttm.nl/wp-content/uploads/2025/01/rapport-netcongestie-voor-wat-hoort-wat_tcm16-251949.pdf.

[25] R. A. Rodriguez, S. Becker, G. B. Andresen, D. Heide, and M. Greiner, "Transmission needs across a fully renewable european power system," *Renewable Energy*, vol. 63, pp. 467–476, Mar. 2014. DOI: 10.1016/j.renene.2013.10.005. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0960148113005351.

[26] J. Jurasz, F. A. Canales, A. Kies, M. Guezgouz, and A. Beluco, "A review on the complementarity of renewable energy sources: Concept, metrics, application and future research directions," *Solar Energy*, vol. 195, pp. 703–724, Jan. 2020. DOI: 10.1016/j.solener.2019.11.087.

[27] "Landelijk Actieprogramma Netcongestie (LAN)." Accessed: 21 July 2025. [Online]. Available: https://actieprogrammanetcongestie.nl/.

[28] Netbeheer Nederland. "Capaciteitskaart." Accessed: 21 July 2025. [Online]. Available: https://www.netbeheernederland.nl/netcapaciteit-en-flexibiliteit/capaciteitskaart.

[29] S. Schulte and M. Altaghlibi, "Dutch grid delays cost up to EUR 376 million every year," ABN AMRO ESG Economist, Group Economics & Financial Markets & Sustainability Research, Tech. Rep., Sep. 2024, Accessed: 21 July 2025. [Online]. Available: https://assets.ctfassets.net/1u811bvgvthc/1mhRzNXmyiHVoS7OKOsNeE/77c73a9d7c3edcf6c38b3dfda6ed224e/ESG_Economist_Costs_of_Grid_delay_ENG.pdf.

[30] E. Rooijers. "Wachtlijst stroomnet verdubbeld: ruim 20.000 bedrijven wachten op aansluiting. "[Online]. Available: https://fd.nl/bedrijfsleven/1549501/wachtlijst-stroomnet-verdubbeld-ruim-20-000-bedrijven-wachten-op-aansluiting.

[31] Ecorys, "Maatschappelijke kostprijs van netcongestie," Ministerie van Economische Zaken en Klimaat & Rijksdienst voor Ondernemend Nederland (RVO), Eindrapport, Apr. 2024, Accessed: 21 July 2025. [Online]. Available: https://www.ecorys.com/app/uploads/2019/02/Maatschappelijke-kostprijs-van-netcongestie-2024.pdf.

[32] Aedes. "Brief Aedes tbv CD Klimaat en Energie, 14 maart 2024. "[Online]. Available: https://aedes.nl/media/document/brief-aedes-tbv-cd-klimaat-en-energie-14-maart-2024.

[33] Team IO+. "New research: In these Dutch towns, all residents suffer from grid congestion." Accessed: 21 July 2025. [Online]. Available: https://ioplus.nl/en/posts/new-research-in-these-dutch-towns-all-residents-suffer-from-grid-congestion.

[34] Netherlands Authority for Consumers and Markets (ACM). "About ACM: Our Organization. "[Online]. Available: https://www.acm.nl/en/about-acm/our-organization/the-netherlands-authority-for-consumers-and-markets.

[35] "Elektriciteitswet 1998. "[Online]. Available: https://wetten.overheid.nl/BWBR0009755/.

[36] Netherlands Authority for Consumers and Markets (ACM). "Voorstel codewijziging alternatieve transportrechten." Geraadpleegd op 22 juli 2025. [Online]. Available: https://www.acm.nl/nl/publicaties/voorstel-codewijziging-alternatieve-transportrechten#:~:text=Netbeheer%20Nederland%20heeft%20een%20voorstel,transportrecht%20en%20het%20tijdsblokgebonden%20transportrecht..

[37] Netbeheer Nederland. "FAQ alternatieve transportrechten," Accessed: Sep. 23, 2025. [Online]. Available: https://www.netbeheernederland.nl/sites/default/files/2024-03/qa_atr_2201.pdf.

[38] Rijksdienst voor Ondernemend Nederland (RVO). "Pas uw stroomgebruik aan en krijg een vergoeding," Accessed: Sep. 23, 2025. [Online]. Available: https://www.rvo.nl/onderwerpen/netcongestie/capaciteitsbeperkend-contract.

[39] GOPACS B.V. "Redispatch Free Bidding (intraday). "[Online]. Available: https://www.gopacs.eu/en/products/redispatch-free-bidding-intraday/.

[40] GOPACS B.V. "Mandatory Bidding Contract Redispatch (intraday). "[Online]. Available: https://www.gopacs.eu/en/products/mandatory-bidding-contract-redispatch-intraday/.

[41] J. van Schelven, "The Right Tool for the Job: Development of a Congestion Management Strategy for the Dutch Transmission System," Eindhoven University of Technology, Tech. Rep., Jun. 2025.

[42] J. G. Dedecca, M. Ansarin, C. Bene, T. van Delzen, L. van Nuffel, and H. Jagtenberg, "Increasing Flexibility in the EU Energy System," European Parliamentary Research Service, Tech. Rep. EPRS_STU(2025)769347, 2025. [Online]. Available: https://www.europarl.europa.eu/RegData/etudes/STUD/2025/769347/ECTI_STU%282025%29769347_EN.pdf.

[43] Netbeheer Nederland. "ACM geeft groen licht voor cable pooling. "[Online]. Available: https://www.netbeheernederland.nl/artikelen/nieuws/acm-geeft-groen-licht-voor-cable-pooling.

[44] Rijksdienst voor Ondernemend Nederland (RVO). "Directe Lijn. "[Online]. Available: https://www.rvo.nl/onderwerpen/netcongestie-en-opslag-van-energie/directe-lijn.

[45] Enexis N.V. "Gesloten distributiesysteem. "[Online]. Available: https://www.enexis.nl/zakelijk/mijn-stroomnet/wat-kunt-u-samen-met-anderen-doen/gesloten-distributiesysteem.

[46] Netbeheer Nederland. "Groepstransportovereenkomst maakt onderling delen van transportcapaciteit mogelijk. "[Online]. Available: https://www.netbeheernederland.nl/artikelen/nieuws/groepstransportovereenkomst-maakt-onderling-delen-van-transportcapaciteit-mogelijk.

[47] Autoriteit Consument & Markt (ACM). "Codebesluit faciliteren meerdere leveranciers op een aansluiting. "[Online]. Available: https://www.officielebekendmakingen.nl/stcrt-2017-39821.html.

[48] Autoriteit Consument & Markt (ACM), "Overzicht en inzicht congestiemaatregelen ACM (geüpdatet juli 2025)," Autoriteit Consument & Markt, ACM/INT/526935, Jul. 2025. Accessed: Sep. 23, 2025. [Online]. Available: https://www.acm.nl/system/files/documents/overzicht-congestiemaatregelen-acm-juli-2025.pdf.

[49] TenneT. "Acm publishes tennet's tariff proposal 2025," TenneT, Accessed: Sep. 23, 2025. [Online]. Available: https://www.tennet.eu/nl-en/news/acm-publishes-tennets-tariff-proposal-2025.

[50] Autoriteit Consument & Markt (ACM). "Codebesluit prioriteringsruimte bij transportverzoeken. "[Online]. Available: https://www.acm.nl/nl/publicaties/codebesluit-prioriteringsruimte-bij-transportverzoeken.

[51] Enexis Netbeheer, *Capaciteitsbeperkingscontract*, Accessed: 2025-09-06, 2025. [Online]. Available: https://www.enexis.nl/zakelijk/netcapaciteit/flexibiliteitsoplossingen/capaciteitsbeperkingscontract.

[52] Autoriteit Consument & Markt (ACM), "Vragen en Antwoorden – Congestiemanagement en het Capaciteitsbeperkingscontract," Autoriteit Consument & Markt, ACM/INT/457084, 2022, Zaaknr. ACM/22/179576; Openbaar.

[53] R. J. Hennig and coauthors, "Congestion management in electricity distribution networks: Smart tariffs, local markets and direct control," *Utilities Policy*, vol. 85, p. 101 660, 2023.

[54] R. J. Hennig, M. Jonker, S. Tindemans, and L. J. de Vries, *Capacity subscription tariffs for electricity distribution networks: Design choices and congestion management (pdf copy)*, PDF uploaded, File: Capacity Subscription Tariffs for Electricity Distribution Networks Design Choices and Congestion management.pdf, 2020.

[55] M. Müller, J. Rodler, and W. F. Jooß, "Avoiding low-voltage grid overloads via curative dso intervention (peak smoothing)," in *6th E-Mobility Power System Integration Symposium*, 2022.

[56] M. Thomas, T. Alexander, and M. M. D., *Designing decentralized markets for distribution system flexibility (pdf copy)*, PDF uploaded, 2019.

[57] C. E. Group, "Connectedsolutions: An assessment for massachusetts," Clean Energy Group, Tech. Rep., 2021, Bring-Your-Own-Device DR/battery program assessment.

[58] J. C. Fransoo and W. G. M. M. Rutten, "A typology of production control situations in process industries," *International Journal of Operations & Production Management*, vol. 14, no. 12, pp. 47–57, 1994. DOI: 10.1108/01443579410072382. Accessed: Sep. 24, 2025. [Online]. Available: https://pure.tue.nl/ws/portalfiles/portal/1485805/413156.pdf.

[59] U.S. Environmental Protection Agency, "Energy efficiency in water and wastewater facilities: A guide to developing and implementing greenhouse gas reduction programs," U.S. Environmental Protection Agency, Washington, DC, EPA-430-R-09-038, 2013. Accessed: Sep. 24, 2025. [Online]. Available: https://www.epa.gov/sites/default/files/2015-08/documents/wastewater-guide.pdf.

[60] R. Hennig, M. Jonker, S. H. Tindemans, and L. de Vries, "Capacity subscription tariffs for electricity distribution networks: Design choices and congestion management," in *17th International Conference on the European Energy Market (EEM)*, 2020. DOI: 10.1109/EEM49802.2020.9221994.

[61] T. Dronne, F. Roques, and M. Saguan, "Local flexibility markets for distribution network congestionmanagement in center-western europe: Which design for which needs?" *Energies*, vol. 14, no. 14, p. 4113, 2021. DOI: 10.3390/en14144113.

[62] R. Yasmin, B. M. R. Amin, R. Shah, and A. Barton, "A survey of commercial and industrial demand response flexibility with energy storage systems and renewable energy," *Sustainability*, vol. 16, no. 2, p. 731, 2024. DOI: 10.3390/su16020731. [Online]. Available: https://doi.org/10.3390/su16020731.

[63] D. B. Nguyen, J. M. A. Scherpen, and F. Bliek, *Distributed optimal control of smart electricity grids with congestion management*, PDF uploaded, 2017.

[64] T. Omrani and coauthors, "Dynamic congestion management with chance-constrained mpc in networked microgrids under consumers-related uncertainties," *IEEE Transactions on Consumer Electronics*, 2024. DOI: 10.1109/TCE.2024.3476200.

[65] T. S. Jingran Maa S. Joe Qina, "Application of economic mpc to the energy and demand minimization of a commercial building," *TODO*, 2014.

[66] K. Chakravarthi, P. Bhui, N. K. Sharma, and B. C. Pal, "Real-time congestion management using generation re-dispatch: Modeling and controller design," *IEEE Transactions on Power Systems*, vol. 38, no. 3, pp. 2189–2203, 2023.

[67] P. Pfeifer, J. Tran, S. Krahl, and A. Moser, *Modelling of uncertainty for smart-grid congestion management*, Conference paper / report, 2020.

[68] D.-T. Hoang, S. Olaru, A. Iovine, J. Maeght, P. Panciatici, and M. Ruiz, *Predictive control for zonal congestion management of a transmission network*, PDF uploaded, 2021.

[69] Autoriteit Consument en Markt (ACM), *Netcode elektriciteit*, Accessed: 2025-08-05, 2025. [Online]. Available: https://wetten.overheid.nl/BWBR0037940/2025-09-01.

[70] Belastingdienst, *Energiebelasting*, Accessed: 2025-07-15, 2025. [Online]. Available: https://www.belastingdienst.nl/wps/wcm/connect/bldcontentnl/belastingdienst/zakelijk/overige_belastingen/belastingen_op_milieugrondslag/energiebelasting/.

[71] Rijksoverheid. "Overheid start nieuwe campagne: Oproep om tussen 16.00 en 21.00 minder stroom te gebruiken." Nieuwsbericht, Accessed: Jul. 16, 2025. [Online]. Available: https://www.rijksoverheid.nl/actueel/nieuws/2025/04/28/overheid-start-nieuwe-campagne-oproep-om-tussen-16.00-en-21.00-minder-stroom-te-gebruiken.

[72] I. Nieuwenhuijse et al., "Energy sources and demand: Can eu27+3 be self-sufficient in its demand for co_2-free energy sources and competing feedstocks in 2040?" CE Delft, Delft, Tech. Rep. 24.220479.079, Jun. 2024.