UNIVERSITY OF AMSTERDAM

MASTERS THESIS

---

# Investigation, simulation and real options valuation of the GOPACS flexibility market

---

*Author:*
Charel FELTEN

*Supervisor:*
Nikita GRANGER

*Examiner:*
Prof. dr. Drona KANDHAI

*Second assessor:*
Jori HOENCAMP

*A thesis submitted in partial fulfilment of the requirements for the degree of Master of Science in Computational Science*

*in the*

Computational Science Lab
Informatics Institute

July 7, 2022

UNIVERSITY OF AMSTERDAM

# *Abstract*

Faculty of Science

Informatics Institute

Master of Science in Computational Science

**Investigation, simulation and real options valuation of the GOPACS flexibility market**

by Charel FELTEN

GOPACS is a Dutch local flexibility market for electricity where sealed-bid pay-as-bid auctions can be launched by distribution system operators when they face congestion in their grid that they are unable to solve themselves. We investigate all relevant aspects related to GOPACS in detail, focusing on an in-depth exploration and presentation of the available data. We find that there is some correlation to related data which could be explored further. Having established a sufficient understanding of the market, we choose a stochastic approach to simulate it, based on a statistical analysis of the features of interest. We develop a stochastic finite-difference model, based on a Markov chain, that combines the arrival and departure times with the volumes and prices of electricity achieved in the auctions. We present two methods of defining the prices, one based on estimated distributions, the other consisting of simulating the bid ladder. Using this model representing the behaviour of our underlying, we continue with the valuation of a flexible asset participating in this market, using a real options strategy inspired by related literature. We approximate participation in GOPACS using digital options with a fixed payoff, justified by GOPACS being an auction-based market in which the payoff of participation is fixed and depends on the bids submitted into the process. Akin to related works we define the value of the flexible asset participating in GOPACS using a lifetime strip of options, one for each hour of GOPACS. We derive an analytical solution of this asset, applicable in some specific conditions, and simulate it with Monte Carlo otherwise. We investigate the behaviour of the prices of the market with respect to the choices possible in the bidding process and external factors, using scatter-based sensitivity analysis, and focusing on the average return based on the mean asset price and the risks which we measure with VaR and ES. We perform a case study of prices achievable with such an asset in a Dutch province and find values of around 60,000 €/MW for 1 year of participation with a 5 MW asset and lower prices per MW if the asset is larger or additional costs of participation apply. We end the work with an extensive investigation into the model risks introduced by our assumptions in the model development, as well as other risks and uncertainties faced when participating in GOPACS. A thorough validation of the achieved results is unfeasible due to the insufficient quantity and quality of the available GOPACS data.

# *Acknowledgements*

I would like to thank my supervisor, Nikita Granger, for their guidance and support throughout my thesis, in particular for introducing me to this area of research and the work at Shell. I would also like to thank my examiner, Prof. Drona Kandhai, for their feedback and expertise on the various topics of this thesis, and my second assessor, Jori Hoencamp for assessing my work. I also want to acknowledge the opportunities and support received at Shell throughout the research internship that produced this thesis, in particular by my manager Vasileios Giagkoulas. Finally, I would also like to thank my family and friends for the support received during my studies and thesis, in particular Gilles for help in the data collection process, and Alex and James for the teamwork in university classes and the support during the thesis work.

# Contents

# List of Abbreviations

| | |
|---|---|
| ABM | Agent-Based Model |
| ACC | Accuracy |
| AUC | Area Under the Curve |
| APX | Amsterdam Power Exchange |
| BPL | Bid Price Ladder |
| BRP | Balance Responsable Party |
| CB | Cleared Buckets |
| CDF | Cumulative Distribution Function |
| CM | Congestion Management (problem) |
| CP | Congestion management Portal |
| DA | Day-Ahead (market) |
| DCF | Discounted Cash Flow |
| DER | Distributed Energy Ressources |
| DR | Demand-Response |
| DRA | Demand-Responsive Asset |
| DSM | Demand-Side Management |
| DSO | Distribution System Operator |
| ENTSOE | European Network of Transmission System Operators for Electricity |
| EPEX | European Power EXchange |
| ETPA | Energy Trading Platform Amsterdam |
| EX | EXpenses |
| FA | Flexible Asset |
| FCR | Frequency Containment Reserves |
| aFRR | automatic Frequency Restoration Reserve |
| mFRR | manual Frequency Restoration Reserve |
| FSP | Flexibility Service Providers |
| FT | Free Text (problem) |
| GBM | Geometric Brownian Motion |
| GOPACS | Grid Operators PlAtform for Congestion Solutions |
| HP | HandelsPlatform |
| ID | Intra-Day (market) |
| IDCONS | Intra-Day CONgestion Spreads |
| ITM | In The Money |
| LFM | Local Flexibility Market |
| LSMC | Least Squares Monte Carlo |
| MA | Market Announcements |
| MAE | Mean Absolute Error |
| MC | Monte Carlo |
| MSE | Mean Squared Error |
| MW | Mega Watt |
| MWh | Mega Watt hour |
| NPV | Net Present Value |

| | |
|---|---|
| OTM | Out of The Money |
| PDF | Probability Density Function |
| PM | Performance Metrics |
| PP1 | Price Process 1 |
| PP2 | Price Process 2 |
| PTU | Program Time Unit |
| RM | Regulating Margin |
| SDAC | Single Day-Ahead Coupling |
| SDE | Stochastic Differential Equation |
| SIDC | Single Intra-Day Coupling |
| TP | TransPort (problem) |
| TSO | Transmission System Operator |

# Chapter 1

# Introduction

## 1.1 Motivation

In recent years, the energy landscape in Europe has been undergoing major changes on both the supply and demand side, impacting the stability of the electricity grid that is increasingly struggling to keep pace with these developments, and resulting in increased congestion. Various developments have emerged to address these problems, such as the grid operator platform for congestion solutions (GOPACS) in the Netherlands. GOPACS is a platform created with the intent of solving congestion in the electricity grid by connecting grid operators with flexibility providers [28]. GOPACS has only been introduced in 2018 and has since then established itself as the main flexibility market in the Netherlands, used by distribution systems operators, flexibility providers and electricity traders across the country. However as the congestion issues plaguing the electricity grid are still as prominent as ever, further growth of GOPACS is likely to continue.

GOPACS is an auction-based market that has no regular schedule and instead takes places whenever a distribution system operator (DSO) expects a congestion in the electricity grid that they are unable to solve themselves. They outsource solving this congestion problem by finding participants willing to sell power in one location and participants willing to buy power in another location. These participants need to be found at all costs, and as GOPACS takes place after the more liquid Day-Ahead (DA) electricity market, the prices achieved in these auctions can reach more extreme levels.

The continuing growth of GOPACS combined with the high prices achievable makes participation in GOPACS financially interesting. In particular for energy market participants that have some flexibility available, participation in GOPACS can result in immediate benefits with very limited downside, if any. Focusing on this optionality component – the right, but not the obligation to participate in GOPACS, we perform a real options (RO) valuation of a flexible asset that is participating in GOPACS.

## 1.2 Research Questions

We present three research questions that all cover different aspects of our work. A first question, or rather set of questions, is all about the various aspects of the GOPACS flexibility platform, specifically, we ask: *Who / What / When / Where / Why is GOPACS?* We want to answer these questions by doing extensive research on these aspects of GOPACS, in particular focusing on analysing and quantifying the available data.

Once we have established some understanding of the market, we will move on to the second part of our work, which is the development of a model to simulate GOPACS followed by the valuation of a flexible asset based on this developed model. Specifically, we want to answer: *What is the value of a flexible asset that is participating in GOPACS?* To determine this value, we focus on the optionality component of the participation, which we do with a RO approach. We will review some of the literature on the topic and then reformulate existing approaches to suit our problem. We also take a deeper look at how this flexible asset valuation is influenced by both external aspects of GOPACS and internal choices made by us, the participant. We develop a model to simulate GOPACS and formulate an option pricing problem to answer this research question.

However, the valuation of a flexible asset that is participating in GOPACS is not complete without taking a look at the risks associated with this problem. This leads us to our third research question: *What are the risks and uncertainties in participation in GOPACS and our valuation thereof?* To answer this question, we take a closer look at the different types of risks faced with our approach and GOPACS in general. While focusing on quantifying the distribution of outcomes generated by our model, we also examine the model risk itself introduced through various assumptions and simplifications as well as other risks and uncertainties.

## 1.3 Outline

We finish this introduction with an outline of the next chapters of this work. In the literature review in Chapter 2 on page 4 we first look into electricity and what distinguishes electricity from other commodities in an economic sense, focusing on the various electricity markets in the Netherlands. Next, we go over the topics of congestion and flexibility and how flexibility markets are used to address these issues. We investigate how flexibility is typically valued and then focus on one such method that we deem to be the most suitable for our situation, which are real options. We end the chapter with a small review of Monte Carlo methods relevant to our model.

Once we have reviewed the literature on relevant topics, we take a deeper look into GOPACS in Chapter 3 on page 17. We cover all relevant aspects of the GOPACS process by investigating who the participants are, why and when auctions are called, what happens during these auctions and when they take place, addressing our first research question. We also visualise all the data that is available to convey how GOPACS works in practice. We end the section with a discussion on the data and also a small dive into some statistical properties of the data that will be of use later on.

Next, we have the methods in Chapter 4 on page 42 in which we first present the model we develop to simulate GOPACS and outline the assumptions that this model relies on. We go over the various steps made in the development of the model and the individual subprocesses, connecting them to the statistical investigations made in the previous chapter. Once we have established the model for the underlying, we develop the RO pricing framework on top of it, which we use to value participation in GOPACS. We end the chapter with a presentation of some risk measures we use to quantify the risks associated to our approach.

After developing the model, we present the results achieved with this model in Chapter 5 on page 61. We first give a qualitative overview of the developed model and its features, followed by a validation of the model based on the stylised facts observed in the empirical data. After exploring the model, we turn our attention to the RO pricing approach. We perform some experiments to convey a feeling for these options and also investigate how we can influence the option value. We end the section, addressing our second research question by performing a valuation of a flexible asset in GOPACS and present the results.

We end this work with a discussion in Chapter 6 on page 79. We first readdress the various assumptions and simplifications made in our approach and discuss how they influence the model risk in order to answer our third research question. After quantifying and outlining this risk and other risks and uncertainties, we turn our attention to what could be done better and what interesting future work could build upon our work. We end the section and the thesis with a conclusion.

# Chapter 2

# Literature Review

## 2.1 Electricity

### 2.1.1 Electricity as a commodity

Energy can be defined as "the ability to do work" [26] and is expressed in joule (J). Various forms of energy exist, such as electric energy, nuclear energy, thermal energy and chemical energy [26]. "Work is done when energy is transferred from one form to another" [26]. The total amount of energy is conserved during this transfer, however, part of it may be transformed into a potentially undesirable form, such as e.g. heat. Power is "the rate at which work is done" i.e. "at which energy is transformed" [26]. Power is expressed as the amount of energy transformed per unit of time and is measured in watts (W), with 1 W = 1 J/s. Conversely, energy is defined as the product of power and time, hence the more common unit of megawatt-hour (MWh).

Electricity is generated when energy in some form is transformed into electrical energy. In practice, this conversion is done by a generator, such as e.g. a fossil fuel power plant converting thermal energy into electrical energy. The abilities of this generator can be quantified in two ways. The generator has a certain capacity, which is the maximum amount of power it can produce in theory [75]. Capacity is power and measured in MW. However we can also measure the actual amount of work produced during a certain period, this is the energy, and is measured in MWh.

In the past, once electricity was produced, it was injected into a high-voltage grid suitable for long-distance transportation [5]. This is the transmission grid, for which a transmission system operator (TSO) is responsible. In substations closer to the energy consumption locations, this high-voltage electricity was transformed into lower-voltage electricity more suitable for distribution of electricity to various destinations [5]. This is the distribution grid, for which a distribution system operator (DSO) is responsible. Nowadays, a lot of electricity generation takes place on the medium and low-voltage distribution grids as well.

From an economic perspective, electricity is a commodity that has two particular features that differentiate it from other commodities, namely (1) electricity can not be directly stored and (2) transporting electricity is constrained by specific physical laws [1]. These features result in various technical and practical characteristics which are of interest to our research.

While electricity can not be stored, it can however be converted into other forms of energy that can be stored. Among the most commonly deployed storage systems are pumped hydro storage, compressed air energy storage and various batteries such as Li-ion [76]. However, this conversion comes at a loss of energy[1], which is one of the key factors distinguishing electricity from other commodities. Among others, the non-storable nature of electricity leads to the emergence of various stylised facts in electricity prices, notably (1) daily, weekly and annual seasonality, (2) mean reversion, and (3) extreme price spikes [31, 1, 35], that differentiate it from other assets.

In addition, the non-storable nature of electricity also has consequences on the way that countries manage their grid. In order to have a properly working electricity network, a country needs to have enough capacity to satisfy peak demand. If the supply can not keep up with the demand, the frequency of the grid will decrease which can have drastic consequences on the stability of the network and lead to blackouts [1]. For this reason, the TSO, who is responsible for maintaining the balance of the electricity grid is required to have operating reserves at disposal [1]. "Operating reserves are generation capacities that can be mobilised within a given notification time" and that are organised according to their response time and manner (i.e. automatic or manual) [1].

The second feature of electricity is that its transport is constrained by physical laws. These laws, called Kirchhoff's rules, state that in a network, "the intensity at each node should be zero" and "the tension in each loop should also be zero" [1]). In practice, this implies that electricity does not simply flow from one node to the other, but instead spreads out evenly across all paths of the grid [3]. The consequence of this is that for any electricity transfer between market zones, the available capacity can not be fully allocated to said transfer only, as some of the capacity will be used by parallel flows from other zones [3]. The result is a complex issue called cross-border capacity allocation and requires careful planning of electricity generation across the grid. In the connected European grid, this issue is approached with Flow-Based Market Coupling (FBMC), used in both day-ahead markets and intraday markets [3, 12]. Of relevance to this thesis is only the consequence of this coupling, meaning that all power networks across Europe are more or less connected, and often issues in one part of this network can originate from complications in other parts of the network, across countries. Of note is also that this whole zone pays the same price for electricity if the electricity is able to spread evenly across this grid. However, in the case of grid congestion, prices may fluctuate across regions. We will go more in-depth into congestion later.

### 2.1.2 Electricity markets in the Netherlands

The Dutch electricity market, much like many other markets in Europe, can be divided into two parts: Electricity trading and power balancing / ancillary services [10]. A general overview of the various markets is shown in Figure 2.1 on the next page.

In Europe, electricity trading can be divided into three types of markets based on the time the delivery is expected: the Intraday (ID) market, the Day-ahead (DA) market and the more long-term futures and forward contracts market [10]. The electricity trading market is non-exclusive and multiple trading platforms and exchanges exist: The Amsterdam Power Exchange (APX), owned by the European Power Exchange (EPEX), handles day-ahead and intraday markets, ICE ENDEX handles futures and forwards [10]. Alternative platforms exist as well, such as the Energy Trading Platform Amsterdam (ETPA) platform which offers short-term intraday trading specifically aimed at optimising the energy capacity of participants, such as, among other, ex-post trading to reduce imbalance costs [23].

The DA market in the Netherlands is handled by EPEX via APX. The Netherlands participate in the European hourly day-ahead coupled auction, SDAC [12], together with various other

[1] technically, no energy is lost, but part of the energy is transformed into forms that are hard or impossible to use, such as heat, hence in practice we can consider it a loss

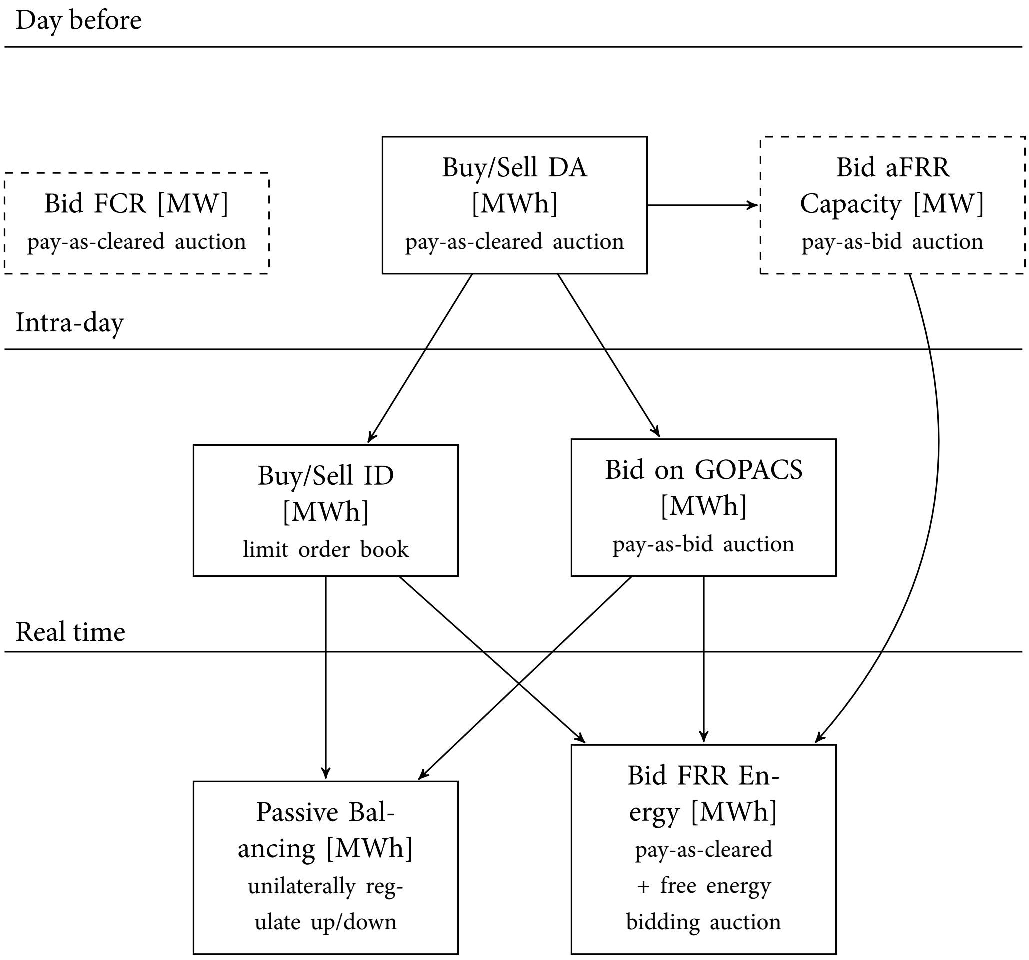



FIGURE 2.1: A flow chart of the various Dutch electricity markets. The document is inspired by a presentation made by Nikita Granger at Shell. Each market is represented by a rectangle, a dashed border means capacity market, a solid border means energy market.

EU members [21]. The day-ahead market is a blind auction, opening 45 days in advance and closing at 12:00 the day before the delivery [21]. The traded products are 24 1-hour contracts corresponding to the 24 hours of the following day, traded in blocks or single hours.

The Dutch ID market is operated continuously, 24/7 all year round, up until five minutes before delivery [21]. It is also part of the Europe-wide coupled market, SIDC [12]. Tradeable products come in greater granularity than on the DA market - in addition to the 1-hour contracts are 30-minute and 15-minute contracts [21]. Both DA and ID markets (as well as forward markets) are energy auctions - that is, the traded products are energy.

Power balancing is the task of the transmission system operator (TSO) and consists of adjusting, balancing, the supply and demand in the grid on short notice [8]. In order to satisfy this demand, the TSO needs operating reserves. However, the TSO typically do not have their own energy generating capacity [8], hence they need to buy it from market participants who do, referred to as balance responsible parties (BRP). These trades take place in the ancillary services market, which is a single-buyer market where the TSO procures operating reserve and capacity from BRPs [41]. The BRP offer their available capacity (e.g. generators) and are financially compensated for offered capacity and/or energy provided with said capacity, depending on the type of reserve [8].

Among the various products traded as ancillary services, there are three categories of operating reserves [41]: Primary, secondary and tertiary reserves. Primary reserves, also called frequency containment reserves (FCR) are the fastest to respond to a contingency event, they have to be fully operational within 30 seconds and are used to stabilise the system frequency [41]. FCR participation used to be mandatory and unpaid until 2014, nowadays it is voluntary and BRPs are paid for their capacity, but not for the energy[2] [10]. Next in line are secondary reserves which are deployed automatically and are also called frequency restoration reserves (aFRR). These reserves are key and participants are paid both for capacity and energy. First, participants can bid their capacity and get paid for making it available. Secondly, in the situation where secondary reserves are called for, market participants that sold their capacity provide energy and are paid for it, but also market participants that did not sell capacity can still provide energy and then also get paid, however only for the energy [10]. In addition, participants can also regulate the grid up or down without being called, and get paid for their service, however only if they regulate in the right direction. For example, if the grid frequency is going down, a customer that is regulating up will get paid afterwards, however, if they regulate down they will get fined. This is called passive balancing. Finally, tertiary reserves are deployed manually with the aim of further supporting the secondary reserves in stabilising the available operating reserve capacity in the concerned area. They are also frequency restoration reserves (mFRR) and replacement reserves (RR) [41]. More details about the current implementation of ancillary services are available on TenneTs website[3].

Among all the mentioned markets, the European DA market is the most liquid one, even more liquid than the European ID market [1, 13, 21]. The DA market volume has reached from 427 to 507 TWh yearly between 2014 and 2020, while the ID market has only reached from 47 to 111 TWh in the same span. In addition, the DA market is also considered the "reference price" across Europe [21]. The DA market is fundamental-driven, while other markets are deviations from the DA market, or markets for residuals, serving as means to "manage forecast errors or unforeseen events" [21, 8].

## 2.2 Congestion and Flexibility

Currently, the whole energy sector, particularly in Europe, is undergoing a major long-term evolution both on the supply side (more renewable energy, decarbonisation, decentralised and distributed energy resources (DER) and storage) and on the demand side (increasing load, increased electrification, technological advances allowing greater demand flexibility) which requires rethinking of the current electricity network design, management and organisation [44, 55, 72, 18, 64, 59]. The more distributed and variable nature of DER such as wind and solar is complicating traditional grid management used for the centralised, dispatchable and predictable fossil-fuel-based electricity generation [55]. In particular, the increase of DER in distribution instead of transmission networks leads to increasing grid congestion [64, 18, 55]. Traditional ways to deal with this congestion, such as reinforcing grid topology and reactive power management are slow, becoming increasingly expensive and reaching their limit with the fast-paced growth of DER [18, 55]. Other solutions to congestion are policies such as energy or capacity tariffs [67]. An evolution of market redesign is required [18]. One solution to the congestion problems caused by decreasing control and predictability on the supply side is by increasing the flexibility of both consumers and producers [67, 55, 34]. Flexibility can be

[2] to clarify, this means that BRPs are paid to have generating capacity available (e.g. a 10 MW battery at 50% charge. This battery is then automatically contracted to charge and discharge, which may happen frequently. However, the battery is unlikely to be fully depleted or charged, hence not a lot of energy is transferred. BRPs are paid to offer the capacity of this battery, but not for the energy that is being transmitted via the battery.

[3] https://www.tennet.eu/electricity-market/dutch-ancillary-services/

defined as “the ability to purposely deviate from a planned / normal generation or consumption pattern” [16]. Flexibility can be monetised and applied in various ways [16, 67]. More traditional ideas are policies that make the energy price depend on factors such as peak hours or peak load and reward users to decrease their needs during peak hours [67]. Combining these ideas and going further is the increasingly more popular concept of flexibility markets / platforms, enabling and coordinating flexibility trading [67, 16, 18, 65, 22, 55, 34, 72, 59].

A flexibility market or platform is a venue that allows users to directly trade and monetise their flexibility in terms of consumption or production [67]. Flexibility can be traded on various levels, such as balancing flexibility at the transmission grid, balancing flexibility at the distribution grid and flexibility for the distribution grid [73]. GOPACS is part of the latter and we will hence focus on flexibility markets for the distribution grid, which also tend to be referred to as local flexibility markets (LFM) in the literature [18, 55, 34, 59]. In particular, [59] define LFMs as “long-or short-term trading actions for flexibility in a specific geographical location, voltage level, and system operator (DSO and TSO), given by grid conditions or balancing needs, where participants in a relevant market can be aggregated to provide flexibility services”. Currently lots of flexibility market designs are being actively tested throughout Europe, often in pilot phases. Comparisons between those markets can be found in the literature, [72] analysed eighteen different designs, [6] analysed eleven market-based congestion management models, and [65] and [18] analysed four flexibility markets respectively. Among the most advanced initiatives according to [65] are Piclo Flex (UK) [57], GOPACS (NL) [28], Enera (DE) [20] and NODES (DE, NO) [54]. Apart from Enera (case study ended in 2020, and results are evaluated and soon to be published [20]), all three other markets seem to be still actively used for flexibility trading as of early 2022. Moreover, all of these projects are aimed at solving congestion on the DSO level, i.e. on the medium and low voltage grid instead of the high voltage transmission grid. Flexibility markets focused on the high voltage grid are still a frontier. An in-depth investigation into the characteristics of, and differences between those different LFMs is out of the scope of this work and we refer the reader to the cited literature.

One way to provide flexibility or participate in flexibility markets is by using flexible assets (FA), which provide flexibility from the demand side or supply side. On the supply side, Kondziella and Bruckner [38] mention, among other options, “Highly flexible power plants that could cope with increasing ramping requirements”. Typical flexible production assets (FA) are coal and natural gas power plants, in particular, there are three aspects that quantify their flexibility: (1) Their start-up and shut-down speeds and capabilities, (2) their output range while running and (3) the ramping speed it takes them to change between these output levels [44, 30]. There are also other types of FA such as Combined-Heat-and-Power Plants (CHP) [44]. On the supply side, flexibility is synonymous to demand response. Demand response is defined by Albadi and El-Saadany [2] as “changes in electricity usage by end-use customers from their normal consumption patterns in response to changes in the price of electricity over time”. They argue that there are three ways customers can apply demand response, namely (1) reducing their electricity usage during peak hours, (2) shifting their power-intensive operations from peak hours to off-peak hours or (3) producing their own energy during peak hours [2]. Sezgen et al. [66] also mention the first two options, referring to them as (1) load curtailment and (2) load displacement. The main benefit of demand response for customers is to save money by avoiding the usage and henceforth the purchase of energy during peak hours, i.e. during hours in which the demand is large and typically prices are higher than usual. There are two types of costs associated with demand response customers, namely initial costs like installing the technology, and running costs that incur during usage such as e.g. the lost business due to doing DR. Demand response is also sometimes referred to as demand-side management (DSM) in the literature.

Finally, one last aspect we care about is how flexibility is typically valued. As flexibility and FA encompass a wide range of potential applications, the valuation depends on the type of flexibility that is offered. Sezgen et al. [66] mention that demand-side DR is often valued with discounted cash flow methods, either by estimating benefits using long-term forecasts of future prices or by estimating historical hourly or daily prices, however, they argue that both are flawed and instead advocate for a real options approach. Hentschel et al. [30] mention a 5-step process for a supply-side flexible power plant valuation, consisting of (1) marginal cost calculation (2) intersection between price signal and marginal cost, (3) optimisation of the dispatching scheme, (4) determination of power plant ramping and start-up events and (5) a comparison of yearly cost and profit. In general, however, a growing body of research uses real options to value the optionality inherent to all FA, and it is this approach that we will follow as well.

## 2.3 Real Options

Derivatives are financial instruments whose value depends on the value of an underlying asset. Options are a kind of derivative which convey to their owner the right, but not the obligation to purchase or sell the underlying asset at a given time or in a given time period and for a fixed price. One example would be a call option, whose value $V_T$ (at maturity $T$) depends on the price $S_T$ of the underlying and a strike price $K$ as follows

$$V_T = \max(S_T - K, 0) = (S_T - K)^+ \tag{2.1}$$

While plenty of option types exist, they all share the optionality component: "The asymmetry deriving from having the right but not the obligation to exercise the option lies at the heart of the option's value" [70]. Options are found in various fields, the most common one being finance. Financial options[4] are standardised instruments that are traded on established and liquid markets and used for a variety of purposes ranging from hedging investments to leveraged speculation. However, another type of option of particular interest to this thesis are real options (RO) which are used to quantify and valuate investment and managerial decisions of companies. As opposed to financial options, RO are not standardised instruments, but instead, they represent financial situations / opportunities that depend on some decision to be made. They are called options, because this decision bears an element of optionality, and can be analysed and quantified with tools similar to the one used for financial options.

In his 1977 work, Myers [52] deconstructs a firm into two asset types, one being "real assets, which have market values independent of the firm's investment strategy" and the other being "real options, which are the opportunities to purchase real assets on possibly favourable terms", henceforth coining the term 'real options'. In particular, he gives the example of some investment or growth opportunity, which he regards "as a call option on a real asset" [52]. In fact, various investment opportunities have an inherent optionality factor. Trigeorgis [70] mentions the "option to defer", "option to alter operating scale", "option to abandon" and others as RO. He argues that " management's flexibility to adapt its future actions in response to altered future market conditions" provides "upside potential while limiting downside losses" which "traditional discounted-cash-flow (DCF) approaches such as the standard net-present-value (NPV) rule, cannot properly capture" [70].

Across the reviewed literature, the term 'real options' is also commonly associated with the valuation of an asset with an optionality component, whose value depends, at least partially, on

[4] Financial options are more commonly referred to simply as 'options', however, we will refer to them as 'financial options' to differentiate them from 'real options'

some underlying uncertainties, and borrowing methodology from the financial option world. Due to the nature of our work, we consider mainly approaches where at least one of the underlying uncertainties is electricity in some electricity market. Paradigmatic examples are the valuation of physical assets such as electricity generation or transmission assets [32, 17, 24, 71, 31, 13], but also the valuation of less tangible assets such as e.g. the value of demand-response, willingness to be subjected to load curtailment or the provision of reserve capacity [56, 66, 62, 49].

Gedra and Varaiya [25] propose "callable forwards" which they define as a "bundle of a forward and a short call" on a unit of energy. If this contract is entered on the long side by a consumer and on the short side by a utility, the utility now has the right, but not the obligation to provide energy to the consumer, and may choose to exercise this optionally in case prices are unfavourable. The work of Gedra and Varaiya constitutes, to our knowledge, thus one of, if not the first valuation of demand response (which they refer to as demand-side management) using a financial options approach.

Hsu [32] suggests the usage of spark spread call and put options "as a tool for valuing natural gas power plants". He connects the "ability of a power plant owner to elect whether or not to operate" with the characteristics of options in finance, and presents this decision as a financial option:

$$V = \max(P_e - hP_g, 0) \tag{2.2}$$

with $V$ the profit per MWh, $P_e$ and $P_g$ the spot prices of electricity and gas, and $h$ the heat rate representing the gas-to-electricity conversion efficiency. Given that both $P_e$ and $P_g$ and uncertainties, Hsu argues that Equation (2.2) is a spread call option. In particular, he arrives at the conclusion that "ownership of a natural gas power plant is equivalent to holding a series of spark spread call options with different expiration dates (e.g., monthly maturity) and with a strike heat rate equal to the unit's operating heat rate", implying that "gas power plants are simply "derivatives" of power and gas prices" and hence justifying the RO approach "we can look at the value of natural gas power plants through a financial framework" [32]. Continuing his exploration, he proposes option positions that can be used to hedge against unfavourable market conditions. He then moves on to the valuation of a natural gas plant, in particular, after making some assumptions on the behaviour of market prices, comparing the differences arising between a DCF approach and a RO approach. In his example he finds that "by ignoring the option piece the company undervalues the power asset by almost 20 percent" [32]. He concludes his research by investigating the sensitivities of the spark spread option with respect to price volatility and gas-to-electricity price correlation, and highlighting potential liquidity issues that could emerge in the, at that point, less established electricity markets.

Deng et al. [17] use RO to determine the value of electricity generation and transmission assets. They first derive the value of both spark spread and locational spread options of an underlying asset following either Geometric Brownian Motion (GBM) or a mean-reversion process, using futures-based replication. Using these valuation results, they then establish the equivalence between the spark spread option and the right, but not obligation, to operate a generation asset as well as the equivalence between the locational spread option and the right to operate a transmission asset. More concretely, they prove that the value of 1 unit of capacity of a generation asset $V_{gen}$ with lifetime $T$ is defined by the value of a spark spread call option value $C_{ss}$ [17], as

$$V_{gen} = \int_0^T C_{ss}(t)dt \tag{2.3}$$

Similarly, they also show that the value of a transmission line $V_{trans}$ between locations $a$ and $b$ can be replicated by the value of two locational spread options $C_{ls,a \to b}$ and $C_{ls,b \to a}$ covering both possible directions [17], as follows:

$$V_{trans} = \int_0^T [C_{ls,a \to b}(t) + C_{ls,b \to a} t] \, dt \tag{2.4}$$

Finally, for the spark spread option, they verify their result by fitting their model and using it to estimate the price of a particular generation asset, then compare their estimation with both an estimated plausible range of the real price and a standard DCF valuation. They find that the RO-based valuation gives a value close to the estimated real price and higher than the computed DCF price.

Oren [56] aims at quantifying the value of flexibility arising from customers willing to curtain their load when electricity prices are high. Acknowledging the earlier work by [25], they extend the therein presented 'callable forward' in a way to make it more sensitive to the timing of the warning issued to the customer that their service will be disrupted. In particular, instead of the simple American call proposed in [25], Oren suggests a more exotic 'double-call' that can be exercised at two separate dates, with the earlier date having a lower strike price, representative of the lower cost that a customer faces when being notified early of their curtailment. Oren goes on to assume that the electricity price follows a GBM process and then derives an analytical valuation of the proposed contract, however, he acknowledges that this assumption may be inadequate and refers to later work following up on this shortcoming.

Frayer and Uludere [24] quantify the value of electricity generation asset operation in a market with volatile prices using a RO-based approach. They argue that "under a RO framework, a flexible generating plant can be modelled as a series of European call options on the spread between electricity prices and variable costs" [24], connecting the key elements of RO to those of financial options. They justify this RO approach saying that "a plant has the right, but not the obligation, to burn fuel and produce electricity" [24]. They continue estimating the parameters for their model from real data, and for the valuation they use an adjusted version of the Black-Scholes formula [4], modified to take into account the partial predictability of electricity price movements and leaving volatility as the main unpredictable component. Valuing spark-spread options between energy and various fuel sources, they conclude that the "real options component of a valuation model captures incremental gains" and "reduces incremental losses", leading various assets exhibiting "relatively low value under traditional DCF methods" to "turn out to be much more valuable when analyzed using a real options approach" [24].

Tseng and Barz [71] also make use of RO to value power plants. In particular, they address that related works have "overlooked the plant's unit-commitment constraints", such as assuming no startup time and no up/downtime constraints, stressing that "the error made by failing to consider such physical constraints is often significant" [71]. They incorporate the constraints by formulating the valuation as a multistage stochastic problem. They assume that the underlying electricity and fuel prices both follow an Itô process and are Markovian. They use an indifference locus to map a given pair of fuel and electricity prices to an optimal decision. Tseng and Barz simulate their model with MC and investigate the sensitivities of their power plant valuation with respect to the imposed physical constraints and time, concluding that not taking into account physical constraints may overestimate the value and risk of power plants.

Keppo and Lu [37] extend the RO approach by considering the particular case of valuing the investment made by a large, market-affecting producer. Concretely, they restate this problem as an optimisation task in which the optimal strategy of an agent is to "maximise the value of the whole company", highlighting the importance of considering that a new market-price-affecting investment may have repercussions on existing assets of said participant in the same market [37]. They find that the price effect is reduced either by the participant itself hedging their portfolio by locking in a future cash flow using financial instruments (e.g. futures contracts), or in the case where there is significant competition on the investment opportunity that would cause the price effect irrespectively.

Hlouskova et al. [31] use a RO model to value electricity generation capacity. They identify increased volatility and the introduction of spot prices as emerging from ongoing electricity market liberalisation. As a consequence, they argue that utilities now "have the choice to turn off their own production assets and buy the necessary electricity in the spot market", resulting in the "unit commitment problem" of electricity producers turning into "profit maximisation under uncertain prices"[31]. They state that there is a resemblance between the choice to "turn on or off production assets" and the "exercise decision for an (exotic) financial option" [31]. More concretely, an electricity-producing asset "can be represented by a string of European spread options", respectively American options when taking into account operational constraints such as ramp up and down delays [31]. After discussing various constraints, they state that "in the case of constant fuel prices the [electricity-producing] turbine can be modelled as a call option on electricity, otherwise as a call option on the spread between electricity and fuel" [31]. They formulate this as an objective function representing the expected profits to be maximised. In particular, the choice of objective function mostly comes to their consideration of additional constraints of startup and shutdown times. As a solution to this optimal unit commitment problem, they present indifference loci which indicate the optimal action (turn turbine on or off) for any combination of electricity and fuel prices. For the underlying electricity prices, they use a mean-reverting, seasonal, spike-diffusion, discrete process, however, arguing that their proposed RO model "is flexible enough to incorporate any of the suggested price models" and the only restriction is that they have to be Markovian [31]. They conclude their research by showing how their model can not only be used to value an electricity-producing turbine but also for risk management purposes by simulating the profit-loss distribution.

Sezgen et al. [66] aim to valuate investments into electricity demand response technologies. They argue against the usage of the more traditional DCF methods, as these "underestimate the value of demand responsiveness"[66], and instead suggest an RO approach, saying that "On any given day or in any given hour, a facility operator or manager has the right but not the obligation to alter the operating schedule for loads" [66]. They specifically consider the case of a small, non-market-affecting end-user capable of observing DA prices and deciding at the same point to take a demand-responsive action such as load shits or load curtailment, hence reducing their usage. Defining load curtailment as the ability to "reduce electricity usage during a given time period without causing demand to increase during another period" [66], they reason that it "can be viewed as the equivalent of owning a strip of [European] options, one for each time period during which load can be reduced". In particular, assuming the price $S$ is the "amount the customer receives for curtailing, which is the day-ahead market price"[66], $K$ is the strike price reflecting "the customer's costs that would be incurred because of the curtailment"[66], and the option value $V$ is the payoff for not consuming energy otherwise consumed, they retrieve the formula for a European call option:

$$V = \max(S - K, 0) = (S - K)^+ \tag{2.5}$$

Under the additional assumption that the underlying on-peak DA electricity price follows GBM, they retrieve the Black-Scholes analytical solution for European option value [4]. They compute the option value of a customer willing to curtail 1 MW of load, a choice that depends on the DA prices and the strike price reflecting said customers' variable operating expenses and revenue losses. They conclude by proposing prices for 1 MW of curtailment capacity that are in the realm of real prices.

Cartea and González-Pedraz [13] use a RO approach to value an electricity interconnector between two different regions with two electricity DA markets. They justify using RO for this valuation because the owner of the interconnector "has the right, but not the obligation, to transmit electricity between two locations" [13]. Hence they define the financial value of the interconnector, once its built, as the difference between two strips of European options, more accurately a bull call spread on the spread between the prices of the two markets. They derive no-arbitrage bounds to compute lower bounds for the value. They first compute the value of the strip of spread options. They simulate the price uncertainty of both markets with a mean-reverting jump-diffusion process, arguing that this approach "captures the most important features of the price dynamics" [13]. For the valuation, they focus on pricing the transmission of 2 MWh of electricity per day, 1 MWh during on-peak and 1 MWh during off-peak hours respectively. They also add some seasonal deterministic components to this process. For determining the value of the strips of options, they consider liquidity constraints. They limit the maximum difference in price they consider in their valuation, arguing that beyond a certain limit, such price differentials would most likely be impossible to capitalise on in reality due to a lack of liquidity in such situations. Finally, they present their results for a one-year lease of an interconnector for various European markets.

Schachter and Mancarella [62] compute the short- and long-term RO value of a portfolio of demand-response (DR) customers. Concretely, they investigate the value an aggregator would pay for a portfolio of customers that are willing to reduce their demand during certain periods determined by the aggregator. As such DR assets, Schachter and Mancarella mention "thermal loads (heat pumps, hot water tanks, etc.) [that] are interruptible and can be curtailed through peak clipping" and "deferrable loads (e.g., washing machines) [that] can be shifted throughout the day"[62]. They however underline the importance of not ignoring the load recovery when these DR assets are later turned back on. They use a matrix representation to represent the load schedules and DR capabilities of each customer over a given number of discrete periods and appliances. They define the DR RO value $V_t$, over $N$ periods per day, as the discounted expected value of the optimal number of exercises ($\phi$ being a binary variable indicating exercise or not), and each exercise being computed as a call (or alternatively put) option[62]:

$$V_t = \max_{\phi} \mathbb{E}\left[\sum_{i=1}^{N} e^{-r(t_i-t)}\phi_i \max(S(t_i) - X, 0)\right] \tag{2.6}$$

with $\phi_{\min} \leq \sum_{i=1}^{N} \phi_i \leq \phi_{\max}$, $S(t_i)$ the electricity spot price and $X$ the strike price. While akin to RO models used in other reviewed literature, the introduction of upper and lower bounds on the number of exercises and maximisation thereof is novel. Also of interest is that Schachter and Mancarella retrieve upper and lower bounds for Equation (2.6) by defining the intrinsic and extrinsic values of the contract. Next, by comparing different load response policies, they analyse the potential profits and the 95% and 5% value at risk (VaR) of the DR portfolio of 1000 customers. Studying the impact of energy price volatility, they determine that volatility has a positive influence on the average profit as losses are limited due to the optionality of the contract, while profits are not. Lastly, investigating the long-term potential of DR using the Datar-Mathews method [45], they find that a DCF approach greatly undervalues DR compared to a RO approach.

Moriarty and Palczewski [49] investigate the valuation of providing reserve capacity in an imbalance market using a RO approach. They investigate the interactions of two agents, one being a network operator that may need electricity at times to ensure the stability of their grid, the other being a storage facility operator that is able to supply electricity. There is an optionality factor, as the network operator does not require electricity at all times, instead they need the option to request electricity when there is an imbalance. Moriarty and Palczewski model this scenario as an optimal stopping problem of American call options, in which the storage operator, who writes the option, aims to maximise their profit by optimising the purchase timing of the underlying and writing of the option. Furthermore, the underlying is electricity in an imbalance which is represented by a shifted Brownian motion stochastic process with zero mean [49]. Additionally, they also argue that the 'lifetime valuation' of the energy storage unit in the imbalance market is represented by an infinite sequence of these options, each option being written once the previous one expired. Later, they identify the conditions required for the storage operator have a positive expected economic profit, and also present the analytical value of both a single contract and the lifetime value of a sequence of contracts representing a storage facility. We note that this work by Moriarty and Palczewski [49] is among the few works that do research in ancillary markets and do not focus on the typical DA and ID electricity markets. In general there seems to be a dearth of research on the valuation of assets or services in ancillary markets.

A critical analysis of the RO approach to smart grid / flexibility investments is undertaken by Schachter and Mancarella [63]. They begin their review by highlighting the shortcomings of traditional methods such as DCF and NPV encountered in literature, in particular their controversial assumptions on the reversibility of investments, the passive nature of management, and the certainty of uncertainties and future cash flow. They refer to literature criticizing analytical solutions used to value RO as these are often limited to a single source of uncertainty and moreover unrealistically assume that said uncertainty follows GBM with constant drift and volatility and without jumps and mean-reversion. With respect to financial options, Schachter and Mancarella discuss, among others, the special considerations required for volatility and the discount rate of RO. Unlike financial options where volatility is given by a single parameter, RO volatility encompasses "uncertainty of the project's cash flows, which is a combination of multiple interrelated variables (not only prices) that cannot be separated and is allowed to change over time as a result"[63]. Discount rates are also different in RO. Unlike financial options where the "risk-free discount rate is used on the basis of a number of assumptions that are reasonable in stock markets" [63], this same reasoning does not apply for RO, despite being used by many. Instead [63] suggest as alternative the advocate the discounting approach found in the Datar–Mathews method [45]. However, they inform that deciding on the 'correct' risk premium or discount rate to apply in RO requires consideration of various aspects.

Another domain where RO are used for valuation is storage valuation. We did at some point consider a valuation of the optionality component of energy storage during this project, but later abandoned this idea in favour of valuing demand-response which is more suited for electricity markets. We can however not ignore the benefits and impact of the reviewed literature on option-based energy storage valuation on our general understanding of RO theory.

Reviewing storage valuation literature, we found that while there is some literature on the valuation of electricity storage such as batteries or pumped storage hydropower (e.g. [76, 61, 48, 68] to name a few), most of this literature is focused on the energy markets and not capacity markets and ancillary services. Furthermore, the field of gas storage valuation seems to be more mature, in particular with respect to real-options derived approaches. As some of these

RO approaches in (gas) storage valuation are closely connected to the other RO approaches discussed before, we will briefly review a few selected works.

Among the most popular works is Boogert and De Jong's [7] use of swing options to quantify the value of gas storage. While not explicitly referring to their methodology as reals options analysis, parallels can easily be drawn. They incorporate various storage constraints into their models such as working volume and withdrawal rates, then compute the optimal position (i.e. inject gas, do nothing, withdraw gas) for a storage owner to take based on market conditions. They establish the link between owning physical storage, a storage or swing option contract and even parallels to American options, resulting in their choice of least-squares Monte Carlo (LSMC) for the valuation [7]. They conclude their work by presenting a pricing algorithm incorporating various volume and other constraints and allowing for optimal control of gas storage.

Another work using a RO approach to value operational flexibility is Kryzia et al. [40] who value a gas turbine investment using RO and MC simulation. They reproduce the operational flexibility of the generation asset by representing the option to shut the plant down when prices are below a certain strike as a put option while starting the plant up when prices are above a certain strike as a call option. They conclude their research with an investigation of the risk profile of the investment in the form of an (extended) net present value distribution.

Thompson et al. [69] also used RO to determine the valuation and optimal operation of gas storage facilities of natural gas storage facilities". They justify the choice of RO valuation, claiming that, as an example, "a natural gas well can be modelled as a series of call options on the price of natural gas, where the strike or exercise price is the total operating and opportunity costs of producing gas" [69]. They represent their valuation problem as an objective function, in the form of a Bellman equation, to be optimised, and solve said objective function with a partial differential equation approach.

For further investigation of RO valuation in the energy sector using LSMC, we refer to the review of [53] who compares various approaches found in the literature.

## 2.4 Monte Carlo methods

Monte Carlo (MC) methods consist of repeatedly "sampling randomly from a universe of possible outcomes and taking the fraction of random draws that fall in a given set as an estimate of the set's volume" [27]. MC is commonly used when one wants to know more information about a random variable sampled from a distribution, and the distribution is either unknown or known but no analytic approach exists. In this context, MC consists of repeatedly sampling this random variable and then using estimators to deduce the desired statistics such as the mean and variance. The law of large numbers ensures that these statistics from a sample converge to the correct value of the population as the number of samples is increased, and the central limit theorem can be used to quantify the magnitude of the error remaining after a given amount of sampling [27].

While the concept behind MC has been used by various previous researchers, the origin of MC as it is known today can be attributed to the work of physicists Stanislaw Ulam and John von Neumann working on the Manhattan Project, and the term 'Monte Carlo' has been coined by Nicholas Metropolis [47]. Since then, MC methods have been used in a variety of domains. Given the nature of our work, of particular relevance is the work of Irish economist Phelim Boyle [9] who was the first to use MC for option pricing in 1977. They priced path-independent European options. It is worth pointing out however that the Black-Scholes-Merton model, published earlier in 1973, already presented analytical solutions for European options [4], hence limiting the usefulness of MC for these options. Arguably more important was the latter application of Monte Carlo methods to price path-dependent American options (for which no closed-form analytical solution exists), pioneered by Longstaff and Schwarz [42]. They combined MC methods with least squares regression to find the optimal exercise time for each simulated option, which in turn is used to value American options. They refer to this technique as least-squares Monte Carlo (LSMC).

# Chapter 3

# GOPACS

GOPACS is a Dutch flexibility market serving as the platform for all congestion- and flexibility-related issues in the Netherlands and providing access points to current flexibility auction announcements and historical data. GOPACS is not a trading platform itself and any orders are cleared via other market platforms. In this chapter, we will cover in-depth all relevant aspects related to GOPACS and dive deep into the available data.

## 3.1 GOPACS technicalities

### 3.1.1 Congestion in Dutch law

We will start our investigation of GOPACS by giving a quick overview of congestion in Dutch law, defining some concepts that will be of importance later. In the Dutch regulation 'Netcode elektriciteit' [60], there are two definitions of congestion[1]. The first definition is presented in §9.1 about the resolution of physical congestion and referred to as 'transport problems' (TP). It specifies that it is the job of the grid operators to identify said problems based on the received transmission forecasts. Inter-grid congestion is to be resolved in combination with the responsible grid operators. The procedure for solving these transport problems foresees grid operators sending out requests to the connected parties to produce/consume more or less energy, indicating the length of the problem period [60]. During this problem period, further restrictions on changes made to transport forecasts may be issued by grid operators. Additional information to be shared by the grid operators among connected parties are the direction of the restriction, the location of the network and bidding zones and the expected duration of the restriction [60].

The second definition of congestion is presented in §9.2 about the resolution of, amongst others, structural congestion, and referred to as 'congestion management' (CM) [60]. CM is applied, among other reasons, to bridge the period that remains until the concerned part of the grid has been strengthened, modified or expanded in a way that the requested transport can be made available in full. Again, a notice to concerned parties is to be made via means of a pre-announcement, containing the expected congestion area, the problem period, the cause of the congestion, and other information [60].

Lastly, in both cases of congestion, according to §13.5, the exchange of said information is to be organised via a web portal [60]. While not explicitly mentioned in the netcode, this web portal is GOPACS, https://gopacs.eu.

[1] In electrical engineering, two types of congestion are voltage congestion and current congestion. The two definitions of congestion in Dutch law are not related to these types of congestion.

### 3.1.2 Participants

According to GOPACS[2], participation in GOPACS is interesting for any market party or large consumer that is able to influence their electricity consumption or generation and thus has flexibility available to trade [28]. In particular, there are 3 groups of participants: Distribution system operators (DSO), 'Handelsplatform' (HP) participants and 'Congiestiemanagement Portaal' (CP) participants. Since GOPACS is a flexibility market, we refer to both HP and CP participants together as flexibility service providers (FSP).

The DSOs are participating in GOPACS because they are required to solve congestion in their grid. Technically all Dutch DSOs should solve any congestion issues encountered via GOPACS, however, the exact list of DSOs listed on the GOPACS website varies is inconsistent. Tennet, Liander, Stedin and Enexis are represented in the available data sources, however, Westland Infra, Coteq, Rendo and Enduris appear also in various mentions on the website.

'Handelsplatform Deelnemers' or HP participants are market parties that trade energy on a structural basis, such as energy suppliers, aggregators and PV parties [28]. They are referred to as HP participants as they submit their GOPACS bids via a Trading Platform [33]. Currently, the only trading platform that supports submitting bids to GOPACS is Energy Trading Platform Amsterdam (ETPA), however, additional platforms are planned for the future [28].

'Congiestiemanagement Portaal Deelnemers' or CP participants are large consumers who typically do not trade on energy trading platforms, such as horticulturists (tuinders), solar and wind farms and industrial companies [28]. They are referred to as CP participants since they submit their GOPACS bids via a GOPACS-specific congestion management portal, which is intended particularly for market parties that can provide flexibility, but do not have access to energy trading platforms and thus need a separate platform to submit their offers [28].

The relation between the two access points (HP and CP) and two types of congestions mentioned in the Dutch grid code (TP and CM problems) is not very clearly stated on the website. It is mentioned[3] that the CP is 'mainly' ('met name' in the original Dutch sentence) intended for cases where CM is applied and the network operators ask specific large business customers in the congestion area in question to place flexible bids [28]. We were unable to find any source explicitly preventing CP participants from also bidding in TP problems.

### 3.1.3 IDCONS

This summary of the GOPACS process is partially derived from the IDCONS manual published on the GOPACS website[33].

Typically, when we talk of congestion, we assume a situation where there is a surplus of energy in one location, a dearth of energy in another location and the grid between both locations is unable to satisfy the transport of energy between those locations[4].

All energy market participants have to provide their planned energy consumption or production to grid operators. Using this information, grid operators can compute in advance whether there may be congestion based on these forecasts received. If they detect congestion, they need to buy flexibility, which they do by organising an auction where they look for energy buyers and sellers. Announcing this auction is done via a market announcement (MA) on the GOPACS website. This market announcement contains a variety of information, such as the

[2] https://www.gopacs.eu/wat-is-gopacs/, near the bottom of the page

[3] https://www.gopacs.eu/hoe-werkt-gopacs/, near the bottom of the page

[4] There are exemptions to this assumption, sometimes there is only a surplus but no dearth, or the opposite. Also, this assumption mostly applies to physical congestion, which is different from structural congestion.

forecasted time span and location of the congestion problem. The time span is divided into 15-minute periods, also called program time units (PTU) [8, 33].

We refer to the location with a surplus of energy as a buy area since the grid operator is looking for energy buyers there. Buyers submit a bid price which is the price they are willing to pay for energy. On the other side, the location with the dearth of energy is the sell area, as the grid operator is in need of energy sellers there. Sellers submit an ask price, the price they are willing to sell energy for. In this 'energy-as-the-product' point of view (POV), the grid operators are both buyers and sellers of energy, depending on location. However, from a 'flexibility-as-the-product' POV the grid operators are the buyer of flexibility whereas the FSP are the sellers of flexibility. This second POV is commonly encountered in literature such as [72].

FSP can either sell flexibility by offering to buy energy (submitting buy offers), or they can sell flexibility by offering to sell energy (submitting sell offers). A buy order is the offer of a market participant to either consume (offtake) more or produce (infeed) less electricity than specified in their electricity program and transmission forecast. A sell order is the offer of a market participant to either consume less or produce more electricity than specified in their electricity program and transmission forecast [33]

Going back to the 'energy-as-the-product' POV, the grid operator tries to combine an order and a counter-order into an Intra-Day CONgestion Spreads (IDCONS) product. IDCONS are defined as the combination of an energy buy order, an energy sell order and the spread between them. Intra-Day refers to IDCONS being traded on the ID market. Both orders have to match in time and duration. There are no minimum or maximum prices or volumes for IDCONS. Grid operators pay the spread between the order and counter-order, hence it is in their interest to minimise it. However, they need to ensure that the created IDCONS solves the congestion problem. IDCONS are limit orders, which means that they can be called partially in terms of power, but not in terms of time. Users can submit 'all-or-none' orders which can not be called partially, however several conditions are to be met before those orders can be submitted, such as sufficient liquidity in the market as well as the user having an installation with technical limitations and having communicated this limitation to the grid operator.

From the 'flexibility-as-the-product' POV, the grid operator is buying flexibility for the spread price, whereas energy buyers are selling flexibility for the difference between the current electricity market price (determined via e.g. DA or ID market) and their bid price for buying energy and energy sellers are selling flexibility for the difference between the current electricity market price and their ask price for selling energy.

IDCONS are then transmitted to an ID trading platform, currently the only platform supported is ETPA. The IDCONS is executed with other ID market orders, and the financial settlement for FSPs takes place on their respective platform. The FSPs are paid / have to pay[5] what they offered in a paid-as-bid style. Once an order is included in an IDCONS, the user who submitted it is obliged to fulfil their obligation. They are not allowed to change their infeed or offtake in the opposite direction to their offer [33]. However parties are only paid the price of their order, and if they supply more, they will not receive additional compensation. The full power of the order must be available at the start of the delivery period, and the contracted volume must be cleared during the agreed PTU and in the agreed location.

[5]even though an FSP sells flexibility, they may do so by offering to buy electricity and hence still need to pay the price for this electricity.

## 3.2 GOPACS data

GOPACS publishes historical data on their website. This data allows us to get a deeper understanding of GOPACS by observing past trends, and the evolution of the market over the years and provides insights into the GOPACS dynamics that can be used to validate a model. As of writing, there are four sources of data for GOPACS, with one having been recently[6] added.

TABLE 3.1: Overview of the 4 GOPACS data sources

| **Official name** | **English name in API** | **Abbreviation** |
|---|---|---|
| Marktberichten | Market announcements | **MA** |
| Afgenomen volumes in MWh | Cleared buckets | **CB** |
| Afgenomen orders | Expenses | **EX** |
| Performance metrics (totalen) | Performance metrics | **PM** |

Table 3.1 mentions the four sources of data. MA and CB contain the times and volumes of GOPACS auctions, and we will investigate them and the other data contained in these two sources first. Afterwards, we will investigate the price data contained in EX and PM.

### 3.2.1 Market announcements and cleared buckets

The first two sources of data are the market announcements[7] (MA) which contain GOPACS auction announcements, and cleared buckets[8] (CB) which contain the volumes ('buckets') cleared during these auctions. Both sources are extensive, continuously updated and contain historic data at the highest level of granularity (i.e. for every PTU).

MA are among the most extensive data in terms of information and contains various details relevant to market participants. MA are issued before a GOPACS auction and contain the volumes of energy forecasted to be needed in each location in order to solve a congestion issue. In addition to the volume, they also contain various timestamps, location data, and other information. They are documented in detail in [43]. There are three different types of market announcements: (1) Transport problems (TP), congestion management (CM) and free text announcements (FT). Out of the 757 announcements published until May 24th 2022, 104 (13.7%) were FT, 218 (28.8%) were CM and the remaining 435 (57.5%) were TP.

CB are less extensive than MA as they only contain time and volume data. They are issued after GOPACS auctions and contain the actual volumes cleared during the auction. The data is structured relatively simple and consists only of the date of the problem period PTU as well as both the volume upwards and downwards of the congestion. In all entries, the downwards volume is equal to the upwards volume, which is also to be expected. Moreover, there are some entries with zero volume, but none with a negative volume.

Figure 3.1 on the next page shows the evolution of both MA and CB announcements over time, grouped by month. First, we see that MA FT messages were common until early 2020 and have since practically not been used anymore. On the other hand, MA TP messages have started to rise at that point and are now the most common MA announcements. Another thing that is interesting is that it seems from this view that the CB announcements represent the same data as both TP and FT combined, as they align somewhat well during some periods, such as

[6] GOPACS performance metrics were added sometime between December 2021 and January 2022

[7] https://idcons.nl/publicannouncements#/announcements

[8] https://idcons.nl/publicclearedbuckets#/clearedbuckets

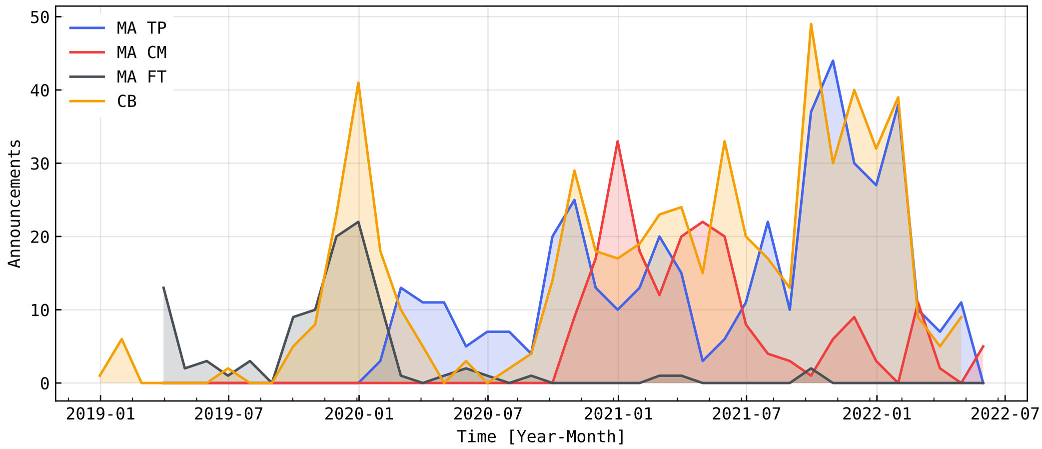



FIGURE 3.1: The number of GOPACS MA announcements and CB periods over time, aggregated monthly. In black, the number of CB problem periods is also given as reference.

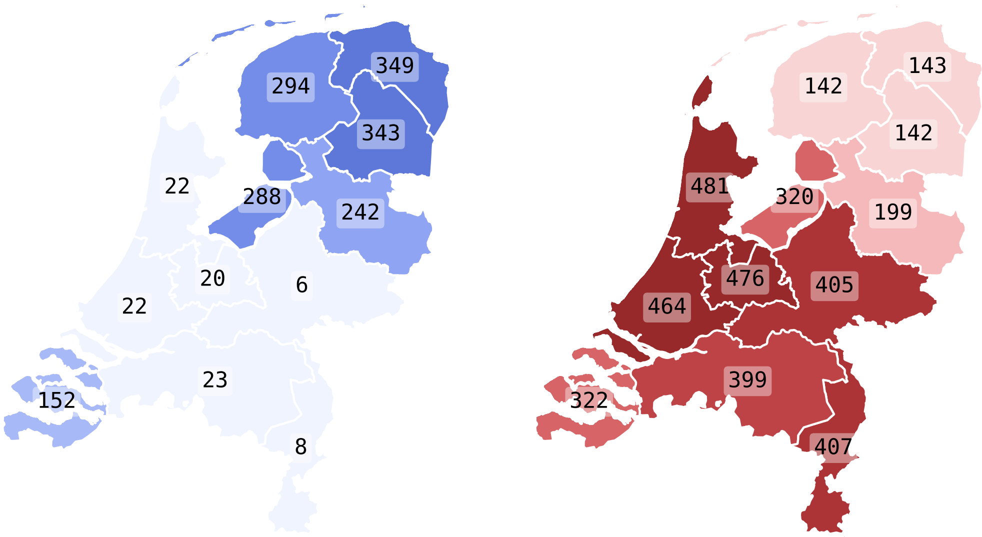
the spikes in the winter months of 2020, 2021 and 2022. However, during other periods they do not align very well, such as e.g. during the summer of 2020 when plenty TP messages were announced but practically no CB messages.

TP and CM represent the two types of congestion problems defined in the Dutch law mentioned in Section 3.1.1 on page 17. FT messages contain lots of empty fields according to the API and their main content is unstructured text. Based on their usage concentrating mostly at the beginning of GOPACS and decreasing over time, we hypothesise that FT messages were used before TP and CM types were introduced and their unstructured nature did not permit us to associate them retroactively to either category. Another assumption is that they may simply be used as a means of communicating information that can not easily be associated with either category.

### Location data

One interesting set of information associated with the market announcements are the location indications. For CM messages, the location of the CM problem is given either in the buy, sell or problem area fields. Interestingly, of 187 CM announcements until February 2022, 173 messages shared the same problem area: Neerijnen. According to the Dutch DSO Liander, "The Neerijnen region in Gelderland is one of the areas where the power grid has reached the limits of its capacity. [...] We are expanding the power grid [...] However, the work will not be completed until 2023" [15]. This observation aligns well with the definition of CM in the Dutch netcode [60] (see Section Section 3.1.1 on page 17). TP messages contain a buy and sell area location, which are always two potentially overlapping sets of one or more Dutch provinces. FT messages also contain buy and sell area locations, however here the data is more unstructured and contained in the free text. However, in most cases, it is again Dutch provinces, and sometimes Neerijnen. The distribution of Dutch provinces mentioned across all FT and TP messages until February 2022 is shown in Figure 3.2 on the following page. CB do not contain location data.

Figure 3.2 on the next page shows the distribution of buy and sell areas across Dutch provinces. If a province is part of a buy location of a GOPACS MA, it means that the DSO issuing the announcements is looking for buyers of electricity in said region, or in other words, they are faced with a surplus of energy and want to sell this energy to participants. The opposite holds true for sell locations. Northern Dutch provinces seem to be mostly appearing as buy areas,




(A) Appearance as buy location (i.e. where DSO is looking for buyers of electricity)

(B) Appearance as sell location (i.e. where DSO is looking for sellers of electricity)

FIGURE 3.2: The amount of times each province was included in any MA announcement.

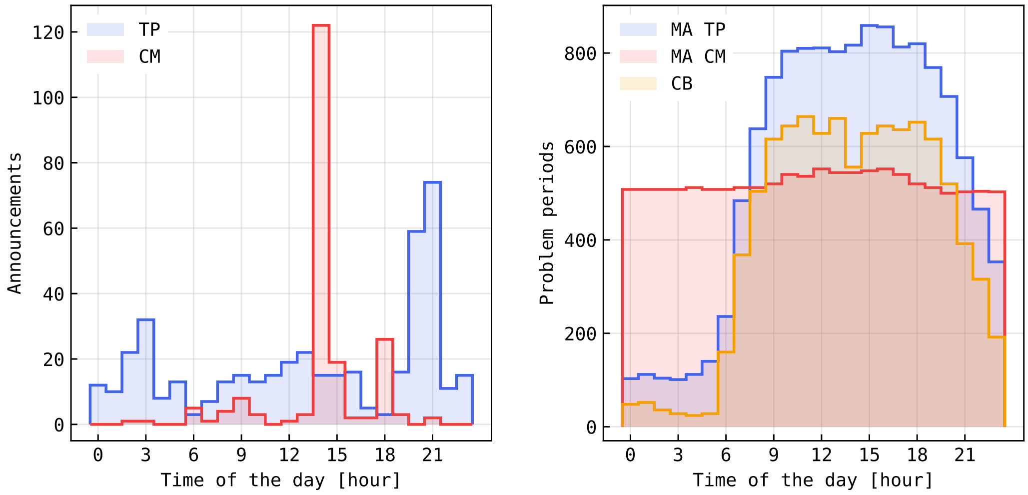
whereas southern Dutch provinces seem to be mostly appearing as sell areas. This implies that congestion issues in the Netherlands seem to be, at least partially, caused by a surplus of energy in the north, a lack of energy in the south, and a power grid unable to solve this imbalance. An investigation into the geographic and demographic factors (such as e.g. location of DER, population density) responsible for this imbalance is out of the scope of this work.

### Time data

The next set of interesting data found are the various times associated with each MA or CB announcement. MA contain timestamps for their creation and last-updated times, for the start and end of their bidding periods and for the start and end of their problem periods. CB contain only the timestamps of the start and end of each particular PTU problem period, as well as the timestamps of the start and end of the full group of problem periods the particular problem period was part of.

Figure 3.3a on the following page shows when announcements are created throughout the day. We can see that TP announcements are distributed somewhat equally throughout the day, with two smaller spikes around 03:00 and in the afternoon as well as the biggest spike from 20:00 to 22:00 in the evening. For the CM announcements, nearly all of them happen in the afternoon, in particular from 14:00 to 15:00.

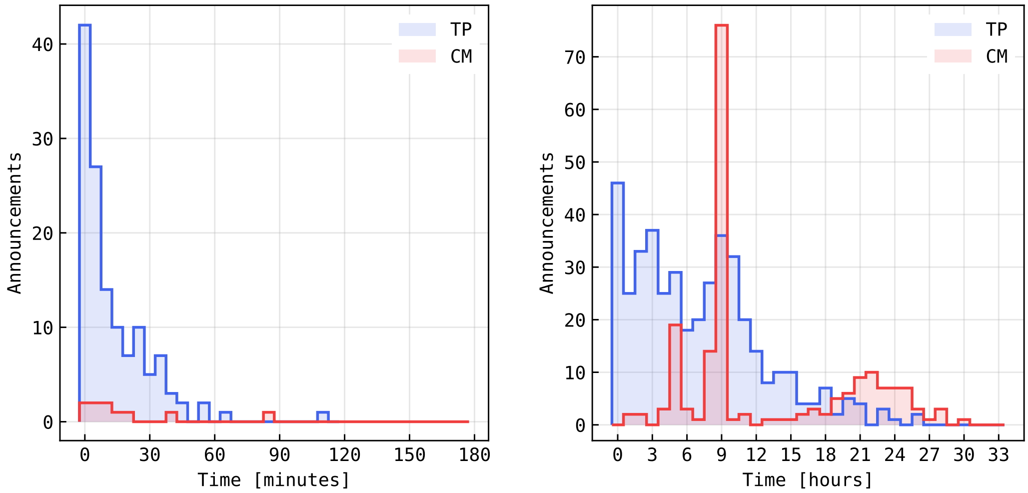
The next set of times is shown in Figure 3.3b on the next page. Again distributed throughout the day, this time we see the hours at which there are problem periods. For TP announcements, nearly all problems take place during the day, starting in the morning and gradually ending around midnight. On the other hand, the MA problem periods always last a full day, with no significant increases at particular hours. It is debatable whether these problem periods really last exactly 24 hours, or whether the responsible DSO just announces them for the full day due to their inability to specify more precise periods. Lastly, CB problem period times are distributed very similarly to TP announcement times.




(A) Announcements per hour.

(B) Problem periods per hour.

FIGURE 3.3: The number of both announcements and problem periods for each hour of the day. The announcements only contain MA data as CB does not include the time of announcement.



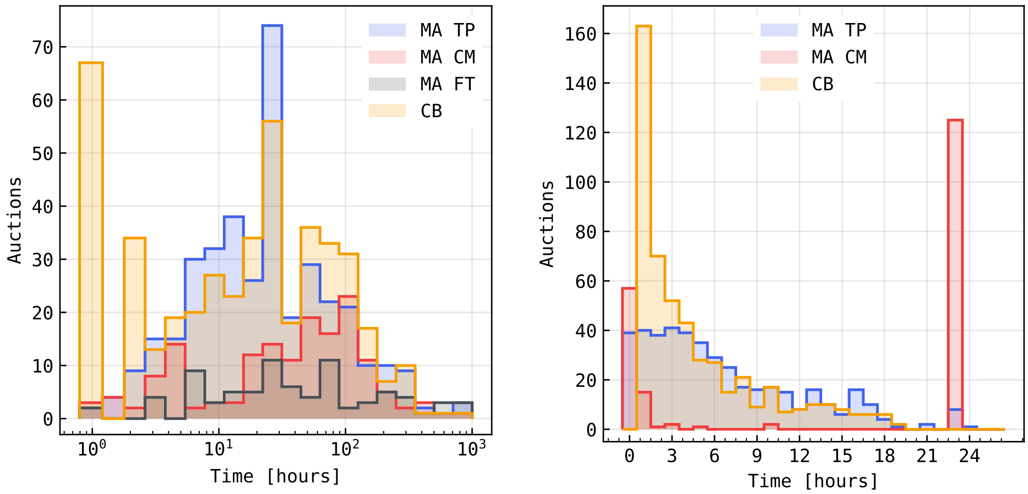

(A) Announcement creation to bidding period.

(B) Announcement creation to problem period.

FIGURE 3.4: The distribution ot the time elapsed between the creation of an announcement and the start of either the bidding period (left) or problem period (right).

In Figure 3.4a we plot the distribution of the time difference between the announcement of a problem and the associated bidding period start. The first thing to notice is that nearly all bidding periods take place less than 1 hour after their announcement, and in most cases they take place within the first 10 minutes. Also note that not all MA contain bidding start and end times, in fact, 40% of TP do, but only 5% of CM do and none of the FT do. Most recent TP messages contain bidding periods. Finally, while not shown on this plot, is that the median bidding period duration is around 30 minutes, the third quantile is 1 hour and the mean is 1 hour 30 minutes. Combining this information, if a GOPACS announcement comes with a bidding period, said bidding period is already over 1 hour after the announcement in most cases.

In Figure 3.4b on the previous page we see the time between the start of the announcement and the start of the problem period. For TP announcements, in most cases, the problem period is following rather soon after the announcement. Only in a handful of cases was the problem period more than 12 hours after the announcement. The median time between the announcement and the start is 6:30, the mean time is close to 7 hours. For CM, we have a different picture again. Here the time is nearly always either 5 or 9 hours. This makes sense, however, as in Figure 3.3a on the preceding page we found that CM are typically announced between 14:00-15:00, adding the 9 hours we arrive at a starting time of 23:00-24:00. In combination with Figure 3.3b on the previous page, we see the full picture: CM announcements are typically announced during the afternoon of a given day and specified as a problem period the whole following day.




(A) Interarrival: Auction end to next auction start. (B) Duration: Auction start to auction end.

FIGURE 3.5: The distribution of GOPACS auction interarrival times and auction durations.

Figure 3.5a shows the interarrival time distribution, i.e. the time between subsequent GOPACS announcements. The $x$-axis is given in logarithmic scale as there is a wide range of interarrival times. For TP announcements, most announcements take place around 24 hours after the previous announcement. The mean time between arrivals is close to 2 days, while the median is 17 hours. However, there are also both shorter and longer interarrival times, up to 1000 hours, equivalent to around 42 days. The CM announcements are similarly distributed albeit slightly more skewed towards the right, here the mean is close to 2 days and the median exactly 24 hours. The distribution of FT messages is also rather spread out with a mean of 9 days and a median of 1 day. Their interarrival times are probably larger since they were predominantly used in the early days of GOPACS before the market became more established. Finally, for the CB we see that often the interarrival times are very low, around 1 or 2 hours and less. This could be due to mistakes in the separation of individual CB periods. There are however also very large interarrival times, and overall the mean is 2.5 days and the median 21 hours.

Lastly, in Figure 3.5b we see the distribution of auction duration times, i.e. the time between the start and the end of each problem period. For TP again, auction durations typically range anywhere from 0 to 7 hours, but longer ones of up to 24 hours are also not uncommon. The mean is 7 hours, the median is 5 hours. Instead for CM, we have a very different picture, nearly all auctions last either 23-24 or up to 1 hour, the mean is 15 hours while the median is 23:45. It is possible of the 1 hour auctions are actually the remnants of auctions that lasted a bit longer than 24 hours and were announced in separate announcements, therefore. For the CB, a lot of auctions last only very shortly, with a mean of 4 hours and a median of 3 hours. Apart from this spike, most auctions are similarly distributed to the TP announcements. In combination with Figure 3.5a on the preceding page, it is likely that for some GOPACS announcements, while they were announced as one continuous longer block in MA, the CB actually displayed various shorter split-up periods, which would both explain the high density of shorter CB interarrival times and shorter CB auction durations – they are simply longer auctions that were split up.

### Volume data

The last set of data contained in the GOPACS market announcements is volume data. For MA, the volume specifies a quantity of power in MWh that is estimated to be needed to solve the congestion problem. Both TP and CM announcements come with volume requirements, while FT announcements do not. CB data contain the actual volumes cleared. In all cases, the mean volume is specified for each hour, as in practically all cases, the 4 PTUs of a given hour have the same volume.

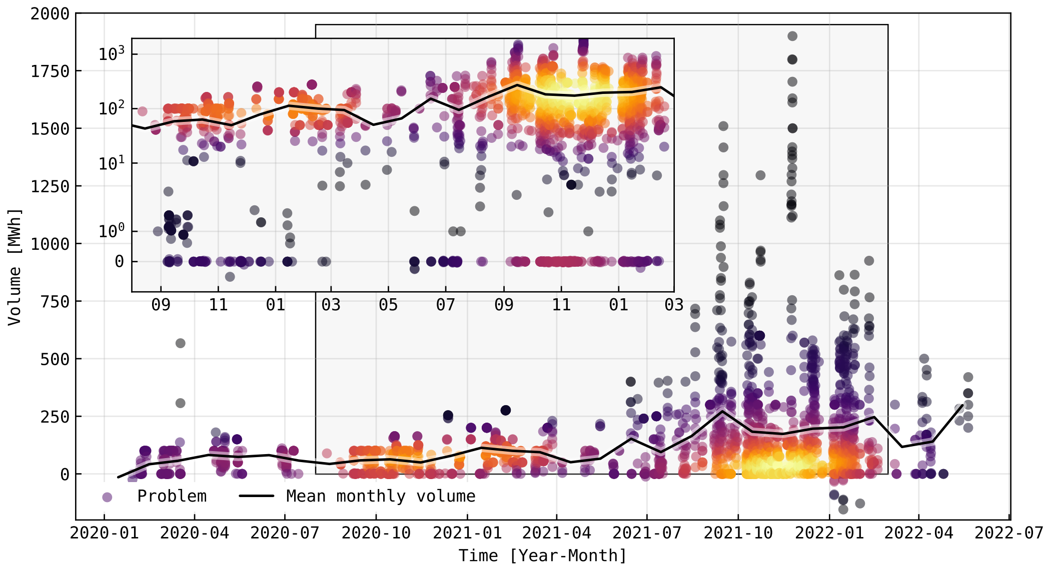



FIGURE 3.6: A scatter plot of the MA TP volume over time. Each dot represents the average volume of a single hour. Also plotted is the monthly mean volume. The inset plot shows the grey area of the main plot, but on a logarithmic scale to convey a different perspective of the same data. Both for the main and inset plot, the dots are coloured according to their density, with black the lowest density and yellow the highest density. The density is estimated using kernel density estimation and is included to better highlight dense regions.

Figure 3.6 shows the evolution of the TP volumes. The mean volume is 140 MWh, the median is 95 MWh. The first and third quarters are 17 and 170 MWh, however, there are also extreme values such as 1800 MWh or negative volumes[9]. We can also see that there is some trend in volume, in particular, a lot of activity happens during the winter months, such as between October 2020 and April 2021, and September 2021 and March 2022.

[9]Technically, the volume should not be negative as a congestion issue is bidirectional (surplus in location $A$, dearth in location $B$ and congested connection), so a negative volume from $A \rightarrow B$ should instead be announced as positive volume from $B \rightarrow A$. Moreover, negative volumes only appear in the 'raw' MA data, there are no negative values displayed on the GOPACS website, and also there are no negative values in the CB data. We are thus unsure where these negative values come from, what they represent or why they are not shown on the website itself, and we will ignore them by treating them as 0.

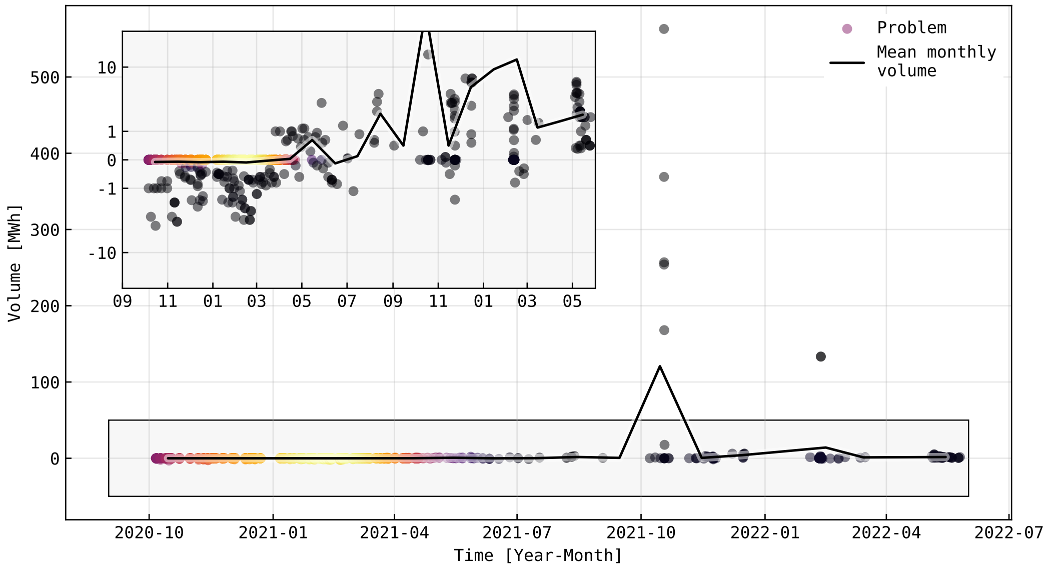



FIGURE 3.7: A scatter plot of the MA CM volume over time. Each dot represents the average volume of a single hour. Also plotted is the monthly mean volume. The inset plot shows the grey area of the main plot, but on a logarithmic scale to convey a different perspective of the same data. Both for the main and inset plot, the dots are coloured according to their density, with black the lowest density and yellow the highest density. The density is estimated using kernel density estimation and is included to better highlight dense regions.

Figure 3.7 shows the evolution of the CM volumes which are very different from the TP volumes. The mean volume is 1.2 MWh, while both the first quantile, median and 3rd quantile are all 0 MWh. While volumes seem to be trending slightly upwards in recent months as seen in the inset plot, they are still very low. Given our understanding of the usage of CM announcements, it makes sense that there is also no seasonal trend visible.

Figure 3.8 on the next page shows the evolution of the CB volumes. They bear some resemblance to the FT volumes, in particular, they also show seasonal trends, such as increased volume during the last 3 winters. The mean volume is 140 MWh (147 MWh for MA), the median is 99 MWh (95 MWh for MA), and the first and third quarters are 5 MWh (17 MWh for MA) and 200 MWh (170 MWh for MA) – all very close to the MA volumes. In order to take a more in-depth look between the volumes of MA and CB, we look at their histogram in Figure 3.9 on the following page.

Figure 3.9 on the next page shows a histogram of the GOPACS volumes, in particular the volumes of the TP announcements and the CB data. Both in normal scale (main plot) and logarithmic scale (inset plot), they look very similar, with a spike around 0, most of the volumes between 0 and 200, and followed by a somewhat exponential decrease. Two main observations can be made. First, MA contain lots of 100 MWh announcements, however, these are not really present for CB. Secondly, MA both contain extreme volume requirements such as negative volumes and high values, but for CB these are not present, here the minimum and maximum volumes were 0 MWh and 1200 MWh respectively.

Finally, MA also contain other information, such as the name of the DSO who called the auction and the compliance type, indicating whether an auction is mandatory or voluntary to participate in. We have not explored this data in much detail as it is not relevant to our endeavour.

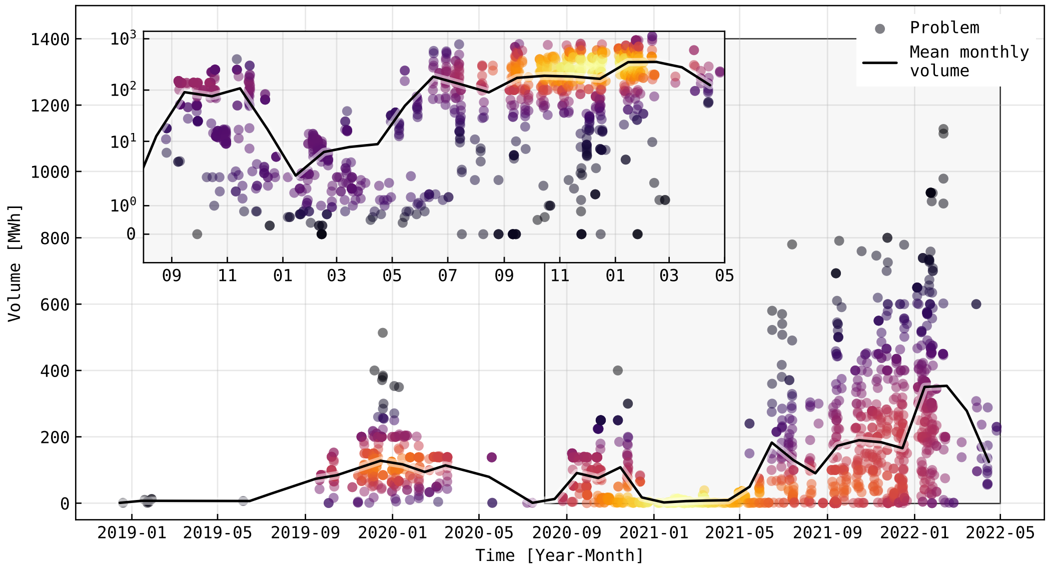



FIGURE 3.8: A scatter plot of the CB volume over time. Each dot represents the average volume of a single hour. Also plotted is the monthly mean volume. The inset plot shows the grey area of the main plot, but on a logarithmic scale to convey a different perspective of the same data. Both for the main and inset plot, the dots are coloured according to their density, with black the lowest density and yellow the highest density. The density is estimated using kernel density estimation and is included to better highlight dense regions.

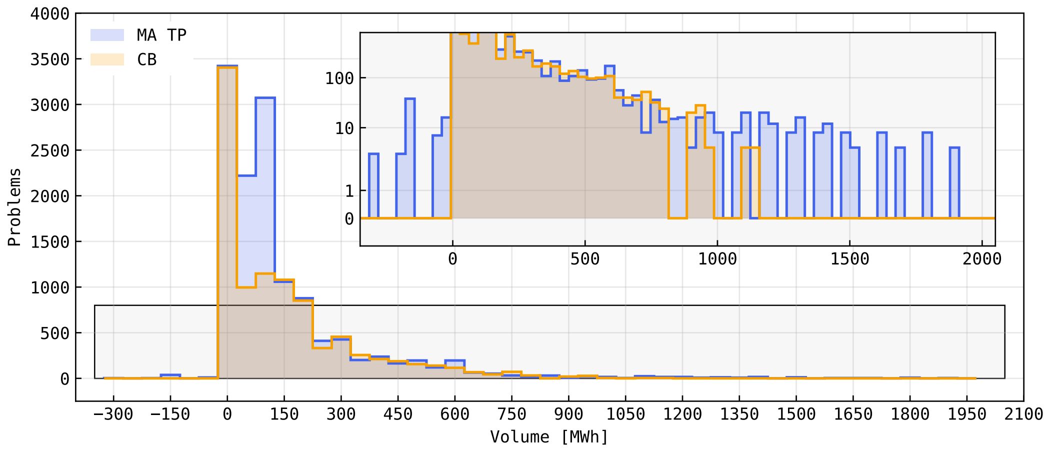



FIGURE 3.9: Histogram of both MA TP and CB volumes for each PTU of a problem period (i.e. for each problem). Again an inset plot is featured, showing an extract (highlighted by the grey area on the main plot) of the same data, but in logarithmic scale.

### 3.2.2 Expenses and Performance Metrics

Gopacs expenses (EX) contain the monthly total of expenses and volume for each of the participating grid operators. Performance metrics (PM) also contain the monthly total of expenses and volume, as well as buy prices, sell prices and the number of currently active grid connections. Both EX and PM contain identical volumes and expenses on a monthly basis. The differences are that EX contains this data split among the participating grid operators, while PM contains the additional buy and sell price data.

The expenses represent the spread price paid by the grid operators in order to combine two orders into an IDCONS. One can get the monthly average spread price (in €/MWh) when dividing the expenses by the volume. Similarly one can get the monthly average buy and sell prices. The evolution of monthly spread, buy and sell prices are shown in Figure 3.10. In addition, we plot the DA prices as a reference electricity price.

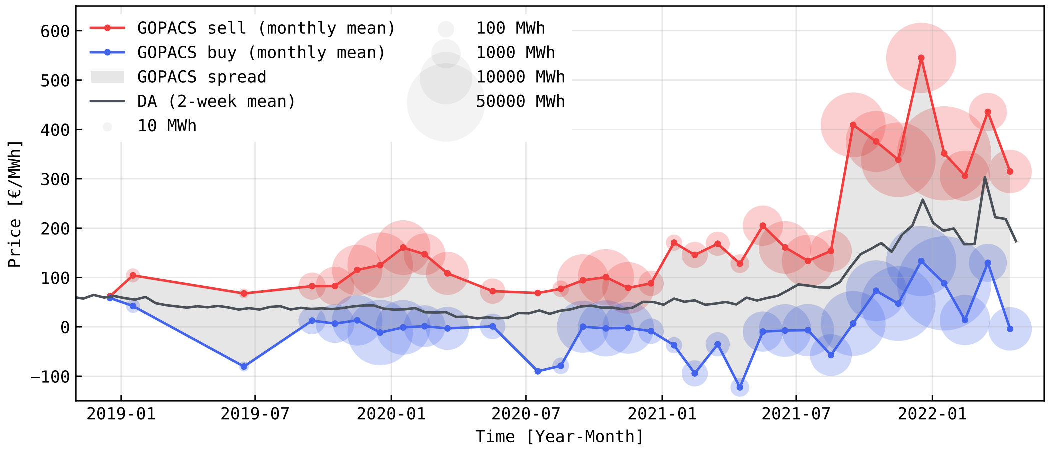



FIGURE 3.10: The monthly mean of the buy, sell and spread price on GOPACS. The monthly average DA price is given as reference. Each dot represents a month, with the size of the circle surrounding it indicating the volume. Some months do not contain any data, as there were no recorded GOPACS auctions.

Figure 3.10 shows the monthly buy, sell and spread prices as well as the volume and also the DA price. Various interesting insights are revealed. First, it looks like GOPACS prices are a deviation of DA prices, in particular, we see that rising DA prices in the last years have resulted in rising GOPACS prices as well. Even particular price spikes of DA, such as the one during January 2021 are reflected in spikes of GOPACS prices. The Pearson correlation coefficient between monthly GOPACS sell prices and monthly DA prices is 0.929, and between GOPACS buy prices and DA it is 0.669, so there is a relatively strong correlation between both markets, and since DA is by far larger, it leads us to the conclusion that GOPACS prices are at least to some degree dependant on DA prices. Moreover, another interesting observation to be made are that GOPACS buy prices are relatively often even negative, meaning that participants get paid to buy energy.

Figure 3.11 on the following page shows a different aspect of GOPACS prices, namely the data from EX. While EX does not contain the buy and sell prices but only the spread price, it does however contain the expenses per DSO. With that information, we can explore their relative contribution to GOPACS. It becomes immediately apparent that TenneT is by far the largest DSO participating in GOPACS, responsible for 99.938% of the volume transacted via GOPACS, with a total of 353,854.8 MWh cleared (from start to April 2022 included). Liander cleared 185.9 MWh (0.053%), Enexis 24.8 MWh (0.007%), Stedin 9.5 MWh (0.003%) and both Enduris and Westland Infra have not cleared any volume in GOPACS according to this data source, hinting that their grid may be less impacted by congestion. Other than this, some extreme prices can be seen, such as in October 2021, where the grid operator Liander paid 4089.82 € for 1.1 MWh, resulting in a price of 3718 €/MWh. This value is in contrast with the average price of 163.7 €/MWh and max price of 442.9 €/MWh respectively achieved in the DA market in the same month. Also in the same month, TenneT paid on average 302.2 €/MWh for a total of 18,218 MWh exchanged via GOPACS.

So far, we have only shown monthly GOPACS prices, as both PM and EX only publish monthly aggregated price data. More granular price data is intentionally not published because it "could

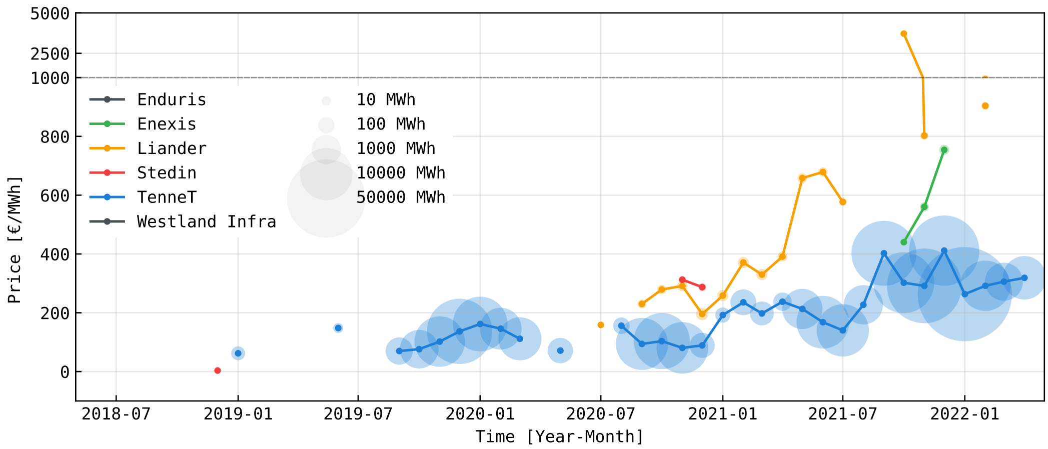



FIGURE 3.11: The monthly mean spread price over time paid by each of the 6 Dutch DSOs participating in GOPACS. In the top part of the graph, the scale is changed to include the very high price paid by Liander without distorting the main content.

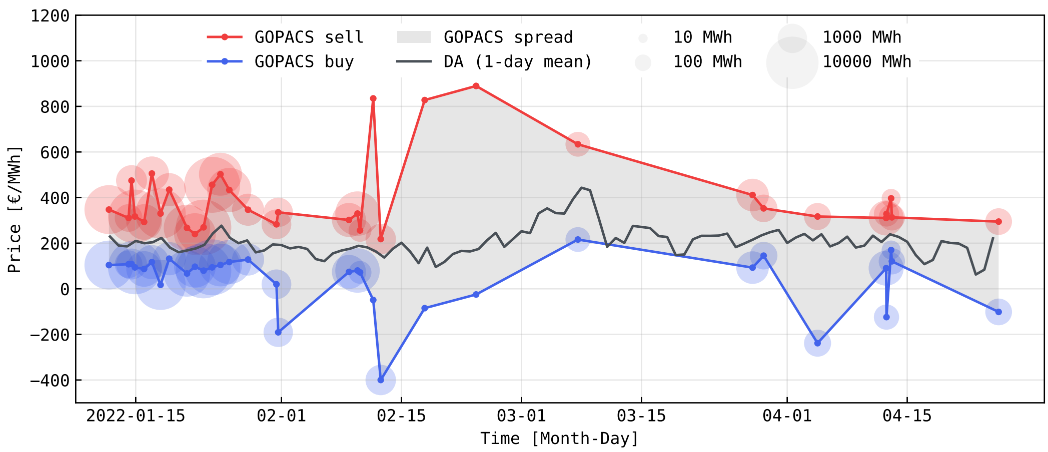



FIGURE 3.12: The scraped GOPACS buy, sell and spread data. The daily average DA price is given as a reference. Each dot represents new data published in the GOPACS EX data source, which may correspond to a whole auction. However, this data source is scraped and there is no official documentation so a given dot may also be only part of an auction or multiple auctions. The value itself is the mean price, given at the time the update was published in the EX data source, which may be different from the time the auction was held or the problem period took place. Again, the size of the circle indicates the volume.

potentially reveal competition-sensitive information" [14], according to TenneT in response to our email asking about the availability of more extensive price data. There is however a workaround that allowed us to get more granular GOPACS price data for a limited period.

We have created a Python script that downloads the latest PM data from https://idcons.nl/public/businessperformance once per hour. This endpoint is updated regularly, we believe these updates to come in roughly after each auction, however, we are not certain of this since this behaviour is undocumented. In each update, the statistics on the current month's bid, ask and spread prices, and volume are incremented to reflect the most up-to-date information. We read the data from this endpoint in regular intervals, saving each such snapshot. Then we take the difference between any snapshot and its preceding snapshot. These differences represent the per-auction statistics. We have been running a script on a dedicated server performing the task of scraping this endpoint every hour set up with a cron job.

The data obtained this way is shown in Figure 3.12 on the previous page. We see that daily GOPACS prices seem to be more volatile and result in higher differences than monthly mean prices which is also to be expected. It also seems that in particular, some lower volume auctions can lead to very extreme prices ranging from -400 €/MWh to 1000 €/MWh. Additionally, the daily or per-auction GOPACS prices seem more detached from DA prices. The Pearson correlation coefficient between daily GOPACS sell prices and daily DA prices is 0.017, between GOPACS buy prices and DA it is 0.430, a weaker correlation as was the case for the monthly prices.

## 3.3 GOPACS prediction

In Section 3.2 on page 20 we have visualised all the GOPACS data available. In this section, we will go a bit more depth into the relation between GOPACS data sources and various other related data.

In total, we collected 148 other features made up of various data of the Dutch and German electricity grids, most of it derived from ENTSO-E[10] and TenneT[11]. We cleaned and resampled all of this data into 15-minute buckets (1 PTU), then analysed their relation to GOPACS data. For the GOPACS data, we focused on four features: (1) The MA TP volume, (2) The CB volume, and (3) the state of GOPACS, given by a binary variable indicating if a MA TP problem period is taking place and the (4) same thing for the CB problem periods. We did not look into price data as not enough granular data is available at this point. We do expect more data, specifically price data, to be made available by the DSOs, as doing so may attract more participants of GOPACS which in turn may increase liquidity and potentially lower the spread that needs to be paid.

In the first step, we compute the Pearson correlation coefficient between the 148 features and the four selected GOPACS features. We report the ten highest (in terms of absolute value) correlated features. In all cases, we give the feature name with the correlation in parentheses. The data is shown in Table 3.2 and Table 3.3 on the following page

TABLE 3.2: The Pearson correlation coefficient between GOPACS MA TP Volume (first column) and GOPACS CB Volume (second column) and various features. The features are ranked according to their absolute correlation. The period of investigation is from January 2020 to May 2022.

| | MA TP Vol | CB Vol |
|---|---:|---:|
| 1 | RM Sub 15 Min (0.402) | DA Price (0.472) |
| 2 | RM 15 to 30 Min (0.367) | Available Capacity (-0.465) |
| 3 | ID Price (0.346) | BPL Neg 100 (0.448) |
| 4 | aFRR Price Up (0.338) | ID Price (0.447) |
| 5 | BPL Pos 100 (0.333) | BPL Neg Min (0.438) |
| 6 | RM 30 Min to 2 Hours (0.320) | BPL Pos 100 (0.415) |
| 7 | Balance Price Mid (0.309) | BPL Pos 300 (0.405) |
| 8 | DA Price (0.297) | Balance Price Mid (0.394) |
| 9 | BPL Neg Min (0.295) | aFRR Price Up (0.390) |
| 10 | BPL Pos Min (0.293) | BPL Neg 300 (0.379) |

[10] https://www.entsoe.eu
[11] https://www.tennet.org/english/operational_management/export_data_explanation.aspx

TABLE 3.3: The Pearson correlation coefficient between GOPACS MA TP problem period (PP) data (binary time series where 1 means there is a problem period at a given time and 0 means no problem period) (first column) and GOPACS CB PP data (second column) and various features. The features are ranked according to their absolute correlation. The period of investigation is from January 2020 to May 2022.

| | MA TP PP | CB PP |
|---|---|---|
| 1 | CB PP (0.622) | MA TP PP (0.622) |
| 2 | ID Price (0.283) | Load NL (0.268) |
| 3 | Load NL (0.279) | Measured exchange (0.250) |
| 4 | Load DA DE (0.277) | Load DA DE (0.248) |
| 5 | Load DE (0.269) | Load DE (0.245) |
| 6 | Measured exchange (0.263) | Gen All DA NL (0.220) |
| 7 | BPL Pos Min (0.251) | ID Price (0.220) |
| 8 | Balance Price Mid (0.245) | Gen All Other NL (0.220) |
| 9 | BPL Pos 100 (0.241) | Gen All Other DE (0.198) |
| 10 | Gen All DA NL (0.239) | Load DA NL (0.189) |

Table 3.2 on the previous page and Table 3.3 give the Pearson correlation coefficient between various features a specific GOPACS feature. For the volume features, we find that the highest correlations are found for Regulating Margin[12] (RM), Bid Price Ladder[13] (BPL), automated Frequency Restoration Reserve (aFRR), Balance[14], DA and ID prices and available capacity[15]. For the binary problem period features, we find the highest correlations for the Dutch (NL) and German (DE) grid load, generation and generation forecasts[16], the exchange between the Netherlands and Germany, BPL, ID and Balance prices.

Most of these features are related to ancillary services, in particular their prices, as well as prices of related markets. In addition, we see also some relation to the general state of the grid, that is, the load, generation and generation forecasts. The high correlation of DA price, ID price and other price data could be an indication that GOPACS auctions may specifically take place when other electricity prices are high and liquidity is thus low. However, as correlation does not imply causation, it may also be the case that there are simply volatile periods in all electricity prices, and all markets are equally affected and thus they are correlated. In order to investigate these relations in more depth, we have attempted a prediction of GOPACS problem periods and a regression on GOPACS volumes.

Specifically, we use gradient-boosted decision trees from LightGBM[36]. Again we trained the model on all of the external features from January 2020 to May 2022. These features represent the independent variables, and for the dependent variable, we chose the same four GOPACS features as used in the correlation studies. For the binary features, MA TP PP and CB PP, we used classification, for the continuous features, MA TP Vol and CB Vol, we used regression.

[12] https://www.tennet.org/english/operational_management/system_data_preparation/Reported_production_capacity/Regulating_margin.aspx

[13] https://www.tennet.org/english/operational_management/system_data_preparation/offering_regulating_reserve_capacity/bid_price_ladder.aspx

[14] https://www.tennet.org/english/operational_management/System_data_relating_implementation/system_balance_information/index.aspx

[15] https://www.tennet.org/english/operational_management/system_data_preparation/Reported_production_capacity/Available_capacity.aspx

[16] The acronym 'DA' in e.g. 'Gen All DA NL' or 'Load DA DE' refers to this feature being a forecast of generation, specifically the forecast made for a particular day on the day before, hence the day-ahead (DA) forecast. 'Other' in e.g. Gen All Other DE means the generation of all types of power plants except for onshore and offshore wind energy

We develop four models for these four features. Other than the choice of model, the setup we use is very similar in all cases.

For the classification task, we have two classes, 1 meaning there is a GOPACS problem period and 0 meaning no GOPACS problem period. GOPACS takes only place rarely, hence class 1 is the minority class. We undersample the majority class such that both classes are equally distributed throughout our dataset. For the regression task, we predict the continuous volume. Here we restrict the data set only to those hours where there is volume, as the focus is on predicting the volume level, and not whether there is or is not any problem period happening. For all four models, we validate our results with k-fold cross-validation, using k=6, running each model on six distinct subsets of the total data. We use a relatively simple model setup, specifically, we chose GBDT with 2000 trees of 30 leaves and a learning rate of 0.002. The loss function of the classifier is binary log-loss, for the regression model we use the least squared error (l2). For the evaluation of results, we use accuracy (ACC) and area-under-the-curve (AUC) for the classifier and the mean squared error (MSE) and mean absolute error (MAE) for the regression model. Visually, we analyse the model performance on a specific out-of-sample test period that is excluded from the dataset the model trained on. We show the out-of-sample model performance in four figures.

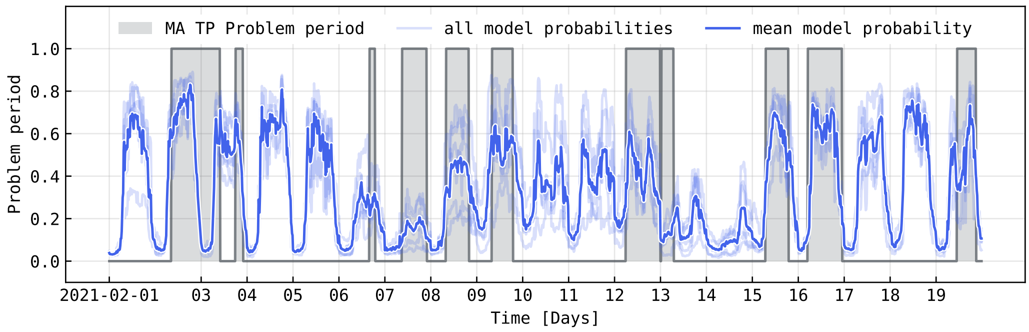



FIGURE 3.13: All classification model probabilities and their mean probability against the actual MA TP problem period, on unseen test data of October 2021

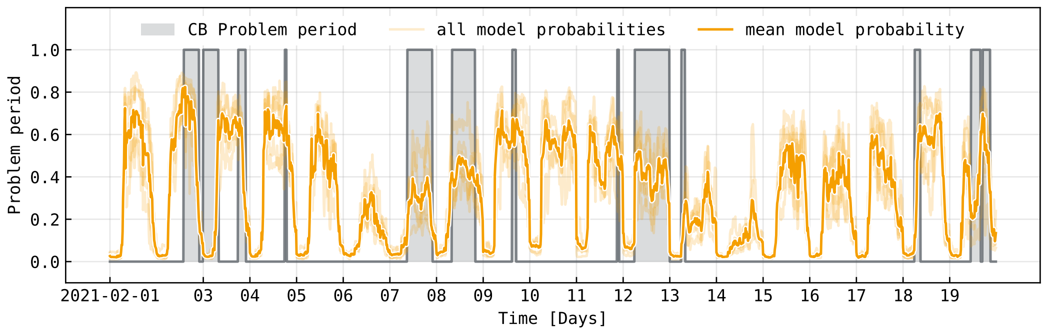



FIGURE 3.14: All classification model probabilities and their mean probability against the actual CB problem period, on unseen test data of October 2021

Across all of the test data, the MA TP PP classifier achieves a mean ACC of 0.739 and a mean AUC of 0.841. The most important features were the buy price for the manual frequency restoration (mFRR Price Up), the offshore wind generation forecast of Germany (Gen Wind-Off DA DE) and the available capacity in the Dutch grid (Available Capacity). The CB PP classifier achieves a mean ACC of 0.735 and a mean AUC of 0.859. Here the most important features were the lowest price on the regulating and reserve capacity bid price ladder (BPL Neg

Max) and the already encountered mFRR Price Up and Gen WindOff DA DE. The LGBM feature importance measures the number of times a feature is used in the trees that constitute the model. We refer to Section A on page 90 for the complete list of feature importances.

Figure 3.13 on the preceding page and Figure 3.14 on the previous page show the performance of the classification models on the unseen data period. We plot the output probabilities of the models. In both cases, we see that the model probabilities have significant daily seasonality. This seasonality also seems to align somewhat with the GOPACS auctions themselves. However other than this seasonality, the models do not present qualitatively good results. There are various instances where the models have false positives and false negatives.

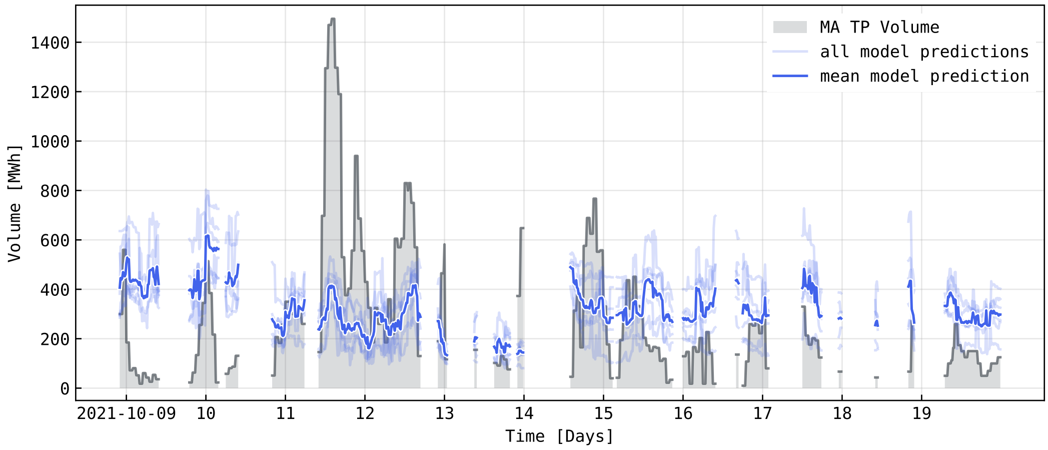



FIGURE 3.15: All regression model volume predictions and their mean volume against the actual MA TP volume, on unseen test data of February 2021

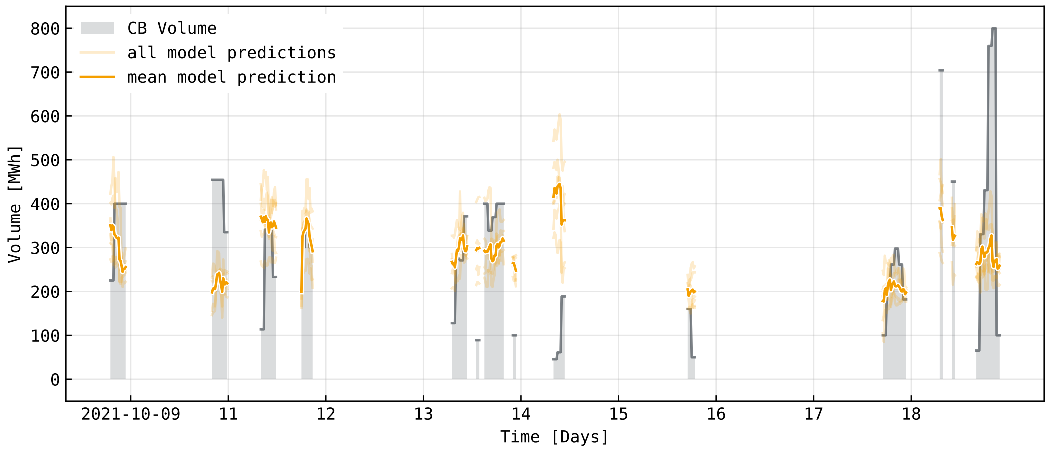



FIGURE 3.16: All regression model volume predictions and their mean volume against the actual CB volume, on unseen test data of February 2021

Across all of the test data, the MA TP Vol regression model achieves an MSE of 57,146 (69,216) and an MAE of 141 (174). In parentheses we give the performance of a naive benchmark, predicting the volume of the test data based on the mean volume of the training data. The most important features were the forecasted 'net position'[17] (Net Pos DA), the generation forecast of all but wind energy in the Netherlands (Gen All Other DA NL) and the offshore wind

[17] https://transparency.entsoe.eu/content/static_content/Static%20content/knowledge%20base/data-views/transmission-domain/Data-view%20Implicit%20Allocations%20-%20Net%20position.html

electricity generation forecast in Germany (Gen WindOff DA DE). The CB Vol regression model achieves an MSE of 35,642 (55144) and an MAE of 137 (186). Here the most important features were the DA price, the net position forecast (Net Pos DA) and the offshore wind electricity generation in the Netherlands (Gen WindOff NL).

Figure 3.15 on the preceding page and Figure 3.16 on the previous page show the performance of the regression models on the unseen data period. The volume predictions of the model are plotted. In the case of the MA TP volume, the model does seem to be able to capture the shape of the volume that is announced, such as is e.g. the case around the 12th, 18th and 20th of October. This is somewhat to be expected, as the MA themselves are forecasts made by a DSO that are most likely at least partially based on similar, albeit more granular data. However the model has trouble predicting the exact volume level, and in total the naive benchmark model has a similar performance. For the CB volume data, the model performance is similar. In some cases, it captures the shape of the volume, e.g. the 19th of October, but the scale is still far off and again the performance of the benchmark is quite similar.

In general, we see that there are indeed some features that are related to GOPACS, either having a high Pearson correlation or having high importance in predicting GOPACS. However, there is no single feature that is predominant in predicting either GOPACS auctions or the volume of those auctions, to a degree worthy of being included in a GOPACS model. Moreover, even if there were such a feature, this feature itself would need to be modelled, and we arrive at the initial situation. For these reasons, we take a stochastic approach to modelling GOPACS, based on various statistical properties of the GOPACS data that we explore in the next section.

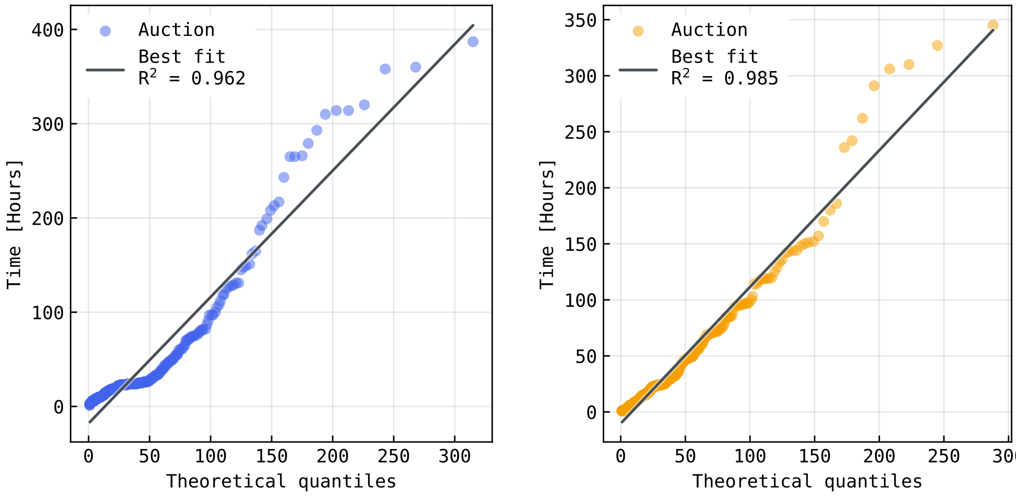
## 3.4 GOPACS modeling

In this section, we will explore the statistical properties of the GOPACS data and how these can be used in the construction of a model that simulates GOPACS. We focus this investigation on three main aspects, namely arrival-departure times, volume and prices.

### 3.4.1 Arrival-departure times

There are a variety of different timings in GOPACS auctions, as discovered in Section 3.2 on page 20. However, the model we construct has to simplify these time dynamics, focusing on those aspects that matter most when we want to simulate GOPACS. For this reason, we focus on the interarrival and auction duration times. These two quantities define both how long auctions last and how many auctions we can expect in a given time frame, thus combined they define 'how much GOPACS' there is in a given time frame.

After limited success with predicting these times, we have decided to use a stochastic approach. In a completely stochastic context, it would make sense to assume that during every hour, or every day, there is some constant probability of a congestion problem appearing and thus a GOPACS auction for this problem period taking place. Such a process would be very similar to a Bernoulli process. Now if we assume that GOPACS follows such a Bernoulli process, then the arrival times and auction durations should be geometrically distributed, as the geometric distribution gives the number of Bernoulli trials needed to have a success (i.e. the number of failures until the first success) [39]. In our case, we define a success as either the start of a problem period (arrival) or the end of a problem period (departure). For this reason, we investigate how good of a fit a geometric distribution is.

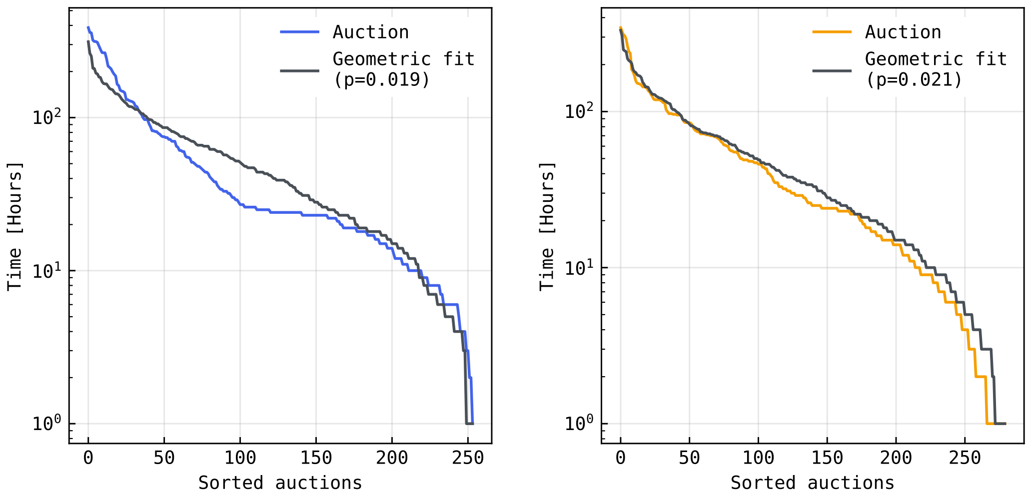



(A) MA TP interarrival times

(B) CB interarrival times

FIGURE 3.17: Probability plot of the MA TP and CB interarrival times against the best fit of a geometric distribution.

Figure 3.17 shows the arrival times for both CB and TP volumes[18]. We display them with a probability plot, which shows the theoretical quantiles on the $x$-axis and the volume of the ordered values on the $y$-axis. This plot thus allows comparing how well data matches a theoretical distribution. A perfect fit would be a straight line. We see on Figure 3.17a that the fit for the TP arrivals works relatively well, however on Figure 3.17b the fit is better. However in general we get a good fit.



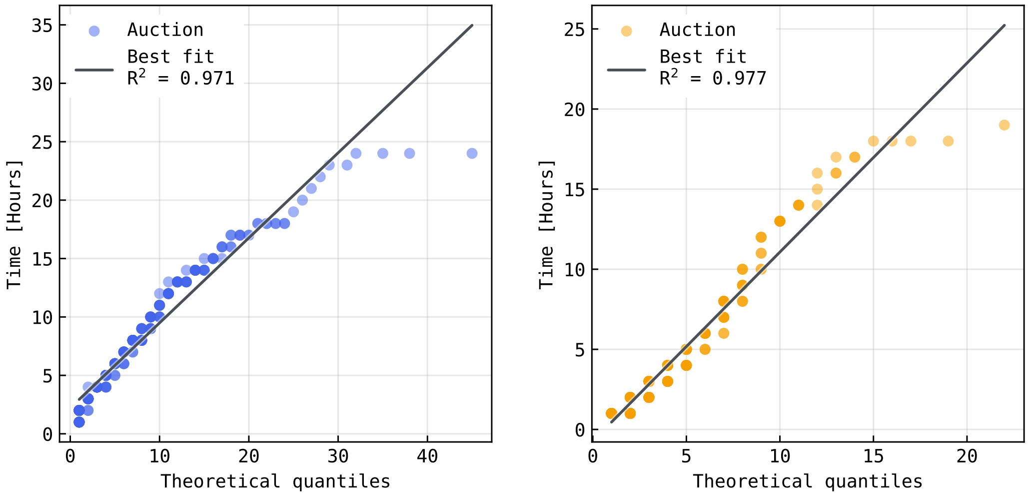

(A) MA TP interarrival times

(B) CB interarrival times

FIGURE 3.18: The sorted interarrival times for both MA TP and CB. The times between each two auctions are sorted and then plotted. Also shown are random values sampled from a geometrical distribution using the probability of success $p$ indicated in the figure

Figure 3.18 shows another point of view of the geometric distribution fitted to the GOPACS arrivals for both TP and CB auctions. In both subfigures, the best fitting arrival rate parameter $p$ is given in the legend of the plot.

[18] we will explain in Section 3.4.2 on page 37 why we model these times only for CB and MA TP volumes

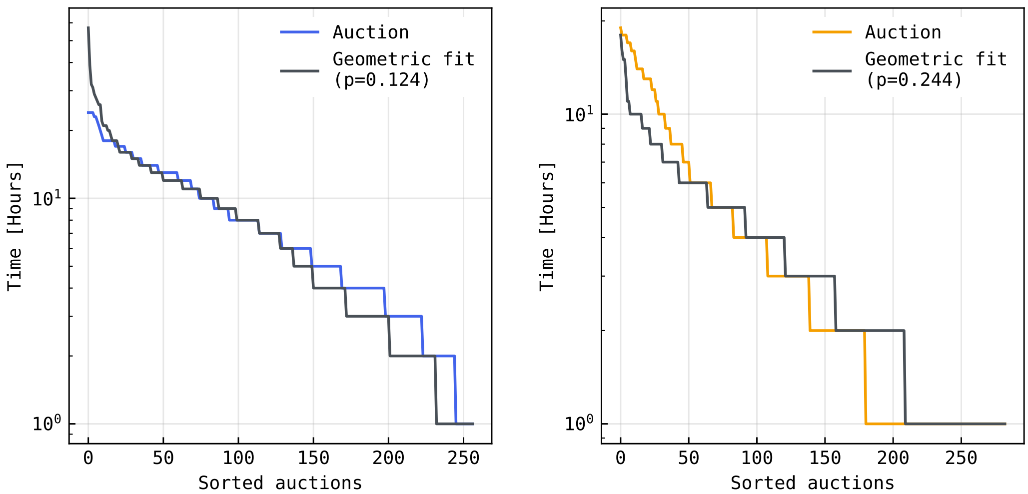



(A) MA TP auction duration times

(B) CB auction duration times

FIGURE 3.19: Probability plot of the MA TP and CB auction duration times against the best fit of a geometric distribution. Due to the discrete nature of the duration times, multiple entries may be shown by a single dot.

We performed the same exploration for the GOPACS departure times. Shown in Figure 3.19 are probability plots for the departure times, again comparing the order values to the theoretical quantiles of a geometric distribution. Again, we find a slightly mismatched fit for the MA TP auction times and a better fit for the CB auction times.




(A) MA TP auction duration times

(B) CB auction duration times

FIGURE 3.20: The sorted auction duration times for both MA TP and CB. The times of each auctions are sorted and then plotted. Also shown are random values sampled from a geometrical distribution using the probability of success $p$ indicated in the figure

Figure 3.20 gives again another view of the GOPACS by showing side by side both the real departure times and a geometric distribution with some probability of departure $p$. Finally, we have explored various of these quantities of interest per Dutch province. These results are shown in Table 3.4 on the following page. Having modelled both the arrival and departure times, we will move on to model the volumes.

TABLE 3.4: Various quantities of interest for each of the 12 Dutch provinces

| location | BUY | SELL | buy-sell ratio | arrival time [HH:MM] | auction duration [HH:MM] | $p_a$ | $p_d$ |
|---|---|---|---|---|---|---|---|
| Drenthe | 343 | 142 | 0.707 | 50:58 | 8:10 | 0.020 | 0.122 |
| Flevoland | 288 | 320 | 0.474 | 58:23 | 7:52 | 0.017 | 0.127 |
| Friesland | 294 | 142 | 0.674 | 61:22 | 8:05 | 0.016 | 0.124 |
| Gelderland | 6 | 405 | 0.015 | 69:06 | 7:31 | 0.014 | 0.133 |
| Groningen | 349 | 143 | 0.709 | 69:06 | 7:31 | 0.014 | 0.133 |
| Limburg | 8 | 407 | 0.019 | 69:06 | 7:31 | 0.014 | 0.133 |
| Noord-Brabant | 23 | 399 | 0.055 | 69:06 | 7:31 | 0.014 | 0.133 |
| Noord-Holland | 22 | 481 | 0.044 | 69:06 | 7:31 | 0.014 | 0.133 |
| Overijssel | 242 | 199 | 0.549 | 69:06 | 7:31 | 0.014 | 0.133 |
| Utrecht | 20 | 476 | 0.040 | 69:06 | 7:31 | 0.014 | 0.133 |
| Zeeland | 152 | 322 | 0.321 | 70:00 | 7:39 | 0.014 | 0.130 |
| Zuid-Holland | 22 | 464 | 0.045 | 70:00 | 7:39 | 0.014 | 0.130 |

### 3.4.2 Volume

Like the time data, there is also a lot of volume data. There are two sources, MA and CB, with MA containing three types of data. First, for the MA, we will only focus on TP announcements. We do not use FT data as it is unstructured and outdated as most of them were sent two years ago, since then GOPACS has evolved a lot. They are also not actively used anymore in practice. On the other hand, CM announcements have been issued recently and are also well structured. However, we have also decided against using them. The reason is that they are mostly focused on very specific locational problems such as the Neerijnen issue mentioned in Section 3.2.1 on page 21. Such congestion issues are simply not relevant for our work which aims to simulate the country-wide auctions. In addition, CM announcements are also less stochastic and thus less interesting from a modelling perspective, and finally, the volumes of energy in them are much smaller than those in typical TP announcements.

This leaves us with two data sources, namely MA TP and CB. Figures such as Figure 3.1 on page 21 convey the impression that CB data is very similar to TP data (and to some extent, FT). However this is not necessarily the case, and when investigating them in detail, various discrepancies are found.

Figure 3.21 on the next page shows the 2-week mean volume for both types of volume investigated. They only rarely line up, often there are major differences between them. Their Pearson correlation coefficients are 0.311 (1-day mean), 0.400 (2-week mean), 0.678 (1-month mean) and 0.911 (3-month mean), so in general, it seems that they share a longer trend, however, vary quite a lot on short term. Some of these differences are to be expected: TP announcements are issued before an auction and contain the forecasts of the volumes that are estimated to be needed. Then the auction happens, and some quantity of volume is cleared. A few hours later, the actual congestion problem occurs and the DSO requests volume from auction participants. Finally, once the problem is over, CB data is published on the volumes that were actually transacted or cleared. We have however also found various instances of hours during which there was recorded TP data but not CB data or vice versa, as an example we refer to Figure 3.15 on page 33 and Figure 3.16 on page 33 which both show the same period, but very different volumes.

Figure 3.22 on the following page shows a scatter plot of the CB volumes plotted against the MA volumes. For this plot, we only took those hours during which both an entry in the MA

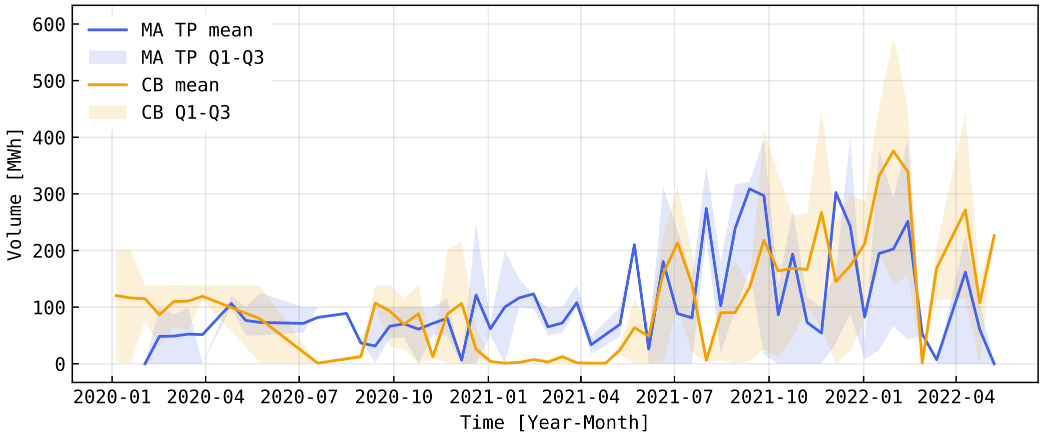
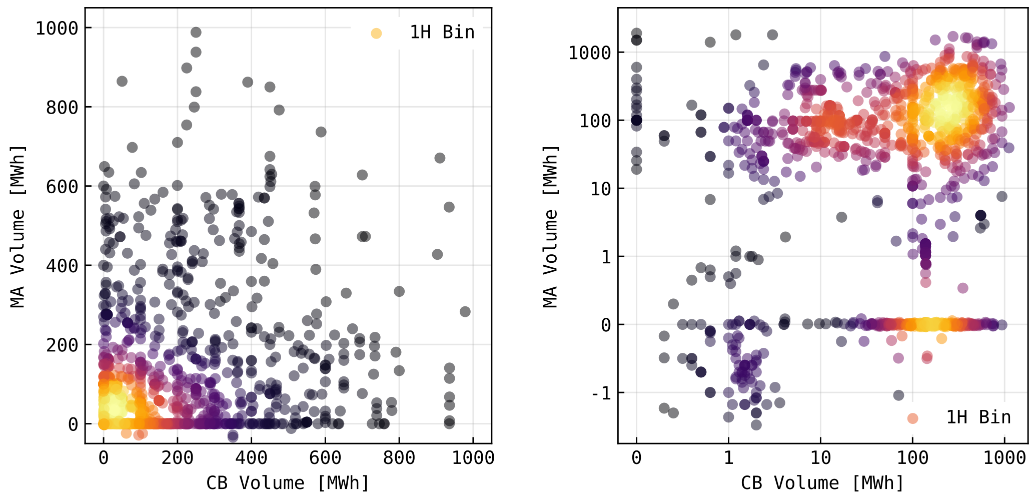



FIGURE 3.21: The 2-week mean volume over time for both MA TP and CB data, shown by the solid line. The colored area represents the difference between the 1st and 3rd quantile of the volume distribution of each month.

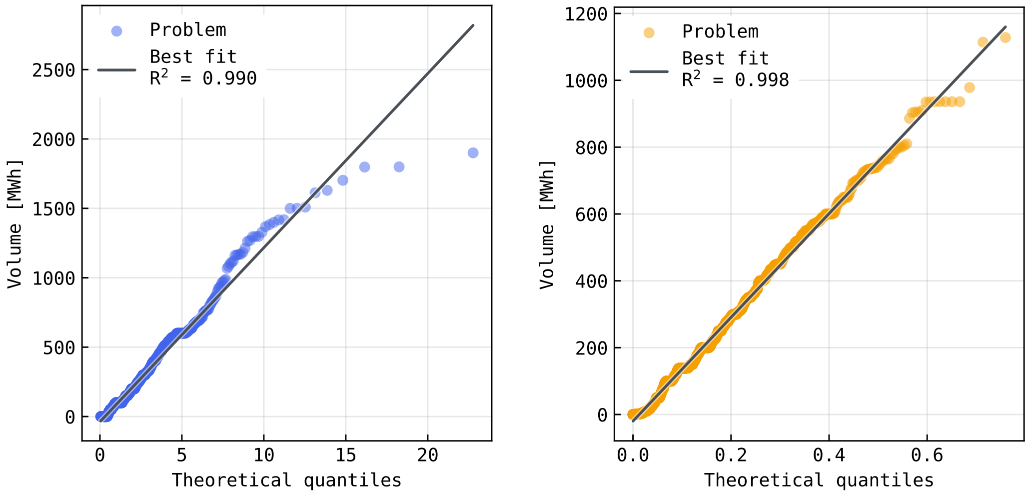



(A) Linear scale (B) Logarithmic scale

FIGURE 3.22: The CB volumes plotted against the MA TP volumes. Each dot represents the relation between the mean CB volume and mean MA TP volume of all hours between 2019 and now. Only hours are shown during which there is both a value for MA TP volume and CB volume.

TP data and CB data exist. The correlation between these two vectors is 0.187. There is a strong relation between them, neither in linear mode nor in logarithmic mode, however at least the dense region in the upper right corner of Figure 3.22b seems to indicate that a high announced volume mostly translates to a high cleared volume, although there are also many cases in which no volume is announced, but some is cleared, or inversely, where volume is announced but none is cleared.

We took these discrepancies into account when modelling the volume. In general, modelling the volume has proven to be a difficult task. Figure 3.21, Figure 3.8 on page 27 and Figure 3.6 on page 25 all showed that there are some (seasonal) trends in the volume, in particular since GOPACS is a new market and still evolving. However, we were unable to find a model that could predict or explain these trends reliably. For this reason, we also attempt a stochastic approach to model the volumes.




(A) MA TP volume against lognorm distribution (B) CB volume against beta distribution

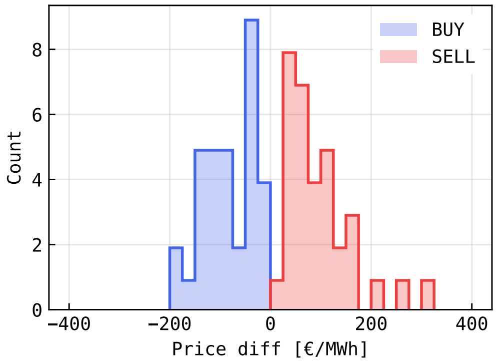
FIGURE 3.23: Probability plot of the MA TP and CB auction duration times against the best fit of either a log-normal or beta distribution.

Figure 3.23 shows two probability plots for the two-volume distributions. In Figure 3.23a we plot the TP problems and their fit to a log-normal distribution. The fit is not ideal, however, the log-normal distribution seems to capture at least some of the MA volume features we care most about, such as both the skewness and excess kurtosis of the values. A better fit is achieved with the beta distribution in Figure 3.23b. Before we used the beta distribution, we had tried both an exponential fit, which failed to capture the long tail of the CB problem distribution, and a power-law fit, which captured the long tail element but did not fit the exponential slope. A beta distribution seemed to be the ideal choice. It is again debatable whether the choice of beta distribution can be realistically argued, however for the purpose of simulating CB it does seem to be an acceptable choice. In both cases, the choice of distribution is more or less arbitrary and there is not really much justification we can give for why we would expect the volume to follow a beta distribution or log-normal distribution.

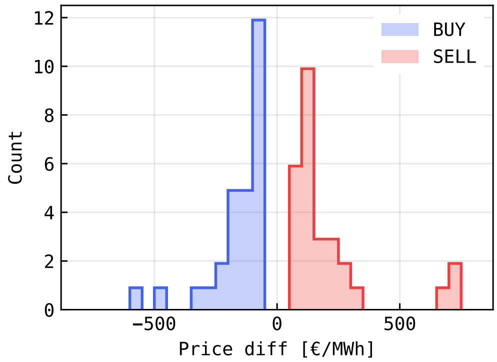
### 3.4.3 Prices

The last aspect of GOPACS we want to explore are the prices. One of the main difficulties with the GOPACS prices are that there is not much data available. As shown in Section 3.2.2 on page 27, prices are only available in monthly aggregated sums, with the sole exception being the scraped prices that we have from the 11th of January 2022 until now. We have also seen in Figure 3.10 on page 28 that GOPACS prices are correlated to DA prices. For this reason, we are mostly interested in the price differences, which we compute by subtracting the DA prices from GOPACS prices.

Figure 3.24 on the following page shows the distribution of both monthly aggregated GOPACS price differences and daily aggregated GOPACS price differences. The monthly prices in Figure 3.24a on the next page resemble more a normal distribution, which can be somewhat expected due to the central limit theorem. Other than that, we see that both buy and sell price differences are mostly distributed in the interval between 0 and 200 €/MWh away from 0, which in this case means they deviate roughly 0 to 200 €/MWh from DA prices. The means are -82 and 99. For the daily prices in Figure 3.24b on the following page we see a bit more

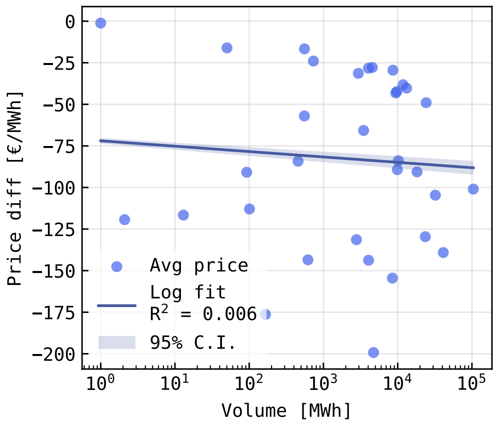



(A) Monthly mean price differences

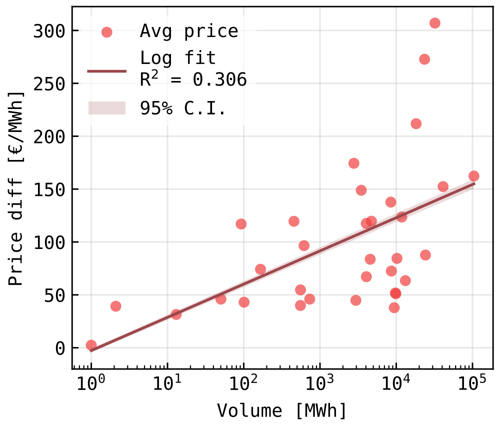



(B) Scraped auction price differences

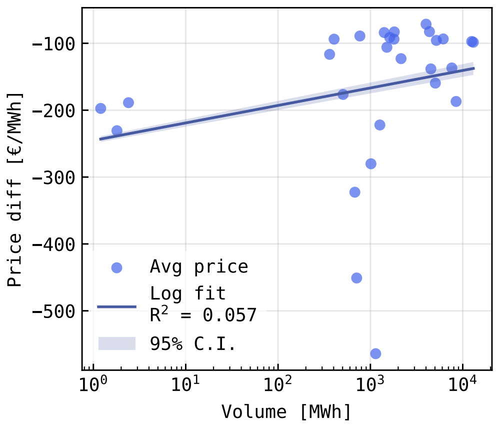
FIGURE 3.24: Histogram of both the monthly mean buy and sell prices differences and the scraped auction price differences. In all cases, the price difference is shown, which is computed by subtracting the DA price from the GOPACS price.



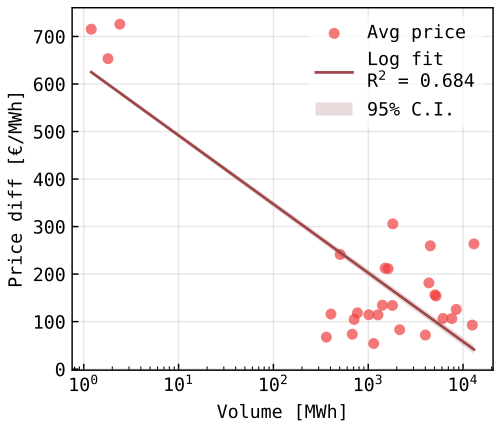

(A) Monthly mean buy price differences




(B) Monthly mean sell price differences




(C) Scraped auction buy price differences




(D) Scraped auction sell price differences

FIGURE 3.25: Regression between the volume in logarithmic space and the price difference in linear space. Each dot represents either the relation between a monthly mean price and the monthly total volume (top row), respectively between a scraped auction price and the associated total auction volume (bottom row).

of an exponential or heavy-tailed distribution. The means are -166 and 203, roughly twice as large as for the monthly prices.

In both cases, the small sample size of available data in addition to price trends make it hard to fit some distribution to these prices. In particular, a distribution would need to generate prices for each auction, however, the best granularity we have at our disposal for a potential validation are the very few scraped auction prices. For this reason, we seek to find other ways to model these prices. One idea is to see if there is a relation between GOPACS price and the volume announced or cleared during the auctions.

In Figure 3.25 on the previous page we explore the relation between GOPACS volume and price. The volume used is sourced from EX and PM data sources, however, we have found that is proportional to the CB volume data instead of the MA volume data, which makes sense as the price is also given as cleared. Moreover, in all plots, we used a logarithmic scale for the volume as values are spread out a lot. For the mean price differences in the top row, there is not really any trend in Figure 3.25a on the preceding page and an upwards trend in Figure 3.25b on the previous page. However, we believe that for the mean price differences, even when we de-trend them using DA prices, they are still taken over a period that is too long to ignore outside effects. For example, much of the lower volume auctions took place in the early days of GOPACS, while most of the higher volume auctions took place more recently, and the GOPACS market has certainly evolved in the years between them. We believe this development of the market itself to play a bigger role in the relation between volumes and prices than the volume itself. Worded differently, we believe that the higher price for the higher volume observable in Figure 3.25b on the preceding page, is, at least partially, because these higher volume months happened a few months ago, at a time were GOPACS was already more established, and not necessarily because higher volume equals higher price per volume.

However on the other hand, for the lower row, the time span of the data is a much shorter interval, and for this reason, we believe that any of the trends from the top row are not present in this data. Moreover, this data is also the prices per auction, and not only the monthly aggregated mean. We see that there is an upwards trend for the buy price differences in Figure 3.25c on the previous page and a downwards trend for the sell price differences in Figure 3.25d on the preceding page. This seems to suggest that as auction volume increases, the price difference paid per volume decreases. Another way to look at this is that auctions, in which the volume that is announced is very low, face liquidity issues as not many participants want to participate, and thus higher price differences are achieved. However, due to the lack of available data, it is difficult to take a decisive conclusion.

# Chapter 4

# Methods

## 4.1 GOPACS model

We have developed a model to simulate the various aspects of GOPACS. This model to simulate GOPACS is required as the base for the RO approach that we employ in order to quantify the value of participation in GOPACS. GOPACS has various aspects that can be modelled. In Section 3.4 on page 34 we narrowed GOPACS down to three main aspects: The arrivals and departures, the volume, and the prices, ignoring other aspects such as the geographical properties of auctions. We combine these three aspects in a model based on a stochastic finite difference equation that describes how GOPACS prices behave over time.

### 4.1.1 Arrival–Departure process

The first aspect of GOPACS that we simulate are the interarrival times between auctions and the duration of auctions. We call them the arrival-departure process as it specifies the arrivals and departures of GOPACS auctions. In Section 3.4.1 on page 34 we found that both for arrival and departure times of GOPACS a stochastic approach can be considered where we model the two aspects as Bernoulli processes.

To make this approach more formal, we define two possible states $X$ of GOPACS. State 0 ($X = 0$) means there is no auction, and state 1 ($X = 1$) means there is a GOPACS auction. In our model, GOPACS is always in one of these two states. If $X = 0$, meaning no auction is currently ongoing, then a GOPACS arrival may happen with a probability of arrival $p_a$ and we transition to $X = 1$. Similarly, if $X = 1$, a departure may happen with a probability of departure $p_d$ and we transition to $X = 0$. We can represent these state transitions with a Markov Chain in Figure 4.1.

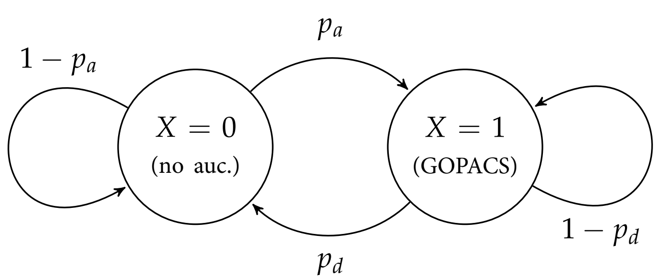



FIGURE 4.1: Markov chain transition probability graph

At every time step, we thus have two Bernoulli trials $\mathcal{B}(p_a)$ and $\mathcal{B}(p_d)$, defined as follows:

$$\mathcal{B}(p) = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{else} \end{cases} \tag{4.1}$$

We sample binary random values $\alpha \sim \mathcal{B}(p_a)$ and $\beta \sim \mathcal{B}(p_d)$ from these Bernoulli trials that indicate whether we change the state or not. Simulating this sampling procedure over time (discretised time steps), the sequence of $\alpha$ and $\beta$ random variables are two Bernoulli processes [19]. We can construct the complete arrival–departure subprocess of GOPACS with a finite difference equation:

$$\Delta X = (1 - X)\alpha - X\beta \tag{4.2}$$

In Equation (4.2), $\Delta X$ can take the values 1 or 0 if $X = 0$, and the values $-1$ or 0 if $X = 1$. Since we want to model this process over time, we denote the state $X$ at time $t$ by $X_t$, and we write the process:

$$X_{t+1} = X_t + \Delta X \tag{4.3}$$

Equation (4.3) shows how we iteratively add $\Delta X$ to $X$, which represents the GOPACS arrival–departure process.

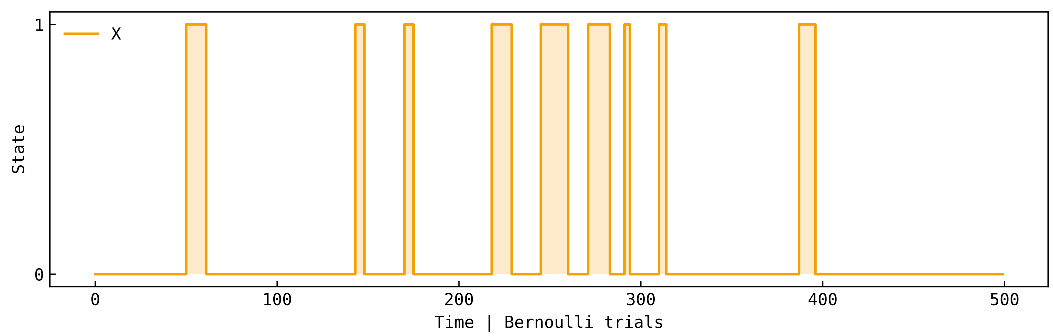



FIGURE 4.2: A sample output of simulating Equation (4.2). The state $X$ can be either 0 or 1, and the durations between state changes are samples from the arrival and departure distributions.

Figure Figure 4.2 shows the simulated arrival–departure process over a set number of time steps and with $p_a = 0.014$ and $p_d = 0.133$. These particular values were chosen as they are the empirical values found for most Dutch provinces as seen in Table 3.1 on page 20.

### 4.1.2 Volume and Price

Every GOPACS announcement comes with details on the volume that is needed to resolve the congestion issue, a bidding period and the locations in which buy or sell orders are asked for. The bidding process is sealed-bid, meaning that one has no insight into the bids of other participants, and pay-as-bid, meaning that each participant gets paid what they bid if their bid got accepted, irrespective of what others bid. Accepted bids may or may not be called. At the time of the expected congestion issue, the bids are called and the electricity must be delivered or taken delivery of. The cleared volume may differ from the previously announced volume (see Section 3.4.2 on page 37). When the bidding process is over, no data is published other than the cleared volume and times of the congestion problem period. Once a month, the mean price achieved in all auctions of that month is published. No data is available on the distribution and prices of bids, the only data participants have at their disposal is thus whether their own bids got accepted or not.

This description of the volume and price aspects of GOPACS highlights their complexity. Any modelling of either volume or price is complicated by the quality (e.g. no link between MA and CB data, which complicates establishing their relation) and quantity of the data available (e.g. only mean prices available publicly, and even through scraping we only get the mean auction price, but we have no idea about the price for a single hour or PTU).

For these reasons, we take a rather pragmatic approach to modelling both price and volume. We develop two price processes that each rely on different assumptions. We keep the first price process relatively simple, focusing on what is most important, namely the prices. The second price process is more complicated as it includes volume, with the aim of giving a more accurate view of prices achieved in an auction. Both processes are explained in depth in their relative sections, Section 4.1.3 and Section 4.1.4 on the following page.

Before we move on to these processes, we will reiterate some considerations taken in both cases. First, as discussed in Section 3.1.3 on page 18, each GOPACS auction contains a buy and sell location or area. We assume that a given region is only part of one type of area, as otherwise the congestion issue in that region could be resolved without the need for an auction. Hence for a given area, we model the auction in that area as either a buy or sell auction. In a buy auction, participants submit the prices for which they would buy energy, which are typically lower than DA prices. In sell auctions, on the other hand, the DSO is looking for sellers of energy, which typically ask for prices higher than DA prices to sell their energy.

Next, we have shown in Section 3.4.3 on page 39 that it is not unreasonable to assume that GOPACS prices are deviations from DA prices. As mentioned in Section 2.1.2 on page 5, it can be assumed that all other markets are merely deviations from the DA to correct potential forecasting mistakes. We believe that the same is true for GOPACS. For this reason, we are actually more interested in modelling the price difference between GOPACS and DA than the GOPACS prices themselves. The main benefit is that we can ignore most daily, weekly or annual trends inherent to GOPACS as these will also be inherent to DA, and hence not captured by the difference between DA and GOPACS. There may still be some trends in GOPACS however, for which we will use other approaches to handle them. Another benefit is that buy auction price differences are now always negative and sell auction price differences are now always positive. Having stated these assumptions, we will now go more in-depth about the two price processes used.

### 4.1.3 Price process 1

In the first price process (PP1), we simulate one single price[1] per GOPACS auction that represents both the mean and the max price difference of the bidding process. This single mean or max price may be a positive value in case it is a GOPACS sell auction or a negative[2] value in case it is a GOPACS buy auction.

This price represents the mean price of the auction as our model is based on the data we have available, which are the mean auction prices. It is however also the max price, as we assume that if we submit a bid into this auction, if the bid price is higher than this price, it is

[1] to clarify, although we call it 'price', it is technically the price difference between the GOPACS price and the DA price

[2] if the value is negative, then by 'max price' we actually mean the max absolute value price which is actually the minimum price

not accepted, so in other words, the price is the max price that we can bid for the bid to be accepted.

In PP1, we model this single price as being a random variable from some distribution. Due to the very few monthly and daily prices available to us (see Section 3.4.3 on page 39) we considered various potential distributions, each with different features, and investigate the resulting prices. The chosen distributions are the normal distribution, log-normal distribution and students-t distribution. With these three distributions, we can explore the effects of the first four moments (mean, variance, skew and kurtosis) on the prices. However careful consideration needs to be taken in the case a distribution with skew is used. If we use a distribution with skew for the positive sell prices and want to use the same distribution for the negative buy prices, we need to account for the direction of the skew. In practice, for most of the cases we investigated, we assume a positive skew for positive prices (i.e. long tail towards $\infty$) and a negative skew for negative prices.

The advantage of PP1 is that it allows us to focus on the GOPACS aspect we care most about, the prices, and gives us granular control over specifying various input distributions. Other aspects such as volume and the relation between announced and cleared volume (MA vs CB), are left out, 'abstracted away'. Moreover, this first process allows deriving an analytical solution of the GOPACS price. PP1 is well suited to give a good approximation of what prices one can expect in typical GOPACS auctions.

The main drawback of PP1 is that we do not take into account the auction element of GOPACS. GOPACS auctions are sealed-bid, so although we can estimate the mean price in an auction, estimating what price we can actually achieve by submitting bids is a different issue, which is not really addressed with PP1. Another drawback of PP1 is that it does not take into account volume. While we mentioned that ignoring volume can be a benefit as we leave out additional complexity that could increase model error, it also comes with several problems. One such problem is that PP1 can give a potential good approximation of the price achievable with small volumes, however, may fail at doing so when we participate with large volume bids in the bidding process. Finally, one last drawback of PP1 is that it also ignores self-cannibalisation. Self-cannibalisation is the effect of submitting a lot of bids into an auction such that the distribution of bids of the auction changes and hence the mean and max clearing prices of the auction differ as well. A practical example is an auction where 50 MWh of volume is auctioned. Assuming that the highest clearing price is 200 €/MWh (before we participate in it) and that all submitted bids are distributed more or less uniformly between e.g. 100 €/MWh and 300 €/MWh, it is certainly unrealistic to assume that if we bid a 50 MWh bid of 180 €/MWh, this bid will be completely cleared. On the opposite, it is much more likely that only a small part of this bid is cleared until the total 50 MWh of volume is reached unless no one else has submitted a bid below 180 €/MWh, which is unrealistic if we assume other people's bids to be more or less uniformly distributed. With such a large bid, we may 'eat into our own profits', hence the term self-cannibalisation. In order to address these drawbacks, we have developed the price process 2.

### 4.1.4 Price process 2

The second price subprocess (PP2) is more complex than PP1. Instead of simulating a single price value per auction, we simulate a fixed number of bids $n_o$ that are uniformly distributed between some minimum and maximum value. For the buy auctions, the minimum value is $o_{\min,b}$ and maximum is $o_{\max,b}$, for the sell auctions we use $o_{\min,s}$ and $o_{\max,s}$. Each bid consists of a size in MWh (typically we use 1 MWh) and a price in €/MWh. We then sample the cleared volume $V$ for each auction as a random variable from the beta distribution with parameters estimated in Section 3.4.2 on page 37. We use this volume as the cutoff for the bids. We perform this cutoff by sorting all submitted bids, those of other participants and our own, according to their price, then accepting enough bids such that the volume of bids is equal to the volume that is to be cleared. The last accepted bid becomes the max price, and we can also determine the mean and min prices from the resulting distribution.

PP2 has various advantages. First, it is a more accurate representation of how GOPACS actually works, as we simulate the auction process to the best of our understanding. Secondly, PP2 includes the volume as well, which allows us to include one more of the GOPACS aspects. Moreover, by simulating the actual bids, PP2 gives a more accurate approximation as the actual price depends on the volume and prices submitted to the auction. In particular, this allows us to correctly take into account the self-cannibalisation issue.

There are however also a fair amount of disadvantages to PP2. One of the main drawbacks is that PP2 introduces many more assumptions and parameters, in particular relating to the volume and bid distribution. For example, it uses only the cleared volume, instead of modelling both the announced and cleared volume. This choice was made because we were unable to determine the relation between the two types of volume (see Section 3.4.2 on page 37). Moreover, we model various aspects of the bidding process, such as the number and size of other bids, and their distribution, which we can not verify or validate as no data is published on them. Even data on related markets is very sparse and no detailed information is available. All of these factors can potentially increase the model error, as there are simply more uncertainties in the whole process. Another drawback of PP2 is that we do not have an analytical solution for it, although there may be one but it is beyond the scope of this work. Moreover, the GOPACS price difference in PP2 is emerging from the model and choice of other parameters and is not directly sampled from some distribution as is done in PP1. While this may be a more correct approach to modelling the situation, it also complicates the reproduction of GOPACS dynamics as these now depend on the choice of other parameters. Finally, PP2 is also more computationally intensive: The complexity of simulating $N$ hours with PP1 is $\mathcal{O}(N)$, growing linearly with the number of hours, as we only need to sample a random variable which is done in constant time. However the complexity of simulating $N$ hours with PP2 is $\mathcal{O}(N \cdot (n_o + n_u)\log(n_o + n_u))$, with $n_o$ the number of other participants bids and $n_u$ the number of our bids, and assuming an efficient ($\mathcal{O}(n \log n)$) sorting algorithm like heap sort is used.

### 4.1.5 Combining the processes

By combining the arrival–departure process and price process(es) we arrive at the combined GOPACS model. This model is again represented by a finite-difference equation. We use $\Delta S$ to denote the price difference in a similar way that we previously used $\Delta X$ to represent the state difference in Equation (4.2) on page 43. Moreover, given that we simulate the GOPACS difference from DA, $S = 0$ is used to represent that no GOPACS auction is currently going on, or alternatively interpreted, that the GOPACS price is equal to the DA price as the difference is 0, hence there is no need to participate in GOPACS. In the case of PP1, we represent the random variable that we sample from the price distribution with $\epsilon$[3]. In the case of PP2, the finite difference equation presented below can not really be used, however, the simulation process is relatively similar as we will see later.

$$\Delta S = \epsilon I(S = 0)\alpha - S\beta \tag{4.4}$$

[3]the difference between $\epsilon$ and $S$ is that $\epsilon$ is always either a positive or negative deviation representing the random value sampled from the price process, while $S$ is the actual GOPACS price difference that also includes the arrival-departure process and thus may take the value 0 depending on whether there is an auction going on or not

with $I(x)$ the indicator function conditional on some binary value $x$ as follows:

$$I(x) = \begin{cases} 1 & \text{if } x \text{ is True} \\ 0 & \text{else} \end{cases} \tag{4.5}$$

Equation (4.4) on the preceding page shows how the value of $S$ can change. If $S = 0$, i.e. the situation in which no GOPACS auction is currently ongoing, then $I(S = 0) = 1$. In this situation, the minuend of Equation (4.8) on page 51 may be the random variable $\epsilon$ when there is an auction arrival ($\alpha = 1$), while the subtrahend will be 0 due to $S = 0$ irrespective of whether there is a departure. Hence when there is no GOPACS auction at time step $t$, $\Delta S$ is equal to 0 with probability $1 - p_a$ or $\epsilon$ with probability $p_a$, and thus $S_{t+1}$ will also be equal to either 0 or $\epsilon$[4] as shown in Equation (4.3) on page 43.

Similarly, in the case where an auction is ongoing, $S \neq 0$, $I(S = 0) = 0$ and hence the minuend of Equation (4.4) on the previous page is necessarily 0. However the subtrahend may take the values $S$ or 0 depending on whether there is a departure $\beta = 1$ or not. Hence when there is a GOPACS auction at time step $t$, $\Delta S$ is equal to 0 with probability $1 - p_d$ or $-S_t$ with probability $p_d$. When $\Delta S = -S_t$, it becomes apparent via Equation (4.3) on page 43 that $S_{t+1} = S_t - S_t = 0$.

A showcase of Equation (4.4) on the previous page simulated over multiple time steps is given in Figure 4.3.

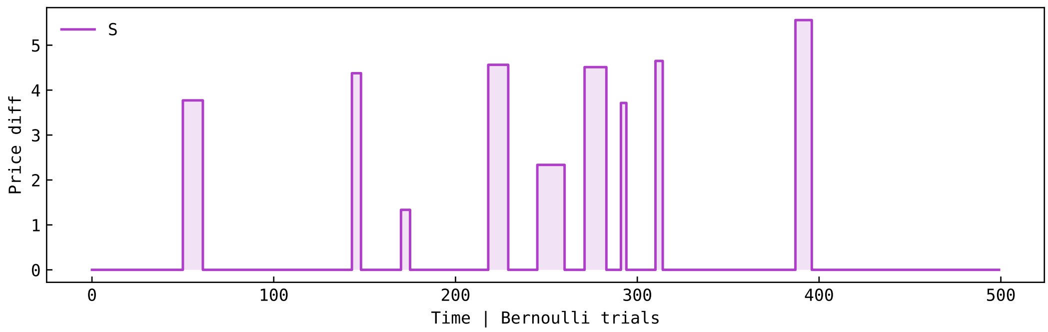



FIGURE 4.3: A sample output of simulating Equation (4.4) on the previous page. The price difference $S$ that is modeled is either 0 or a random value from the price distribution. We also draw attention that the durations between state changes are random values sampled from the arrival–departure process, and they are the same values as used in Figure 4.2 on page 43 as the same random seed has been used for both figures.

One aspect worth highlighting is that, even though combined in Equation (4.4) on the previous page, the arrival-departure subprocess and price subprocess are still independent processes. In practice, this means that parameter changes of one process have no effect on the other process. This independence is intended as it simplifies the model and allows independent investigation of the effects that parameter changes have on the model output. In addition, the available price data is also insufficient in granularity to verify a relation between auction arrivals and the prices achieved in those auctions.

So far, the finite difference equation shown in Equation (4.4) on the preceding page is mostly applicable in the case where the price can be represented by a single value $\epsilon$, which remains the

[4] Note that $\epsilon$, being a random variable, may itself take the value 0 depending on the probability distribution used. However since we only use continuous probability distributions, the probability of $\epsilon$ being exactly 0 is negligible, and even if not, as long as it is significantly lower than $p_d$, will not affect the functioning of the model.

same for the whole problem period. In PP1 this is the case, but in PP2, the price changes for each hour as it depends on the volume which we assume to change for every hour. While the finite difference equation can not directly be used for PP2, the numerical simulation thereof is valid for both price processes:

**Algorithm 1:** Simplified GOPACS model discretisation

$S \leftarrow [N]$
**for** $i \leftarrow 0$ **to** $N-1$ **do**
    $\alpha \leftarrow \mathcal{B}(p_a)$
    $\beta \leftarrow \mathcal{B}(p_d)$
    **if** $((S_i = 0)$ **and** $(\alpha = 0))$ **or** $((S_i \neq 0)$ **and** $(\beta = 1))$ **then**
        $S_{(i+1)} \leftarrow 0$
    **else if** $((S_i \neq 0)$ **and** $(\beta = 0))$ **or** $((S_i = 0)$ **and** $(\alpha = 1))$ **then**
        $S_{(i+1)} \leftarrow \text{PP}$
    **end**
**end**

Algorithm 1 shows a simplified version of a discretised implementation of the GOPACS model from Equation (4.4) on page 46. $S$ is initialised as an empty (i.e. only containing only 0) array or list of size $N$, with $N$ being the number of discretised time steps we want to simulate GOPACS for. For every time step, we proceed to sample both the $\alpha$ and $\beta$ random variables as specified in Section 4.1.1 on page 42. Next, we check if either there is currently no auction going on $(S_i = 0)$ and no arrival $(\alpha = 0)$ or if there is an an auction going on $(S_i \neq 0)$ but also a departure $(\beta = 1)$ – if that is the case, then the next price value $S_{(i+1)}$ is 0, indicating that there is no auction at the next time step as the price difference is zero. In the opposite cases, namely when there is an auction going on $(S_i \neq 0)$ but no departure $(\beta = 0)$, or when there is no auction going on $(S_i = 0)$ but there is an arrival $(\alpha = 1)$, we generate a new price for the next iteration. The generation of this price depends on the price process. In the case of PP1, $S_{(i+1)} = \epsilon$ if there is an arrival or $S_{(i+1)} = S_i$ if we are already in an auction, in the case of PP2 it is more involved as we generate the price by simulating the bids of other participants, sampling the volume, then cutting of these bids and taking the mean.

TABLE 4.1: The general and arrival–departure model parameters

| Symbol | Parameter Description |
|---|---|
| $N$ | Number of hours to simulate |
| $p_a$ | Probability of arrival |
| $p_d$ | Probability of departure |
| $p_b$ | Probability of buy auction |

Finally, we present the parameters of the final model in three tables. Table 4.1 contains the general parameters inherent to all versions of the model. The only parameter not yet mentioned before is $p_b$ which is the probability of an auction being a buy or sell auction. A value close to 1 means that a new auction is very likely to be a buy auction, i.e. an auction where the DSO is looking for buyers, typically the case in northern Dutch provinces as seen in Figure 3.2a on page 22. A value close to 0 means a new auction is more likely to be a sell auction. The value for this parameter is region-specific and is listed as 'buy-sell ratio' in Table 3.1 on page 20.

TABLE 4.2: The PP1-specific model parameters

| Symbol | Parameter Description |
|---|---|
| $\mu_b$ | Mean of buy distribution |
| $\mu_s$ | Mean of sell distribution |
| $\sigma_b$ | Standard deviation of buy distribution |
| $\sigma_s$ | Standard deviation of sell distribution |
| $s_b$ | Shape parameter of buy distribution |
| $s_s$ | Shape parameter of sell distribution |

TABLE 4.3: The PP2-specific model parameters

| Symbol | Parameter Description |
|---|---|
| $V$ | Auction volume |
| $n_o$ | Number of other bids |
| $v_o$ | Volume / size of other bids |
| $o_{\min,b}$ | Other bid price minimum for buy auction |
| $o_{\max,b}$ | Other bid price maximum for buy auction |
| $o_{\min,s}$ | Other bid price minimum for buy auction |
| $o_{\max,s}$ | Other bid price maximum for sell auction |

Table 4.2 contains the PP1-specific parameters, which are mostly related to specifying the price distributions of buy and sell auctions. The shape parameter is used in certain cases, for example, it is used in some numerical implementations of the log-normal distribution or to represent the degrees of freedom in the students-t distribution. Lastly, Table 4.3 contains the PP2 specific parameters, which are defined in in more detail in Section 4.1.4 on page 45.

## 4.2 Real Option valuation

### 4.2.1 Flexible Assets in GOPACS

Having defined the GOPACS difference equation which provides us with a price process $S$ representing the underlying, we are able to build a RO pricing model on top of it. Concretely, our aim is to determine what value can be extracted from participation in GOPACS. Based on our understanding, there are three broad ways to trade on GOPACS: (1) DA arbitrage with storage, (2) ID arbitrage and (3) Flexible assets.

The first possibility would be to use energy storage assets that are then e.g. charged on DA prices and discharged in GOPACS auctions, allowing us to profit from price arbitrage between the two markets. However, this approach has drawbacks as batteries are expensive and not suited to store energy for an extended amount of time due to degradation, while other storage assets are not very common in the Netherlands. The second option would be to do directly arbitrage between GOPACS prices and ID prices. As the Dutch ID market, part of the Europe-wide ID market, is active 24/7, it should be possible to perform arbitrage between GOPACS and ID. However, the ID market is both more volatile and less liquid than DA, which may have an impact on the profit achievable with this arbitrage strategy. Next, it is also possible that this arbitrage between ID and GOPACS is already being exploited and hence offers no value. The lack of granular GOPACS price data prevents us from investigating this possibility in more depth. In general, we can not completely rule out these two options to participate in GOPACS, however, we have outlined why we believe them to be difficult to exploit. They could be a basis for future work and will be reiterated in Section 6.3 on page 83. The last approach is the one we will consider. We will value participation in GOPACS with a flexible asset (FA) that allows changing one's consumption pattern whenever it is financially profitable to do so. We identify four possible types of such FA:

1. When there is a GOPACS sell auction in our region (i.e. DSO needs electricity), we submit a sell bid (i.e. a bid of how much we want to be paid to sell energy to GOPACS).
   (a) **Produce more:** If the bid is accepted, produce (more) energy and feed it into GOPACS. If no GOPACS auction is ongoing, stop (additional) energy production. Examples are flexible power plants such as gas or coal power plants [44].
   (b) **Consume less:** Consume energy all the time and buy this energy on DA or some other regular market. If the bid is accepted, stop the consumption of energy and sell the 'right of consumption' on GOPACS. This is the typical load curtailment or load displacement [66], examples are e.g. commercial heating, ventilation and air-cooling (HVAC).
2. When there is a GOPACS buy auction in our region (i.e. DSO has a surplus of electricity), we submit a buy bid (i.e. a bid of how much we are willing to pay to buy energy from GOPACS).
   (a) **Consume more:** If the bid is accepted, consume (more) energy from GOPACS. If no GOPACS auction is ongoing, stop (additional) energy consumption.
   (b) **Produce less:** Produce energy all the time and sell this energy on DA or some other regular market. If the bid is accepted, stop the production of energy and sell the 'right of production' on GOPACS. This is again an example of supply-side flexibility which can be achieved with various flexible gas or coal power plants.

In all four cases, in order to extract value from participating in GOPACS with a FA, it is necessary to submit bids to the auction process. By submitting a bid to GOPACS, we are paid (or have to pay) for the energy that we deliver or receive. However, inherent to flexibility is also the optionality component which implies that we have the option, but not the obligation to participate in GOPACS. We model this optionality component with a RO approach, similar to the literature reviewed in Section 2.3 on page 9. Concretely, we will model the value of a FA in GOPACS as a lifetime strip of options, with one option per unit of electricity volume (1 MWh) and unit of time (1 hour), similar to the approaches of [17, 31] and others. In the following section, we will present our detailed option pricing approach.

### 4.2.2 Option type

Until now, we have referred to these options merely as 'option' without further specifying the type thereof. The option right, i.e. whether the option is a call or a put, depends on whether the auction is a sell or a buy auction.

Among the four cases outlined above, we will focus our investigation on cases (1b) and (2b), as these are the two strategies that directly on profiting from the relation between GOPACS and DA, whereas Case 1a relies on the relation between GOPACS and the fuel used in the power plant, and Case 2a relies on the relation between GOPACS and whatever goods or services are produced or offered with the energy sourced from GOPACS.

With the FA strategy that we pursue, we make the largest profit if there is a large difference between DA and GOPACS. For example in Case (1b) we sell energy usage contracts bought on DA in GOPACS sell auctions. The larger the difference between GOPACS and DA, or in other words, the larger the modelled GOPACS difference, the larger the profit we can make. This is similar to a spread call option on the difference between GOPACS and DA.

$$C_E = \max(0, S_{GO} - S_{DA} - K_s) \tag{4.6}$$

We have a similar situation in the case of a buy auction, in which the profit we can make depends on the difference between GOPACS and DA. However in this case it is a put option, as the profit increases the lower (negative) the GOPACS price difference is:

$$P_E = \max(0, K_b - S_{GO} + S_{DA}) \tag{4.7}$$

In both Equation (4.6) and Equation (4.7), $K$ is the strike price / cost of participation / running cost of the FA strategy. These costs include things like ramping up or down the production or consumption, or costs incurred by reducing the normal production or consumption in favour of participating in GOPACS. The strike price however does not contain any fixed costs that do not depend linearly on the number of times we participate in GOPACS, such as the investment cost of the FA. $S_{GO}$ is the GOPACS price (not difference), and $S_{DA}$ is they DA price. We outlined in Section 4.1.2 on page 43 the reasons for which we model the GOPACS price difference $S$ instead of the actual GOPACS price $S_{GO}$, that is:

$$S = S_{GO} - S_{DA} \tag{4.8}$$

Combining both Equation (4.8) with Equation (4.6), respectively with Equation (4.7), we can argue that one way to quantify the value of a FA is done with European options on the GOPACS price difference, in particular:

$$\begin{aligned} C_E &= \max(0, S_{GO} - S_{DA} - K_s) \\ &= \max(0, S - K_s) \end{aligned} \tag{4.9}$$

$$\begin{aligned} P_E &= \max(0, K_b - S_{GO} + S_{DA}) \\ &= \max(0, K_b - S) \end{aligned} \tag{4.10}$$

Several assumptions and simplifications have been made to be able to represent participation in GOPACS the European options given by Equation (4.9) and Equation (4.10), which we will defend now. As previously mentioned, both GOPACS and DA do not take place at the same time, so direct arbitrage is not possible. Instead, we suppose that we enter some position on DA, and then potentially change our position in the case it is profitable to do so during a GOPACS auction. In some of the literature reviewed in Section 2.3 on page 9, in particular, Tseng and Bartz [71], a similar approach is taken, but two main differences are apparent: First, they use two separate complex processes, one for energy and one for fuel (in our case, that would translate to two separate processes for GOPACS and DA). Second, they model the asset valuation as a unit-commitment / optimal control problem, considering ramp-up and ramp-down as well as other operational constraints of the asset they value.

Concerning the first point, we do not consider both GOPACS and DA as given by two separate complex processes, instead, we assume that they are roughly proportional to each other allowing us to only focus on modelling one stochastic process, which is their difference. One reason we do not consider them to be separate processes is that apart from short-term random-walk-like behaviour, and long-term trends, the DA model is relatively predictable and seasonal on a daily, weekly and annual basis [1]. On the other hand, for GOPACS, we do not know anything about the short-term behaviour and most of our modelling assumptions rely on educated guesses. For this reason, we make the choice of modelling both markets with the same process, or, from a different perspective, we focus our efforts on modelling the GOPACS difference only.

For the second point, there are a few reasons for which we do not apply a potentially more rigorous unit commitment or optimal control approach to our problem. First, GOPACS auctions take place, on average, around seven hours before the energy transfer that was bid on, as we found in Section 3.2.1 on page 22. These seven hours are long enough that ramp-up and ramp-down constraints potentially present should not pose an issue. On the other hand, the costs of ramping production or consumption up or down are considered by the strike price $K$. Another reason is that for the two FA cases we care about, Case (1b) and Case (2b), taking a position on the DA market is required before we even know if GOPACS will take place, and evidently also before we know whether we will participate in GOPACS or not. Moreover, the reason we consider this FA approach is also because this particular strategy allows participating in DA, even if no GOPACS takes place, without incurring any financial loss. What this means is that we do not need to consider any unit-commitment problem – instead, we simply participate in GOPACS if it is more profitable than not participating, or else we do not participate.

Related to this second point are also the reasons we chose a European option instead of an American option to model the FA strategy. Precisely because we take a position on DA by entering a contract that requires us to deliver or take delivery of electricity of some specific quantity at some specific time. Hence if we participate in GOPACS to exploit any price difference, this also applies only to this specific time of the contract we entered in DA before, which itself is thus best represented by a European option with a single exercise date. In particular, each option we model represents the option, but not the obligation to participate in GOPACS with a FA for some predefined unit of electricity and for a specific hour.

A similar European option valuation has also been used in the literature, see Section 2.3 on page 9, in particular [24, 71, 31, 66, 13]. However, in each of these cases, the markets in which they valued the assets with this approach have publicly available real-time prices. On the other hand, GOPACS is a sealed-bid auction, meaning that no price information is known until after the auction. Moreover, the payout is paid-as-bid, not paid-as-cleared, making the payout that some participant receives entirely dependent on their bidding strategy. The consequence is that the presented European-style options can only be used to either compute the theoretical average value achievable in GOPACS or under the assumption that we have a bidding strategy that allows us to achieve the mean price in each auction for all submitted bids. However in particular when submitting a large number of bids, it is highly unlikely to assume that such a strategy exists.

All things considered, a more realistic approach to valuing the optionality component of participating in GOPACS is by submitting bids that consist of a price $B$ and a volume which we consider to be always 1 MWh into the GOPACS auctions which are either accepted or not. The payoff is then fixed: for a sell auction the payoff is $B_s - K_s$ and for a buy auction the payoff is $K_b - B_b$, in both cases $|B| > |K|$ as it would not make sense to submit a bid lower than the cost of participation. In financial literature, options with a fixed payout are referred to as digital or binary options, and in this case we can define the digital call $C_D$ and put $P_D$ options as follows:

$$C_D = (B_s - K_s)I(\max(B_s, K_s) < S) \tag{4.11}$$

$$P_D = (K_b - B_b)I(\min(B_b, K_b) > S) \tag{4.12}$$

with $I(x)$ the indicator function defined in Equation (4.5) on page 47.

In both Equation (4.11) and Equation (4.12), the condition of the payout depends on whether the GOPACS difference $S$ is higher (respectively lower) than the submitted bid. The threshold of acceptance depends on the way we simulate the GOPACS price process. Using PP1, the threshold of whether a bid is accepted or not is the mean price of the GOPACS auction, $S$. If instead, we simulate the whole bid ladder as done in PP2, the threshold of whether a bid is accepted or not is computed by inserting the bid into the simulated bid ladder, sorting all the bids and using the cleared volume $V$ to determine the highest bid to be accepted. If the inserted bid is either the highest accepted bid or below, it is accepted, otherwise, it is not.

Finally, we want to draw attention to the fact that, while we talk about digital and European options, no actual options are written, traded or used in any way. Instead, we only use these theoretical options as a representation of the optionality value of the FA.

### 4.2.3 Discounting

One aspect of the option pricing approach that can not be ignored is to discount the value of the option to take into account the time value of money. Special considerations need to be made as we need to correctly discount back the profit made with a FA with respect to the various time points in the life of the FA as well as the choice of a discount rate.

First, we will use $t$ to represent the first hour at which the FA is able to be used to participate in GOPACS. In this case, this FA is already ready now or even already in use, $t = 0$. However in the more likely case that this FA is an investment that will take some time to be constructed, $t$ will be equal to the time between the allocation of the funds for the investment and the time the FA is ready to be used, taking into account the construction time in between, which may be significant in the case of a larger investment. Next, we will use $T$ to represent the end of the lifetime of the FA, or, in case we want to limit the valuation to a fixed horizon (e.g. 10 years), to represent that time. Finally, we will also use $\tau = T - t$.

The choice of discount rate or discounting approach, in general, is dependent on the application. Across the RO literature, the Datar–Mathews Method (DM Method) [45] is often encountered for discounting in RO. In the DM method, two different discount rates are used. The price of the underlying is discounted with the so-called hurdle rate or required rate of return in the target market, $R$, while the traditional risk-free discount rate $r$ is used for the strike price $K$ [45]. However after careful consideration of this method, we have decided against using it, as the typical RO examples it is used in are more abstract investment-decision RO, while in our case the options we use are much closer to actual financial options. Across the literature, various works in similar domains as ours, also use the more traditional single-rate discounting. In practice, this means that we will discount the options using only the annual risk-free discount rate $r$.

### 4.2.4 Combined option pricing model

Having specified all parts of the option pricing model, we can combine them to define the lifetime RO valuation of a FA [17, 13, 49].

First, we note that the option value of the FA is dependent on the actions we can take with it. It may be possible to use the same FA to both participate in GOPACS buy auctions and sell auctions, e.g. with a power plant that allows to both produce more (Case (1a)) and produce less (Case (1b)). In that case, we would model the value of the FA as the sum of both the respective put and call options. However, for the general case, we assume that only one strategy is pursued, and the value of the FA is given entirely by either a put or call option.

Next, one more thing to take care of is the volume of energy that we aim to trade with GOPACS. So far, we have assumed that when we submit a bid into GOPACS, this bid is of a fixed size of 1 MWh. In practice, however, the FA that we use may enable us however to submit more bids or bigger bids into GOPACS, and this should be correctly accounted for in its valuation. In PP1, we handle this situation by multiplying the value of a single option with the total volume we may be able to bid in GOPACS, then discounting this value with a cannibalisation factor that we estimate based on observations from related markets. This approach however has obvious flaws and was one of the main drivers behind the development of PP2. In PP2, we handle the situation by simulating the bidding process and performing a cutoff with the volume that will be cleared, as explained in depth in Section 4.1.4 on page 45. In general, however, we will adopt a functional notation to represent the volume of an option, for example, $P_D(25)$ represents a put option whose value is given by the submission of bids of a total volume of 25 MWh into a buy auction.

Combining the pieces, we arrive at the formula that we can use for the valuation of a FA:

$$A = \sum_{i=0}^{N} e^{-r(t+\frac{\tau i}{N})} C_{D,(t+\frac{\tau i}{N})}(v_u) \tag{4.13}$$

Equation (4.13) gives us the lifetime valuation $A$ of a FA based on the option strategy that we can pursue with it, which in this case is a call, but may also be a put depending on the type of FA. The specific option value depends, among other aspects, on the volume that we can trade with this FA in GOPACS, $v_u$, and the current time. The time of maturity of each option is given as subscript, so $C_{D,(t+\frac{\tau i}{N})}$ is a call option expiring at time $(t + \frac{\tau i}{N})$. Equation (4.13) is constructed in a way that the value of the FA is the sum of the values of a strip of options, one option for each hour during which the FA is active, starting at time $t$ and ending at time $T$. Finally, each such option is discounted back to today. There are various ways we can compute $A$. In all cases, we can use MC simulation to get a good approximation of $A$, and in some particular cases, we can derive an analytical value.

### 4.2.5 Monte Carlo Simulation

The previously defined lifetime value of the FA, $A$, is a random value which depends on the underlying GOPACS price difference random variable $S$. We can use Monte Carlo simulation to compute the expected value of $A$ by performing repeated sampling. Each sample $A_i$ consists of computing the value of the FA by performing a realisation of the GOPACS market. We use these samples to construct $\hat{A}_n$ which is the unbiased estimator of the expected FA value using $n$ samples:

$$\hat{A}_n = \frac{1}{n}\sum_{i=1}^{N} A_i \tag{4.14}$$

We note that $\hat{A}_n$ is an unbiased estimator[27], and we can use it to approximate the expected value of the FA, $\mathbb{E}(A)$ as follows:

$$\mathbb{E}(A) = \lim_{n\to\infty} \hat{A}_n \tag{4.15}$$

The advantage of using MC simulation to derive the value of the FA is that by doing this repeated sampling we also create a distribution of possible outcomes. The mean of this distribution will tend towards the expected value $\mathbb{E}(A)$, however, we can also compute other estimators from this distribution such as for example the variance. The variance of this distribution can give us insights into the range of values we can expect for the valuation of the FA, and gives us interest in the risks that we face. We will define how we can quantify these risks in Section 4.4 on page 59.

## 4.3 Analytical solution

Under certain conditions, there is an analytical solution to Equation (4.13) on the previous page. There is a closed-form solution to the arrival-departure process and also for PP1. These solutions give both the value of the European and digital style options. We also only have the European-style analytical solution in the case where the price random variable $\epsilon$ is sampled from a normal distribution.

### 4.3.1 Arrival-departure process

We recall that the arrival-departure process from Section 4.1.1 on page 42 can be represented by a Markov Chain (Figure 4.1 on page 42). We can alternatively represent this Markov chain by its transition probability matrix:

$$\mathcal{P} = \begin{bmatrix} 1-p_a & p_a \\ p_d & 1-p_d \end{bmatrix} \tag{4.16}$$

This matrix describes the probability to move from one state to another. For example let $p_{i,j} = \mathcal{P}$. Then $p_{0,0} = 1 - p_a$ is the probability to stay in state 0 if we currently are in state 0. In total we have two states in the Markov chain, state 0 (i.e. $X = 0$, 'no auction'), and state 1 (i.e. $X = 1$, 'GOPACS'). Next, we will verify if these states are recurrent and periodic, quoting El Gamal [19]:

**Theorem 1** *Let $A(i)$ be the set of states that are accessible from state i, i.e., can be reached from i in n steps, for some n. State i is said to be recurrent if starting from i, any accessible state j must be such that i is accessible from j, i.e., $j \in A(i) \iff i \in A(j)$. Clearly, this implies that if i is recurrent then it must be in $A(i)$.*

The Markov chain represented by $\mathcal{P}$ has only one recurrent class encompassing all states, if $p_a \in (0,1)$ and $p_d \in (0,1)$. When we are in state 0 we can access state 1 if $p_a > 0$, and vice versa, if we are in state 1, we can access state 0 if $p_d > 0$. Since both states are accessible from each other, they are in the same recurrent class, and since these are the only two states of the chain, we can say that all states are in the same recurrent class, and thus the Markov chain only has a single recurrent class [19].

**Theorem 2** *A recurrent class $A$ is called periodic if its states can be grouped into $d > 1$ disjoint subsets $S1, S2, ..., Sd$ with $\bigcup_{i=1}^{d} S_i = A$, such that all transitions from one subset lead to the next subset.*

Our Markov chain $\mathcal{P}$ is also aperiodic, since in the only recurrent class we have, as long as $p_a < 1$ and $p_d < 1$, there is no possibility to create two disjoint subsets where states from one subset must lead to the next subset. Finally, we can state the steady-state convergence theorem, again quoting El Gamal [19]:

**Theorem 3** *If a Markov chain has only one recurrent class and it is not periodic, then $r_{i,j}(n)$ tends to a steady-state $\pi_j$ independent of i, i.e.*

$$\lim_{n\to\infty} r_{i,j}(n) = \pi_j \ \forall i \tag{4.17}$$

Also from El Gamal [19], we note that $r_{i,j}(n)$ are the $n$-step transition probabilities, and constitute the elements of the matrix $R(n)$ which can be defined as follows

$$\begin{aligned} R(1) &= \mathcal{P} \\ R(2) &= R(1)\mathcal{P} = \mathcal{P}^2 \\ R(n) &= \mathcal{P}^n \end{aligned} \tag{4.18}$$

We note that steady-state probabilities $\pi_j$ are uniquely determined by the the balance equations and normalisation equation [19]. The balance equations are:

$$\pi_j = \sum_{k=0}^{(m-1)} \pi_k p_{k,j} \text{ for } j = 0, 1, , 2, ..., (m-1) \tag{4.19}$$

and the normalisation equation is:

$$\sum_{j=0}^{(m-1)} \pi_j = 1 \tag{4.20}$$

In both cases, $m$ is the number of states[5], which is 2 in our case. In our case, the system of equation is the following:

$$\begin{cases} \pi_0 = (1-p_a)\pi_0 + p_d\pi_1 \\ \pi_0 + \pi_1 = 1 \end{cases} \tag{4.21}$$

Solving this system of equations, we find the analytical solution for the steady states:

$$\pi_0 = \frac{p_d}{p_a + p_d} \quad \text{and} \quad \pi_1 = \frac{p_a}{p_a + p_d} \tag{4.22}$$

Finally, we can state this solution in the form of a matrix as well, using the notation from Equation (4.18) and we have:

[5]we use 0-indexing to represent the states, and hence state $m-1$ is the $m$-th state

$$\mathcal{P}^{\infty} = \begin{bmatrix} \pi_0 & \pi_1 \\ \pi_0 & \pi_1 \end{bmatrix} = \begin{bmatrix} \frac{p_d}{p_a+p_d} & \frac{p_a}{p_a+p_d} \\ \frac{p_d}{p_a+p_d} & \frac{p_a}{p_a+p_d} \end{bmatrix} \tag{4.23}$$

In practice, this means that if we simulate our GOPACS forever, the fraction of time spent in state 0 ('No auction') is $\pi_0$, and state 1 ('GOPACS') is $\pi_1$. This solution is also intuitive: For example, when $p_a = 0.2$, and $p_d = 0.4$, it means that leaving the GOPACS state is twice as likely (double the probability) then entering the GOPACS state, as there is a 20% probability of entering it when we are not in it, but 40% probability of leaving it if we are in it. If we do this enough times, it is intuitive that we will spend twice as much time in the 'No auction' state than in the 'GOPACS' state, and this is confirmed by the steady-state probabilities, which in this case are $\pi_0 = 0.66\bar{6}$ and $\pi_1 = 0.33\bar{3}$.

### 4.3.2 Option pricing solution

The analytical solution for the arrival–departure process is one of the ingredients we need for the closed solution of the option pricing problem. The other ingredients are the price process and the option itself. For the price process, we only have an analytical solution for PP1. There may be a solution for PP2, however, we have not derived it as it will most likely take considerably more time to derive due to the complexity of the process. For PP1, the analytical solution of the price is relatively easy: it is simply the mean of the price distribution. I.e., if the GOPACS price difference $S$ is computed with $\epsilon$ sampled from a normal distribution with given mean, $\epsilon \sim \mathcal{N}(\mu, \sigma)$, then if we sample enough values, the mean price will approach $\mu$ as stated by the law of large numbers.

The solution of the European and digital options is a combination of the option formula, the arrival–departure solution and the PP1 solution. Let $f(x; \mu, \sigma)$ be the probability density function (PDF) of some distribution with mean $\mu$ and variance $\sigma^2$, and let $\phi(x; 0, 1)$ be the PDF of the standard normal distribution, defined as:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x)^2} \tag{4.24}$$

Similarly, we define $F(x; \mu, \sigma)$ as the cumulative distribution function (CDF) of some distribution and $\Phi(x; 0, 1)$ as the CDF of the standard normal distribution, defined as:

$$\Phi(x) = \int_{-\infty}^{x} \phi(x) dx \tag{4.25}$$

Moreover, we recall that in the case of PP1, there are two price distributions, one for the buy ($b$) prices and one for the sell ($s$) prices. For the sell prices, we defined the PDF and CDF as $f_s(x, \mu, \sigma)$ and $F_s(x, \mu, \sigma)$, for the buy prices it is $f_b(x, \mu, \sigma)$ and $F_b(x, \mu, \sigma)$. We note however that when distributions with non-zero skew are used, it is important to verify the skewness has the correct sign. Particularly, if we care about long-tail extreme-value prices, we need to have a sell distribution with a positive skew and a buy distribution with a negative skew.

In the case of the European options, a call option expires in-the-money (ITM) if the price of the auction $S$ is above the strike price $K_s$, and a put option expires ITM if the price of the auction $S$ is below the strike price $K_b$. In the case of the digital options, a call option expires ITM if the price of the auction $S$ is above the bid price $B_s$ and consequently also above the strike price $K_s$. A digital put option expires ITM if the price of the auction $S$ is below the bid price $B_b$ and strike price $K_b$. In both cases, we need to compute the likelihood that $S$ is higher (resp. lower) than the strike (resp. bid) price. This probability is given by the CDF of the price distribution. For sell auctions, the probability of $S$ being higher than the bid $B_s$ is given by $1 - F_s(B_s)$ and higher than the strike $K_s$ is given by $1 - F_s(K_s)$. For buy auctions, the probability of $S$ being lower than the bid $B_b$ is given by $F_b(B_b)$ and lower than the strike $K_b$ is given by $F_b(K_b)$.

At this point, we are close to having the closed-form solution. What is still missing is to compute the payoff that we receive if the option is ITM. In the case of the digital option, this payoff is $B_s - K_s$ for sell auctions, $K_b - B_b$ for buy auctions, as defined in Equation (4.11) on page 53 and Equation (4.12) on page 53.

In the case of the European option, the payoff is a bit more complex. The mean payoff is the mean of the truncated distribution, truncated by the strike price $K$. When the price distribution is the normal distribution, there is a closed-form solution for this truncated normal distribution. For other distributions there may also be a closed-form solution, however, we have not investigated it. For the normal distribution, we have, quoting [11]:

$$\chi(\alpha, \beta) = \mu - \sigma \frac{\phi(\beta) - \phi(\alpha)}{\Phi(\beta) - \Phi(\alpha)} \tag{4.26}$$

With $\chi(\alpha, \beta)$ being the mean of the truncated distribution and $\alpha$ and $\beta$ defined as follows [11]:

$$\alpha = \frac{a - \mu}{\sigma} \quad \text{and} \quad \beta = \frac{b - \mu}{\sigma} \tag{4.27}$$

Equation (4.26) and Equation (4.27) give the formula for the mean of a normal distribution truncated on both sides, i.e. the mean of the area of the distribution that is between $a$ (lower end, towards $-\infty$) and $b$ (upper end, towards $\infty$). In the case of a sell auction, where the price difference $S$ is positive, $a = K_s$ and $b = \infty$. In the case of a buy auction, we have $a = -\infty$ and $b = K_b$.

At this point, the last thing we need to consider is the probability of a given auction being a buy or sell auction, given by $p_b$ as defined in Table 4.1 on page 48. This is easily taken care of by multiplying the payoff of a buy auction with $p_b$ and the payoff of a sell auction with $1 - p_b$.

We can finally express the closed form solution of the various options:

$$C_D = \pi_1 \cdot p_b \cdot (1 - F_s(\max(K_s, B_s); \mu_s, \sigma_s)) \cdot (B_s - K_s) \tag{4.28}$$

$$P_D = \pi_1 \cdot (1 - p_b) \cdot F_b(\min(K_b, B_b); \mu_b, \sigma_b) \cdot (K_b - B_b) \tag{4.29}$$

$$C_E = \pi_1 \cdot p_b \cdot (1 - F_s(K_s; \mu_s, \sigma_s)) \cdot \chi\left(\frac{K_s - \mu_s}{\sigma_s}, \infty\right) \tag{4.30}$$

$$P_E = \pi_1 \cdot (1 - p_b) \cdot F_b(K_b; \mu_b, \sigma_b) \cdot \chi\left(-\infty, \frac{K_b - \mu_b}{\sigma_b}\right) \tag{4.31}$$

Equation (4.28), Equation (4.29), Equation (4.30) and Equation (4.31) give the closed form solutions for the four options introduced in Section 4.2 on page 49. These closed form solutions do not depend on the price difference $S$. They are only valid in the case PP1 is used, and the solution of the European options also only applies if the price distribution itself is the normal distribution[6].

## 4.4 Risk quantification

One of the research questions we aim to answer is to quantify the risks and uncertainties faced when participating in GOPACS with a FA. In particular, we want to measure 'what can go wrong', or, put more formally, "the quantifiable likelihood of loss or less-than-expected returns" [46]. In order to quantify this risk, we make use of a risk measure, which "associates a financial position with loss L with a real number that measures the 'riskiness of L' " [46]. Instead of using a loss distribution as is commonly done in the literature and also referred to by this definition, we will consider what one may call a 'revenue' distribution, as we will only consider the upside of the RO strategy. In other words, we will ignore the initial investment cost and only look at the distribution of the random variable $A$, which due to the underlying options strategy that limits risk, is always above or equal to zero. We do not consider investment costs before we measure risk, as they can be simply subtracted afterwards and do not influence the risk measure itself. Moreover, the actual cost of the investment may depend on various factors that we do not know or that may differ for each situation, and including them does not add any benefit.

To quantify this outcome distribution, which we will represent by $A$, we can use the typical measures such as mean and variance. We also use two well-defined dedicated risk measures that are used in a variety of contexts, namely Value-at-Risk (VaR) and Expected Shortfall (ES)[7].

### 4.4.1 VaR Value-at-risk

Value-at-Risk (VaR) is a widely used risk measure, commonly encountered in the literature but also widely used by financial institutions [46, 51]. VaR measures the highest possible loss that can occur at a certain probability. In that sense, the VaR of a profit distribution (where profit is positive and loss is negative) is formally defined as, quoting [46].

$$\mathrm{VaR}_\alpha(A) = \inf\{x \in \mathbb{R} \mid P(A < x) \leq 1 - \alpha\} = \inf\{x \in \mathbb{R} \mid F_A(x) \geq \alpha\} \tag{4.32}$$

From 4.32 we can see that VaR is a quantile of the distribution, and we will mostly focus on VaR for $\alpha = 0.95$ or $\alpha = 0.99$, as is typically done [46]. The advantage of VaR is that the measure is relatively simple to compute and gives us insights into the "maximum loss that is not exceeded with a given high probability" [46], or in our case, the minimum profit that is not exceeded with a given high probability. However, the drawback of VaR is that it gives no insights into "the severity of losses that occur with a probability of less than $1 - \alpha$" [46].

### 4.4.2 Expected Shortfall

Expected Shortfall (ES) is an alternative risk measure that addresses the main drawback of VaR, that is, it gives us more distributional insights into the worst case. In other words, "ES can be interpreted as the expected loss that is incurred in the event that VaR is exceeded" [46]. The ES of a profit and loss distribution can be computed in various ways. In our case, given that our distribution $A$ is continuous, we can express it using VaR, quoting [58]:

[6] it may work for other distributions, but then one would need to have a different formula for $\chi$.
[7] ES is also sometimes referred to as Conditional Value-at-Risk (CVaR) or Average Value-at-Risk (AVaR)

$$\mathrm{ES}_\alpha(A) = \mathbb{E}(A|A < \mathrm{VaR}_\alpha(A)) \tag{4.33}$$

From the definition in Equation (4.33) we can see that ES is the "expected loss that is incurred in the event that VaR is exceeded", or in other words, ES is the average of all outcomes that are lower or equal to the VaR of the same $\alpha$. The advantages of ES over VaR become apparent when one computes both the VaR and ES of two distributions, with one of them having a longer tail, such as e.g. comparing the normal distribution and students-t distribution. In this specific case, the 95% VaR[8] of both distributions is relatively similar, while the 95% ES of both distributions is rather different [46]. We will use both VaR and ES to measure the riskiness of various experiments.

[8] We use 95% VaR as an alternative notation for $\mathrm{VaR}_{0.95}$

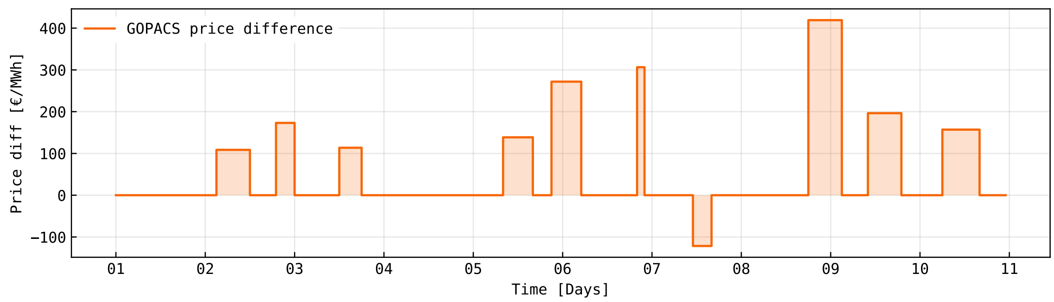
# Chapter 5

# Results

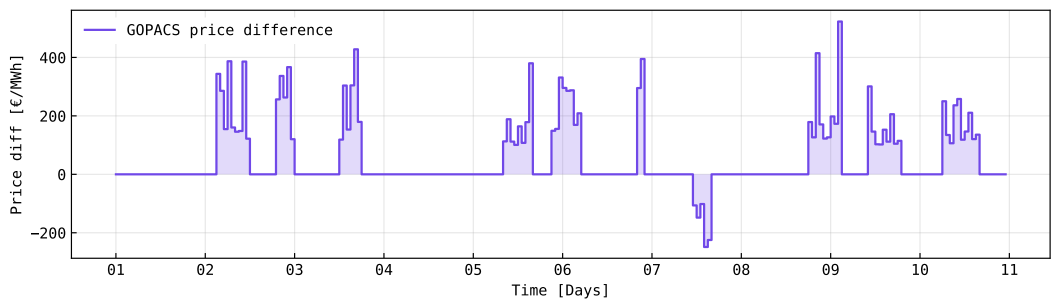
## 5.1 GOPACS model

### 5.1.1 Model overview

To start the results section, we will first take a look at the GOPACS model, leaving the option pricing aside for now. We simulate the model over a few days and show the output.




(A) using PP1




(B) using PP2

FIGURE 5.1: GOPACS price differences generated with both price processes. The x-axis represents time, with a granularity of one hour. The y-axis represents the price differences in €/MWh. 0 indicates that there is currently no auction.

In Figure 5.1 we see eleven days of simulated GOPACS auction price differences using PP1 and PP2. In the case of PP1 (Figure 5.1a), there is a single price per auction that is determined by a random variable. PP1 is a direct simulation of the GOPACS finite difference equation given by Equation (4.4) on page 46. In the case of this experiment, a normal distribution is used, the exact parameters for both the price and arrival–departure process are not of relevance as they were chosen at will in order to get a visually interesting figure. In later sections, we will go more in-depth in the parameter choice. In the figure, we see multiple blocks that appear and then disappear over time. Each block represents a problem period, however, we alternatively refer to it also as auction sometimes, although technically the auction takes place before the problem period. Each such block has a width, which represents the duration of the auction, and a height, which represents the price difference of GOPACS to DA in €/MWh. Between blocks are periods of zeros which indicate that the price difference is zero, meaning there is no GOPACS auction. The height of the blocks can be positive or negative. A positive value indicates the prices achieved during a sell auction, as the price of energy during a sell auction exceeds the DA price. A negative block represents a buy auction, here the DSO looking for buyers of energy typically fetches a price lower than DA for the electricity, and hence a negative price difference is shown. There are both buy and sell auctions, as the shown auction price process is simulated to be in some particular province/region in the Netherlands in which both buy and sell auctions take place.

In the case of PP2 (Figure 5.1b on the previous page), we see various similarities to PP1, but also differences. The differences are that the prices are computed differently, and we also have multiple prices per auction / problem period, in fact, we see one price per hour. This single price per hour is determined by a variety of factors, in particular, the volume and the bids of other parameters. Moreover, in both figures, the GOPACS price differences are shown. We can add the DA price back to these values to get the actual GOPACS prices.

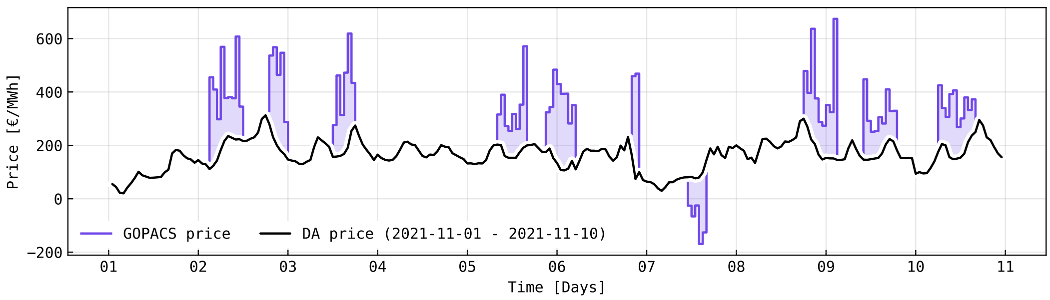



FIGURE 5.2: The GOPACS price shown as deviation from DA prices.

Figure 5.2 shows the GOPACS prices by adding the GOPACS price differences simulated with the GOPACS model (using PP2) and actual (non-simulated) DA prices. The resulting output differs from the outputs of previous figures. First, the GOPACS buy auction prices are now not necessarily negative anymore, instead they may just be prices below the current DA price. Additionally, the short-term trends and seasonality (e.g. price differences between day and night) of DA amplify the GOPACS prices, which we assume to be also the case in the real world. Next, while not visible in this plot, is that the long-term trends of DA prices (e.g. systematic price increase or decrease over multiple years) now are also integrated into GOPACS prices. This latter effect will be explored later in a bit more detail. Lastly, we want to draw attention that this addition of the DA prices influences the output of the $y$-axis, however, the arrival-departure process itself, determining the auction interarrival and duration times, is left unchanged. The blocks and spaces between them are still the same as in the previous figures.

### 5.1.2 Parameter sensitivity and non-stationary processes

In Section 4.1.5 on page 46, particularly in Table 4.1 on page 48, Table 4.2 on page 49 and Table 4.3 on page 49, we outlined the various parameters of our GOPACS model. Most of these parameters can either be input in the form of a single scalar value or with a vector of size

$N$, specifying the parameter value to be used for every single hour of the simulated period. When we input one parameter as a vector with different values, the process that this parameter contributes to will become a non-stationary process[1]. The usage of non-stationary parameters does invalidate some of the underlying statistical assumptions at the cost of simulating more realistic dynamics.

One example of how non-stationary processes could be used is shown in Figure 3.3b on page 23, where we see that the number of problem periods changes throughout the day, with there being more arrivals during the daytime than during nighttime. Instead of using a single arrival rate to model these problem period arrivals, a vector could be used with a seasonal arrival rate that specifies a different rate, or probability of arrival, depending on the time of the day.

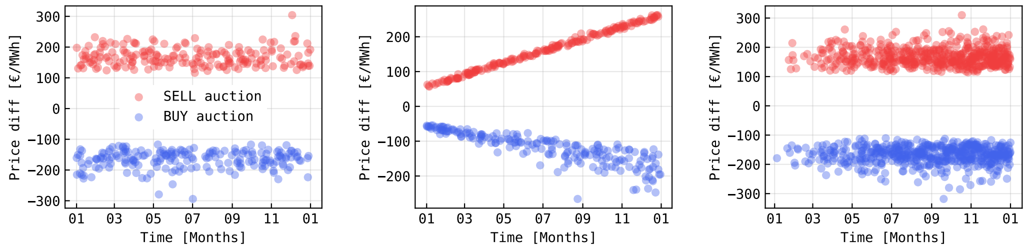
Another advantage of the non-stationary parameter inputs is that they allow to investigate of how the model output, in particular prices, react to changing parameters. Instead of plotting changes of the parameter on the $x$-axis and the model output on the $y$-axis, we can simply specify the input to the parameter as a non-stationary process changing over time, then plot the time on the $x$-axis and the model output on the $y$-axis. In this section, we show how the model output is affected by changes in parameters, with those changes being given by non-stationary parameters.

In general, the non-stationary parameters also allow to explore hypothetical scenarios, such as e.g. investigating what happens when more people start participating in GOPACS, that is, the number of bids from other participants increases. It is these experimental scenarios that we will look at more in-depth in Figure 5.3 on the following page.

Figure 5.3 on the next page shows the effects of various non-stationary processes on the resulting GOPACS prices. Figure 5.3a on the following page shows the default parameters with no non-stationary process. These default parameters are $N = 8760$ (one year), $p_a = 0.05$ (close to the real arrival rate determined from the data), $p_d = 1$, $p_b = 0.5$, $o_{\text{min},b} = -100$, $o_{\text{max},b} = -4000$, $o_{\text{min},s} = 100$, $o_{\text{max},b} = 4000$, $n_o = 200$, $v_o = 20$, $V = 100$. Most of these parameter choices are either somewhat arbitrary or modelled to best reproduce the original dynamics, an endeavour we will go more in-depth in Section 5.1.3 on page 65. We will however note two particularly interesting parameter choices. First, we set $p_d = 1$, which in practice means that every single auction only contains a single problem period of one hour and is then over. This was done to achieve a more clean scatter plot, where each dot is a single auction, which is thus more receptive to changing parameters. The other choice is $V = 100$. Typically we sample the volume as a random variable from some distribution, often one that is empirically estimated. However, in these experiments, we use a simple constant value to reduce the effect of stochasticity that volume has on the prices in order to better highlight the effects of the other parameters.

We investigate various scenarios in the other plots. In Figure 5.3b on the following page the minimum and maximum bid prices of other participants are increased (in absolute terms). This could for example be the case if GOPACS participants collectively increase their prices, or if participants start to ask for higher prices. Specifically, we have: $o_{\text{min},s}$ increasing from 50 to 250, $o_{\text{max},s}$ increasing from 500 to 1000, $o_{\text{min},b}$ increasing from -50 to -100 and $o_{\text{max},b}$ increasing from -500 to -5000. What is interesting is that the sell prices (red) increase higher than the buy prices (blue), despite the range of the buy prices going further, up to -5000 compared to only 1000. It seems that the starting range of the bidding prices seems to have a bigger effect on prices than the end of the bidding range, which mostly just seems to have an influence on the spread of the prices. However, we also do not want to read too much into

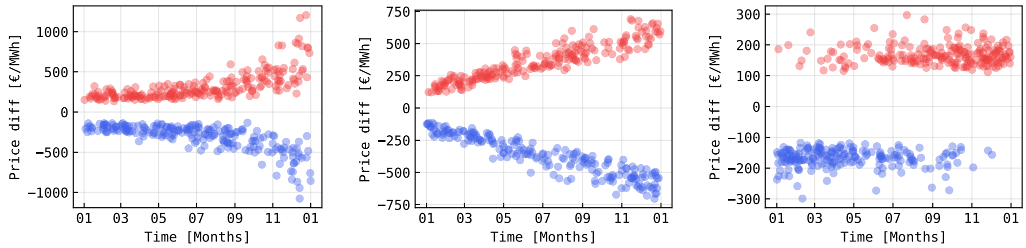
[1] non-stationary processes are also sometimes referred to as non-homogeneous processes or time-varying processes.




(A) Default parameters.

(B) Linearly increasing min and max bid prices.

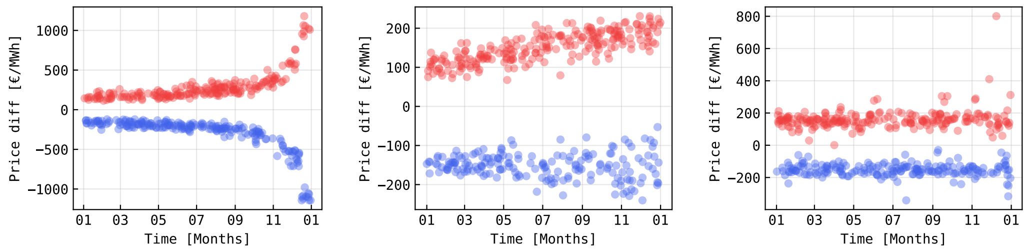
(C) Linearly increasing probability of arrival $p_a$ from 0.001 to 0.3.




(D) Linearly decreasing $n_o$ from 150 to 20 participants.

(E) Linearly increasing $V$ from 50 to 1000 MWh.

(F) Linearly changing $p_b$ from 1 to 0.




(G) Linearly changing the bid size of the bids that other submit from 20 MWh to 1 MWh.

(H) Using PP1, with normally distributed prices and changing the mean and standard deviation.

(I) Using PP1 with prices from a student-t distribution. Changing the degrees of freedom parameter linearly from 5 to 1.

FIGURE 5.3: The GOPACS model simulated over one year. Each dot represents the price difference of a single auction. Each plot shows the effect that changes of some particular parameter have on the model output.

these dynamics as they are not really verified or provable as this data about GOPACS is not available.

Next, in Figure 5.3c we explore the effect of changing the probability of arrival $p_a$ from 0.001 to 0.3, or in other words, from one auction every 1000 hours to one auction every 3.33 hours. An increase in arrivals could happen if congestion in the Dutch grid becomes worse. The outcome is nothing unexpected – the number of auctions increases as seen by the more dense region of the plot. Prices themselves remain unchanged, as expected. It is debatable however whether an increase in auctions in the real world would not have an effect on prices. Figure 5.3d shows the effect of the number of bids of other participants decreasing from 150 to 20. This could be the case where fewer people participate in GOPACS, however, those that do not change their bidding habits. What is interesting is that the effect of this decrease is nonlinear, it seems to accelerate exponentially towards the end. In particular, the spread of the output prices increases a lot when only a low number of bids are submitted. This particular scenario is also realistic – it is not far-fetched to assume that if everything stays the same, only the number of participants goes down, those same participants would start to ask more for their electricity.

Figure 5.3e on the previous page shows the effect of the volume rising linearly from 50 MWh to 1000 MWh. The outcome is that prices would increase, as we assume that the bids of participants remain the same, hence a higher volume would force the DSO to accept higher bids. While it does make sense that a higher volume would increase in higher prices if everything else remains the same, it is most likely unrealistic that volume could increase, but the size of bids and number of participants would not also change. Figure 5.3f on the preceding page shows the outcome of changing the probability of a given auction being a buy auction from 1 to 0. This particular example is not really a realistic scenario, it merely serves to show the effect on the output prices. In Figure 5.3g on the previous page we change the bid size of the bids that others submit, $v_o$, from 20 MWh to 1 MWh. Even though this decrease is linear, the effect on the outcome is exponential. In particular, the effect on the outcome seems to be proportional to the percentage change between subsequent $v_o$. We see that, if other parameters stay the same, decreasing the size of the bids of participants results in increased prices, as now more of the higher-priced bids need to be used to cover the entire volume requirement of the DSO. Again however this scenario is not particularly realistic, in particular, we do not see why existing participants would decrease the volume with which they participate.

Finally, in the two last scenarios, PP1 is used as opposed to PP2. We use the same default parameters where those apply and instead focus on changes to the price distribution. In Figure 5.3h on the preceding page use a normal distribution for the prices and we change the mean of sell prices $\mu_s$ from 100 to 200 while keeping the standard deviation $\sigma_s$ fixed at 25, whereas for the buy prices we keep the mean $\mu_b$ constant at -150 while changing the standard deviation $\sigma_b$ from 10 to 50, all in €/MWh. In Figure 5.3i on the previous page, we use a student-t distribution with mean of 150, resp. -150, and a standard deviation of 25, while decreasing the degrees of freedom via the shape parameters $s_b$ and $s_s$ from 5 to 1. The effects are clearly expected as the output of the model are the prices, which are the exact thing we influence directly. These two examples show that specific situations may be hard to explore with PP2, but trivial with PP1, such as for example the extreme values achieved with the students-t distribution, which we can not easily reproduce with PP2. Finally, for all of the explored scenarios of Figure 5.3 on the preceding page, we want to remind that these are only the outcomes of our model, whereas the real dynamics may be different.

### 5.1.3 Reproducing GOPACS

We developed the model to be able to reproduce GOPACS dynamics and use it for valuation purposes. Using PP1, it is relatively simple to reproduce various GOPACS dynamics as we can look for a good fitting price distribution from the empirical prices and then use that distribution to generate the prices from. It gets however more interesting and challenging to reproduce GOPACS dynamics using PP2 as here the prices can not directly be specified but instead rely on other parameters.

Previously, in Section 3.4.3 on page 39 we found some stylised facts that we would like our model to be able to reproduce. These stylised facts differ slightly between the scraped daily prices and the published monthly prices. In particular, for the daily prices we would like to have:

1. slightly negative correlation between volume and GOPACS price difference, i.e. higher volumes should result in lower prices per volume.
2. prices distributed mostly around some mean value with a few outliers the low-volume regions.
3. prices of up to 3000 €/MWh should be technically possible (see Figure 3.11 on page 29 although unlikely to see as daily mean.

In order to observe these stylised facts we have to specify some model parameters. For the volume to best reflect the volume of real GOPACS prices, we sample it from a beta distribution with parameters as specified in Section 3.4.2 on page 37, that is, we use the SciPy [74] implementation of the beta distribution[2] with parameters $a = 0.623$, $b = 4.978$, $\text{loc} = 0$ and $\text{scale} = 1500$, or alternatively the NumPy [29] implementation[3] with the same $a$ and $b$, and then multiplied by 1500. We also sample both arrival and departure probability parameters from geometric distributions fitted in Section 3.4.1 on page 34, that is, we used $p_a = 0.021$ and $p_d = 0.244$.

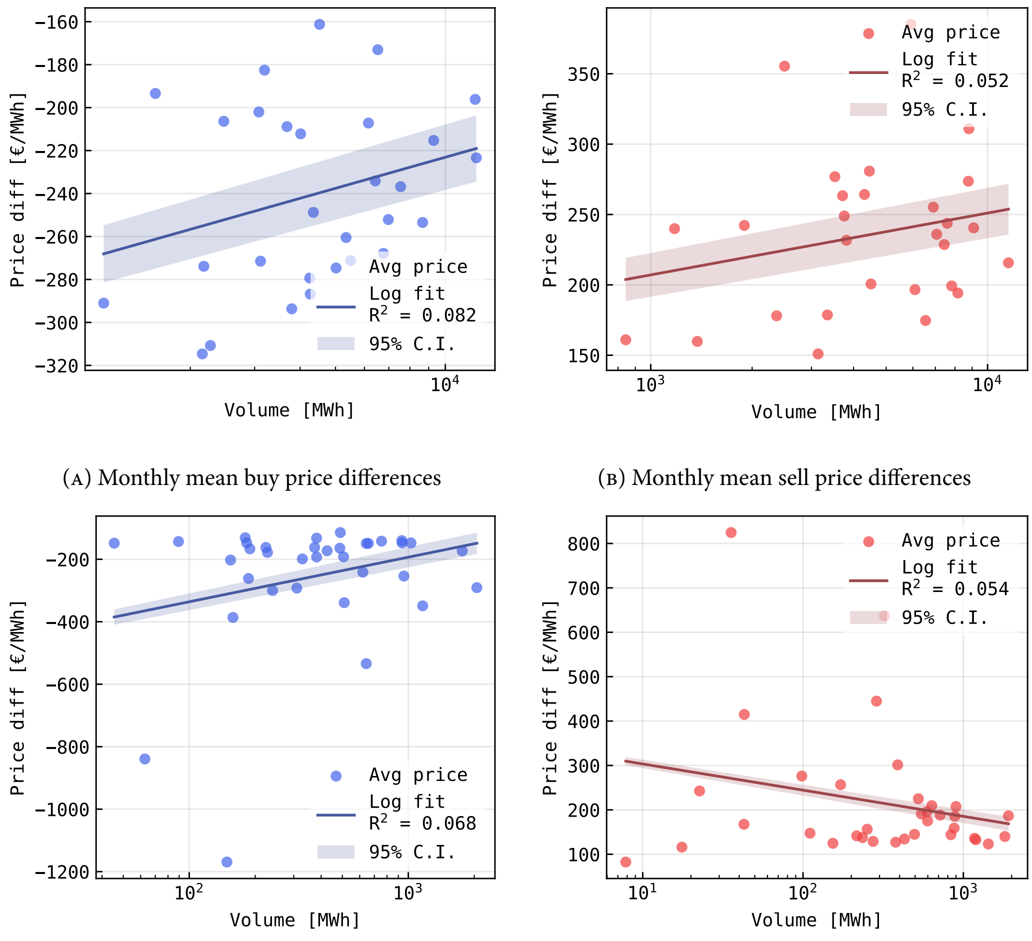
For most of the other parameters, however, we can not set them to the empirical data as there is none, and hence we need to explore which values both make sense and produce the desired results. For the bid distribution parameters, we used $o_{\min,s} = 50$, $o_{\max,s} = 4000$, $o_{\min,b} = -50$ and $o_{\max,b} = -4000$. The max bid price of 4000 €/MWh makes it technically possible to observe very high prices as specified in the third stylised fact, although very rarely. The distribution we use for the bids is still the uniform distribution as we do not know anything about it. Next, we use a bid size $v_o$ of other participants of 20 MWh. This size is chosen because a smaller size either results in too high average prices, whereas a larger size results in too low prices and variance of prices.

Finally, we set the number of other participants' bids to be equal to the volume plus some random variable, that is, $n_o = V + X$. This is arguably the most arbitrary choice, and it was taken as a necessity to achieve prices that resemble the original dynamics. The idea behind setting the number of other participants to be relative to volume is that we assumed that participants may decide whether they even participate by looking at how much volume will be cleared by GOPACS. In GOPACS, bids may get called partially, so a large energy company may not be interested to participate in a low-volume auction out of fear that only a small volume of them is accepted and the expenses of submitting this volume may exceed any benefit gained from it. While we can not verify this theory, the prices we achieve by setting the number of participants to be relative to the volume look convincing. For the random variable $X$, we chose a uniform random integer between 1 and 20, that is $X \sim \lceil \mathcal{U}(0, 20) \rceil$ with $\mathcal{U}$ the uniform random distribution. We chose a random number as technically the volume we model is the cleared volume, however in reality a participant would base their decision of participation on the announced volume, and we use a random variable to reflect the uncertainty between the announced volume and cleared volume.

Summarising these parameters, if a volume of 100 MWh is to be cleared in an auction, we assume that there are 105 bids from other participants, each 20 MWh in size. These bids are uniformly distributed between 50 €/MWh and 4000 €/MWh in a sell auction and -50 €/MWh and -4000 €/MWh in a buy auction. With these parameters, we have run the model and achieved the following results.

Figure 5.4 on the next page shows the outputs of the model in the same fashion as Figure 3.25 on page 40. Focusing on the daily prices from Figure 5.4c on the next page and Figure 5.4d on the following page, they exhibit the stylised facts we expect to see. First, the mean buy price is -200 €/MWh while the mean sell price is 192 €/MWh, close to the values of -166 €/MWh and 203 €/MWh found in Section 3.4.3 on page 39. In addition, most of the prices in the high volume region are close to this mean, with the outliers mostly in the low volume region. Lastly, we also fit the prices and observe a slightly negative correlation. Arguably, however, this relation is weak and was not present in all simulations. A similar impression is given by

[2] https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.beta.html
[3] https://numpy.org/doc/stable/reference/random/generated/numpy.random.beta.html




(A) Monthly mean buy price differences

(B) Monthly mean sell price differences

(C) Single auction buy price differences

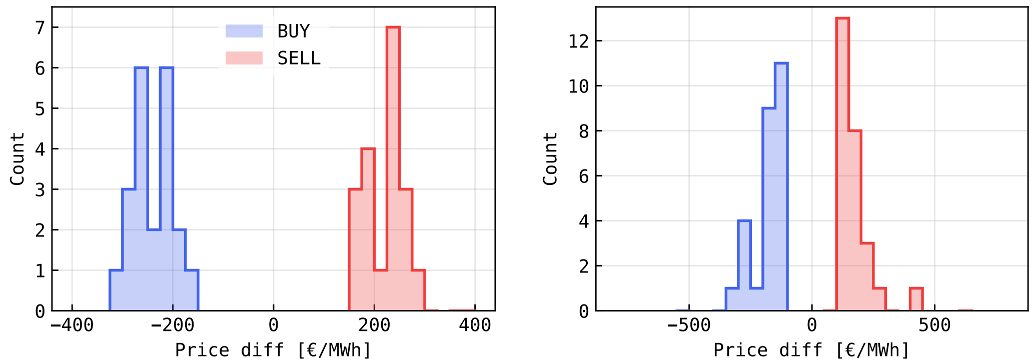
(D) Single auction sell price differences

FIGURE 5.4: Regression between the simulated volume in logarithmic space and the simulated price difference in linear space. Each dot represents either the relation between a simulated monthly mean price and the simulated monthly total volume (top row), respectively between a simulated single auction price and the associated simulated total auction volume (bottom row).

Figure 5.5b on the following page for the daily prices. Again they seem to roughly resemble the empirical data from Figure 3.24b on page 40.

The monthly prices shown in Figure 5.4a, Figure 5.4b and Figure 5.5a on the following page also resemble the real monthly values, but not to the same degree. We believe that the monthly prices contain a lot of trends since they contain the evolution of GOPACS over multiple years, and their reproduction may thus be more difficult without either a more complex model or through the usage of non-stationary parameters. For this reason, we provide a small example of what dynamics can be achieved by specifying non-stationary parameters, shown in Figure 5.6 on the next page.

Figure 5.6 on the following page shows the output of the model run over the same period that GOPACS has been running. A signal was used to specify how both the volume $V$ and probability of arrival $p_a$ should change over time, making them non-stationary parameters. The resulting output bears quite a lot of resemblance to the real GOPACS prices shown in Figure 3.10 on page 28, which is however intended as the signal was chosen to mimic the dynamics in these prices. While this is a good showcase of the capabilities of the model and the




(A) Monthly mean price differences

(B) Single auction price differences

FIGURE 5.5: Histogram of both the monthly mean buy and sell prices differences and the single auction price differences, as simulated by our model. In all cases, the simulated price difference is shown.

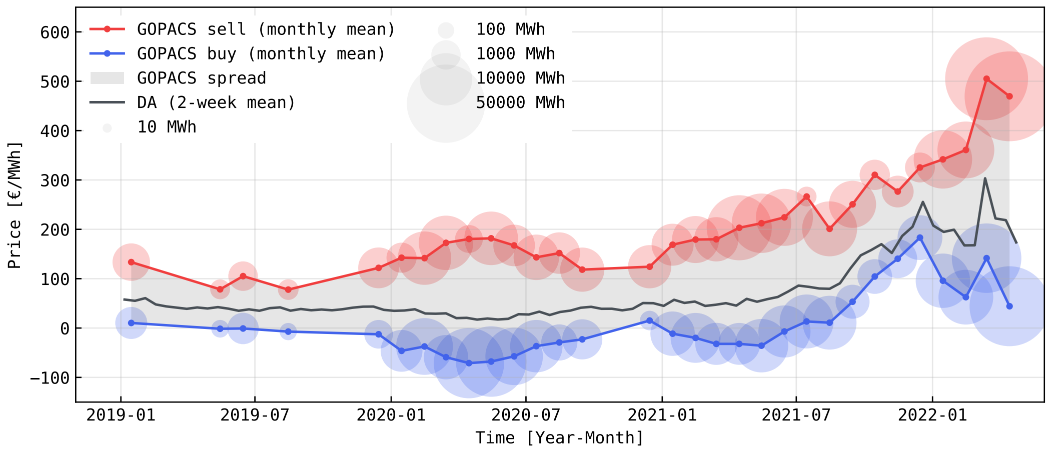


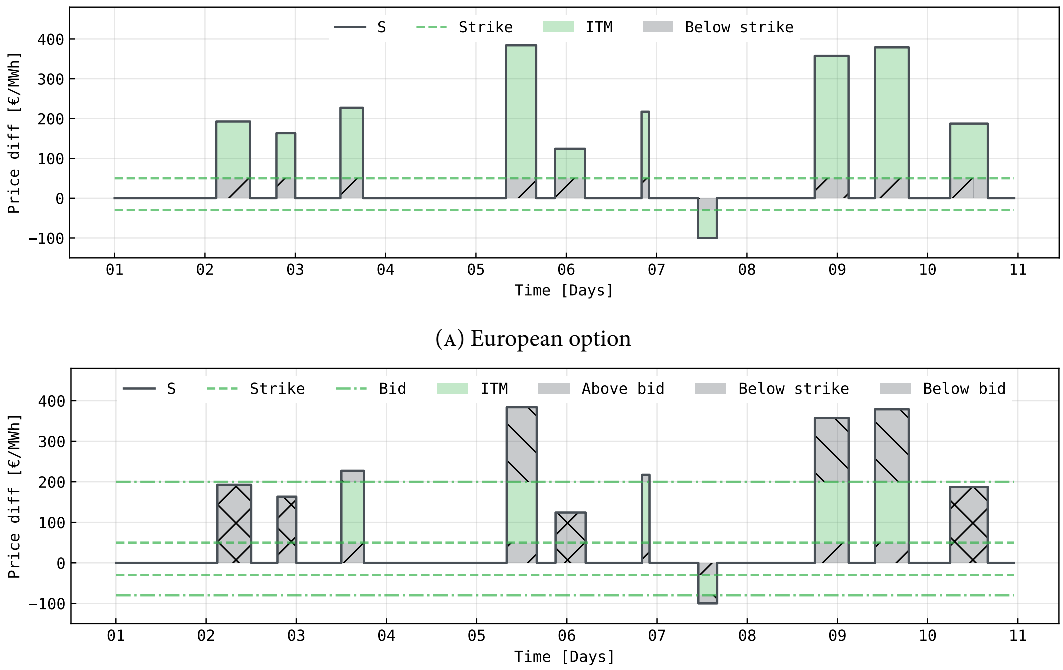

FIGURE 5.6: The monthly mean of the buy, sell and spread price simulated with the GOPACS model using non-stationary parameters for volume and probability of arrival. The monthly average DA price is given as a reference. Each dot represents a month, with the size of the circle surrounding it indicating the volume. Some months do not contain any data, as there were no recorded GOPACS auctions.

benefits of non-stationary processes, it also is arguably not of much use as the desired output depends strongly on the signal used for the non-stationarity, and said signal was not produced by the model but simply manually created with inspiration from real data. In other words, it is very unlikely that the model would reproduce these dynamics on its own. A discussion of this issue can be found in Section 6.2 on page 82.

## 5.2 Real Option valuation

### 5.2.1 Options overview

Besides simulating GOPACS, we also use the GOPACS model for the valuation of a flexible asset (FA), which we approach with a RO approach. This RO approach consists of pricing the asset the same way one would price an actual financial option. We note that no actual financial options are written, we only use a valuation approach inspired by these financial options to value the optionality of the flexible asset. Specifically, the value of the optionality of the FA during 1 hour of 1 MW of capacity is the same as the value of a theoretical financial option maturing at this specific hour. Over a longer period, we thus value the asset the same way as one would value a strip of such actual options, with one option per hour and MW of capacity during this period. We have suggested both the usage of European call and put options and digital call and put options, defined in Section 4.2.2 on page 50. Together with the choice of price process of the model (be it PP1 or PP2), this leaves us with four different approaches for the valuation of a particular FA. We will begin with a visual inspection of how the options are computed in the case of PP1.




(A) European option

(B) Digital option

FIGURE 5.7: Visual representation of how the various parameters of both European options (top figure) and digital options (bottom figure) are computed. The simulated price process $S$ is the same as in Figure 4.2 on page 43 and Figure 4.3 on page 47.

Figure 5.7 shows how both the European and digital option payoff is computed on the model output $S$, that is, the GOPACS price difference to DA. In the case of the European call for the sell prices and European put on the buy prices[4], Figure 5.7a, the options are in the money (ITM) if the price difference $S$ is either above the sell strike price $K_s$ (call) or below the buy strike price $K_b$ (put). We assume that both $K_s > 0$ and $K_b < 0$, hence if there is no GOPACS auction, meaning $S = 0$, the options are always out of the money (OTM). This means that with the realistic arrival and departure rates we estimated, a lot of the options we have will expire OTM, as at any given moment we are more likely to not have a GOPACS auction than to have a GOPACS auction.

In the case of the digital call and put options, Figure 5.7b, the payoff computation is a bit more complex. Here, a call option is ITM if both the strike price $K_s$ and bid price $B_s$ are below the GOPACS price difference $S$. In the case of such an ITM option, the payoff is the fixed quantity

[4]To recapitulate once more the somewhat unintuitive relation between the options and auction type: In a sell auction, the DSO looks for energy sellers, who ask prices higher than DA. We own a FA, so our profit increases the higher we can sell energy for, which is similar to owning a call option on the sell auction price difference to DA. In a buy auction, the DSO looks for energy buyers, who bid prices lower than DA to buy energy. We own a FA, so our profit increases the lower the buy prices, similar to owning a put option on the buy price differences, which also increases in value the lower the prices are.

$B_s - K_s$. The difference between the bid price and the price difference, $S - B_s$, referred to as 'Above bid' in the figure, is wasted profit, as it means that we could have submitted a higher bid, which, as long as it would have been below $S$, would have netted us a bigger profit. $S$ may also be below the bid price $B_s$, referred to as 'Below bid' in the figure, which is also wasted profit, as we could have submitted a lower bid price, and as long as said price would have been below $S$ and above the strike price $K_s$, we could have profited from this auction as well.

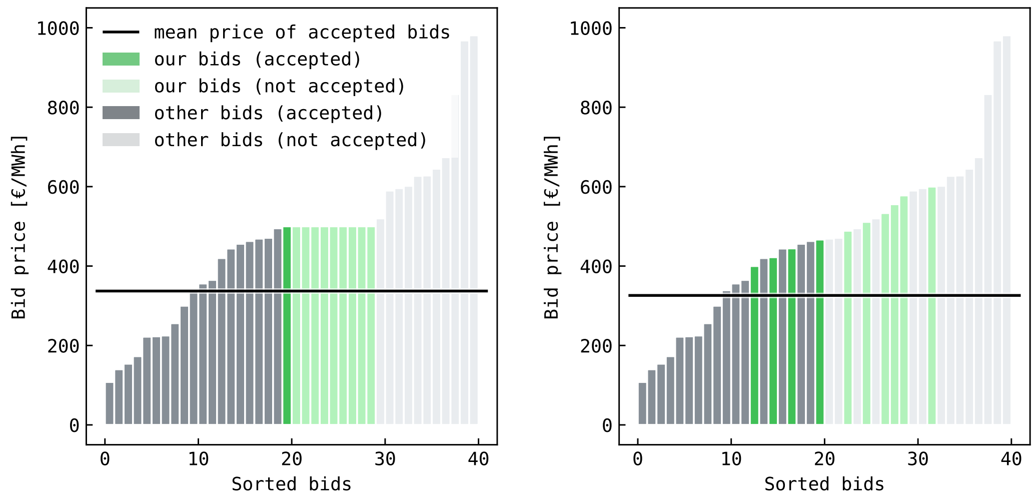
In both the European and digital situation, there is an area between the sell and buy strike, which is essentially lost even if the option expires ITM. It is this value, namely $K - 0$ or simply $K$, which we use to represent the running costs of the FA. These costs are assumed to always occur when we participate, and hence we represent them with the strike price.

In Figure 5.7 on the previous page we have shown the application of the European and digital options on the GOPACS model with PP1. In Section 4.2.2 on page 50 we have discussed why the European option is not ideal, as it represents the mean value that could be extracted from GOPACS without considering the difficulties that a participant faces when actually trying to achieve these prices in GOPACS. We have thus argued that the digital option is a potentially more correct approach to the valuation. However, we have also argued in Section 4.1.4 on page 45 that PP2 offers various advantages over PP1 for the pricing of digital options, in particular, it gives a more realistic valuation when more than a very small number of bids are submitted. We have used PP1 in the above figure as the visualisation of the bidding process with PP2 is not easily done due to the additional dimension added by the bidding ladder. However, in the remainder of this section, we will focus on the usage of the digital option with PP2, which we believe to be the most accurate pricing approach, at the cost of additional complexity.

There are two aspects of the digital options strategy that deserve further investigation. First, irrespective of the price process, is that the bid price $B$ is free to be chosen at will by the participant, unlike the strike price $K$ which is given by the running costs of the FA. There are some soft boundaries to choose for this bidding price – it should not be lower than the strike price, otherwise, participation does not make sense. Additionally, if the bidding price is too low, we will often get called but will not make much money. If the bidding price is too high, we will make a lot of money if our bid is called, however, it being called is unlikely as there are most certainly enough bidders bidding less than us.

The second aspect of the digital bidding approach is apparent only when PP2 is used. When we submit more than one bid into an auction, as is e.g. the case when we participate with a FA that has more than 1 MWh of flexibility, we can compare various ways of submitting these bids. In the case of PP1, submitting more than a single bid of 1 MWh is simply taken care of by multiplying the payoff with some cannibalisation factor, which is hard to estimate for GOPACS. On the other hand, in PP2, we handle this arguably more correctly by simulating the bidding process.

Figure 5.8 on the following page shows the simulated bidding process for a single hour of a problem period, thus corresponding to a single digital option. We see both our own bids (in green) and the bids of other participants (in grey). The bids are sorted according to their bid price, and since this represents a sell auction, the bids with the lowest price are first. In this example, we suppose that all bids, both our own and others, are of a fixed size of 1 MWh, that is, $v_u = 1$ and $v_o = 1$. In total, there are ten bids from us and 30 bids from other participants. In the bidding process, we simulate, these bids are sorted, and the volume $V$ that is needed to solve the congestion problem of this particular hour is used as the cutoff or threshold value for the bidding process. In other words, bids are accepted according to their size, until the total volume of accepted bids is equal to the volume required to solve the congestion issue.




(A) Block

(B) Linear

FIGURE 5.8: Our bids sent into the bidding process containing the bids of all participants. We compare two particular strategies of submitting these bids.

Since we submit more than one bid, there are multiple ways to define the bid prices of these bids. We consider two strategies in general, that is, submitting all bids with the same bidding value (Figure 5.8a), which we will refer to as 'block', or submitting all bids with the bidding value out of a linear range (Figure 5.8b), which we will refer to as 'linear'.

### 5.2.2 Bidding optimisation

In the previous section we mentioned that when considering the valuation of a FA using digital options, there are two aspects that are interesting to explore further, which are (1) the bid price at which we submit bids and (2) the 'shape' of the bids in case we submit more than one bid. In this section, we explore these two aspects in more detail.

Figure 5.9 on the next page shows how both the linear and block style bidding performs for range of possible bid values, and across various scenarios. In the first row, we compare how the linear strategy changes when we increase the start and the end of the linear range that we chose the bids from. For each bid value $x$ on the $x$-axis, the block bid is ten bids of $x$ MWh, the linear bid is ten bids valued from $x$-25 to $x$+25 (Figure 5.9a on the following page), $x$-75 to $x$+75 (Figure 5.9b on the next page) or $x$-150 to $x$+150 (Figure 5.9c on the following page).

The output, on the $y$-axis, is the distribution of prices achieved per accepted bid, across 1000 simulations. The second row is practically the same experiment as the first row, with the only difference that this time the output is the price achieved per submitted bid. These two outputs were chosen as they both represent meaningful metrics. We recap that GOPACS auctions take place on average seven hours before the start of the problem period. We submit some bids to the auction, and once the auction ends we are told how many are accepted. The price per accepted bid shows what price we achieve per actually delivered MWh of energy. The price per submitted bid, when multiplied by the number of bids we submitted, gives us the actual total value of our FA that we managed to achieve in GOPACS. Both quantities are interesting, however since the focus of our work is on valuing the FA, we will focus on the price per submitted bid for future experiments, as this is closer to the total value of the FA. Moreover, since the output is a distribution, which we do not plot in full, we instead plot the mean of the distribution as well as the mean of the all the values higher than the median and the mean of

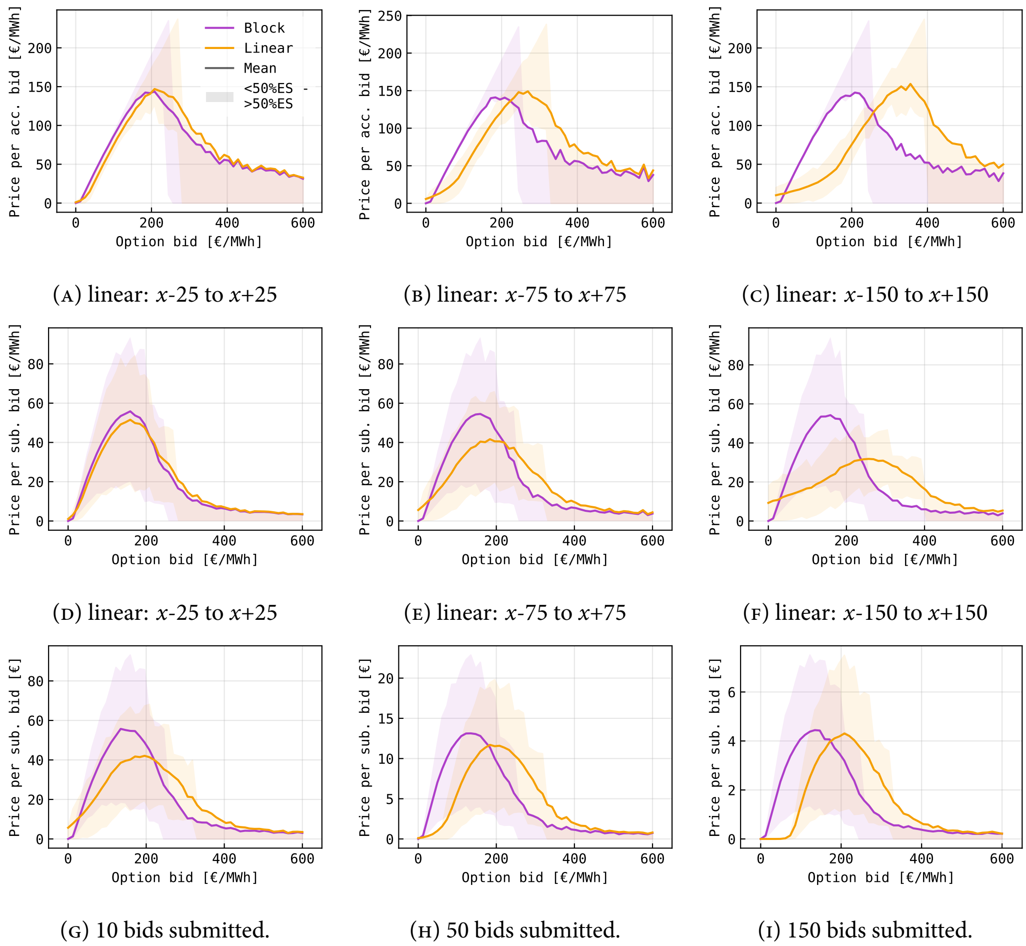


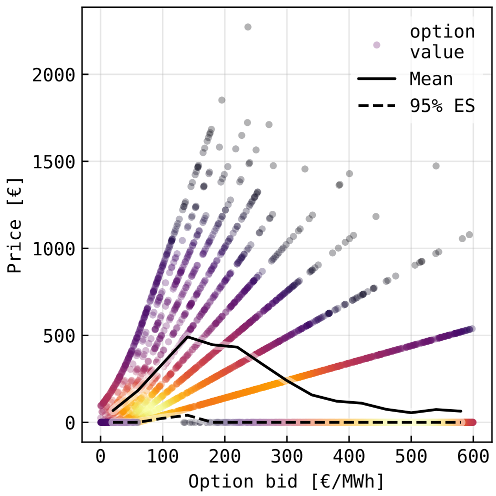

FIGURE 5.9: Comparison of the block vs linear bidding strategies. In the first row, the price per accepted bid is shown as a function of bid size. Each column differs by the start and end of the range (indicated below each subplot) of the linear bids; the block bids are the same for the three columns. The second row is nearly the same as the first row with the only difference being that the output is the price per submitted bids. In rows one and two, the number of submitted bids is always ten, in row three, the range is kept at x-75 to x+75 but we increase the number of bids submitted.

all the values lower than the median, which the same as the expected shortfall, particularly, $\mathrm{ES}_{0.5}(P)$ and $-\mathrm{ES}_{0.5}(-P)$ when $P$ is the distribution that we have. We chose the mean and the area delimited by the two ES values because we find it to transmit well the shape of the underlying distribution.

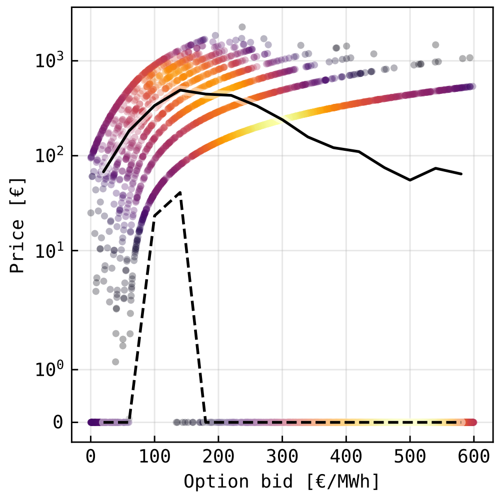
Going back to the figures themselves, first, we see that the price at which we submit the bids does matter, as in all figures the price per bid goes up, then peaks, before it goes down again. Next, comparing the linear bid with the block bid strategies, we see that when the linear bid range is narrow, it behaves very similar to the block bid strategy, which is to be expected. Only when the range grows larger do the two strategies diverge. In the case where the output is the price per accepted bid, in the most extreme case, Figure 5.9c, the linear bid and block bid strategies look very similar, just scaled on the $x$-axis. It looks like the linear bid strategy achieves a slightly higher mean price per accepted bid, though not by much. In the case where the output is the price per submitted bid, which is relative to the total price for the FA, the linear bid strategy performs much worse on average than the block bid strategy especially when the range is wide, as is the case in Figure 5.9f. However even though the mean in this particular case is worse for the linear bid strategy, we can also look at the mean of the lower half of the distribution, which fares well against the block strategy. For example, at $x = 200$, the lower ES of both strategies is relatively similar, while the means are still very different. This means that the block strategy, even though having a higher mean, has a non-negligible risk compared to the linear strategy.

Finally, we will take a look at the third row of Figure 5.9 on the previous page, which shows a different experimental scenario. This time, the range of the linear bids is kept at $x$-25 to $x$+25, however, the number of bids that are submitted by us is changed. In the first two rows, the number of bids that are submitted was always ten, now we change from 10 to 50 to 150. Otherwise the configuration is the same, hence Figure 5.9e on the preceding page is equivalent to Figure 5.9i on the previous page. There is not much change with respect to the ideal bid value as the number of our bids is changed. However, the output price per submitted bid changes drastically – the more bids we submit, the lower on average the price we will achieve per bid. This does make sense, as it is a showcase of the self-cannibalisation factor discussed in Section 4.1.3 on page 44, for which we developed PP2, as in PP1 the output would, unrealistically, not change unless we would discount it manually. Next, we also see that as the number of bids increases, the linear bid strategy becomes more competitive. When we submit a very large number of bids as is the case in Figure 5.9i on the previous page, the linear bid strategy performs on the same level as the block bid strategy.

So far in Figure 5.9 on the preceding page we have only looked at the distribution of option prices as represented by the mean and the area between the mean of the upper half and lower half of the distribution. However, it is also interesting to take a look at how exactly the individual values are distributed. Using a probability of arrival $p_a = 1$ and a probability of departure $p_d = 0$, then simulating for $N = 1$, we practically take a look at the range of possible values that a FA may take. The FA we use is represented by ten linearly spaced digital call options, that is, $A = C_D(10)$, and for the GOPACS prices, PP2 is used.




(A) Linear scale




(B) Logarithmic scale

FIGURE 5.10: The various valuations that the FA can take for one hour of GOPACS. The strategy consists of ten linearly spaced digital call options, $A = C_D(10)$. In addition to the various outcomes, we also plot the mean across the option bid values as well as the 95% ES. In both plots, the dots are colored according to their density, with black the lowest density and yellow the highest density. The density is estimated using kernel density estimation and is included to better highlight dense regions.

Figure 5.10 on the previous page shows how the value of a FA depends on the bid value. We view the plot both on a linear scale (Figure 5.10a on the preceding page) and logarithmic scale (Figure 5.10b on the previous page). Since the options used are digital options, whose payoff scales with the bid size as the strike price is kept constant, we see linear growth of the FA price at first. In particular, for every bid value, we see exactly eleven possible values of the FA, which correspond to how many of the ten digital options that we submitted got accepted, a value in [0, 10]. Each of these outcomes has a certain probability of occurring. For low bid values, there is a high probability that most of the options get accepted, however, as we increase the bid value, the lower the number of accepted bids goes. We colour the individual dots by their density, and one can see that close to bidding values of around 600, we practically only have two outcomes left, namely that one option of the bunch is accepted or none, with the probability of none becoming higher the higher we bid. This again highlights the importance of choosing an adequate bid value for participation in GOPACS.

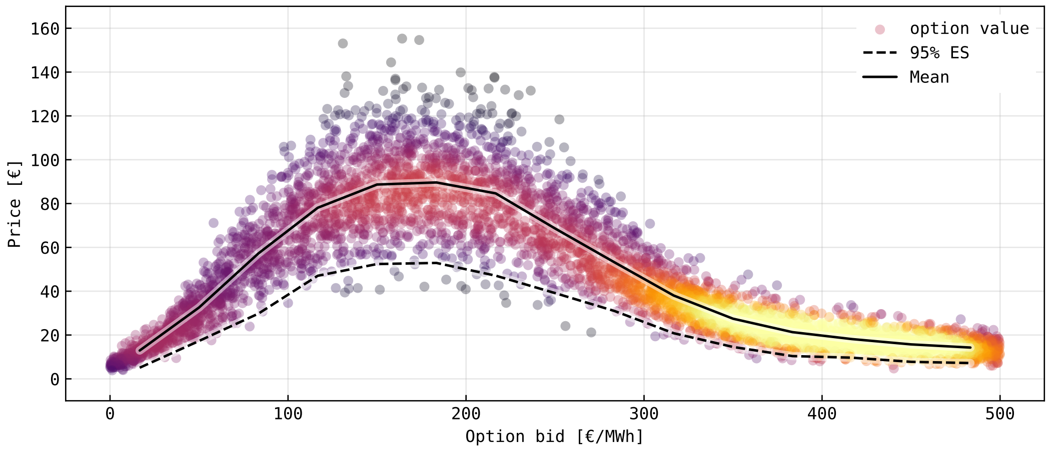



FIGURE 5.11: The various valuations that the FA can take for one month of GOPACS. The strategy consists of ten linearly spaced digital call options, $A = C_D(10)$. In addition to the various outcomes, we also plot the mean across the option bid values as well as the 95% ES. In both plots, the dots are colored according to their density, with black the lowest density and yellow the highest density. The density is estimated using kernel density estimation and is included to better highlight dense regions.

In this next experiment in Figure 5.11, we keep the same parameters for the price process and option pricing, however the arrival-departure process is changed this time, in particular, we use the realistic $p_a = 0.021$ and $p_d = 0.244$ from Section 5.1.3 on page 65. In addition, instead of plotting the value of the FA for a single option, we look at the mean over one month, precisely we have:

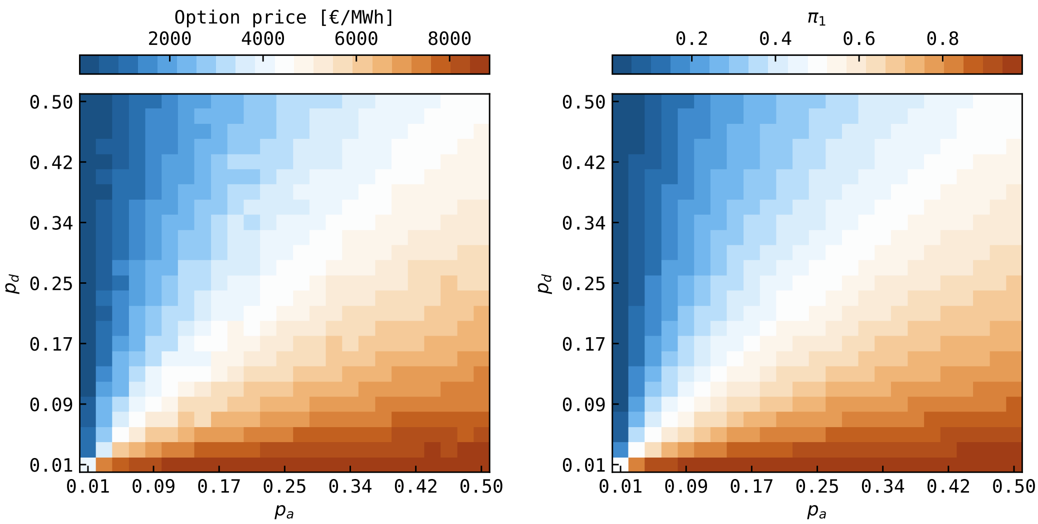
$$A = \frac{1}{720} \sum_{i=1}^{720} C_D(10) \tag{5.1}$$

Equation (5.1) shows how we compute the asset value. We take the mean of all of the options of one month, which is 720 hours. The options used to represent the asset are ten linearly spaced digital call options. Figure 5.11 shows how the mean value of a FA changes depending on the bidding value, with the mean taken over one month, and the output being the total price of the 10 MWh FA per hour. If we multiply the $y$-axis by 720 hours in a month, we get the option value of the FA per month.

We see that the spread of the distribution of the values in Figure 5.11 on the preceding page is already much tighter than before in the single-hour FA value. In particular, we also do not see the streaks anymore that we saw before, which makes sense as we now take the mean across a month so the underlying options can take on much more values than just the fixed payoff. The mean itself however is similar to before. We see that as the option bid goes up, so does the price of the FA, although only to a certain degree. At around a bid of 175 €/MWh in this particular case, we level out and the value of the FA starts to decline. Interesting is also the 95% ES which we plot to convey an idea of the potential risk one would face in this context. At its highest, we have a 95% ES of 50 €/MWh, which in practice means that even in the worst 5% of outcomes in this simulation, we still average around 50 €/MWh per hour for the FA during every single hour of the month, in the best case with respect to the bid value. However if the bid value is not well chosen, the value of the FA declines.

### 5.2.3 Arrival and Departure rates

In Section 4.1.5 on page 46 we have argued that the GOPACS price process and GOPACS arrival–departure process, albeit combined in the model, are actually separate in the sense that changes made to one of them do not affect the other process. While we have so far focused mostly on the pricing aspect of GOPACS and the option pricing model, we want to briefly draw attention to the arrival-departure process and how it affects the prices.




(A) Option price (B) Steady-state probability

FIGURE 5.12: How both the FA value and steady-state probability changes as the parameters of the arrival–departure process change. The mean of 50 simulations are used for the valuation of the FA, which is computed with 20 digital call options taken over one month.

In Figure 5.12 we see how the price of a GOPACS option (Figure 5.12a) and the steady-state probability of being in the GOPACS state $\pi_1$ (Figure 5.12b) change with respect to changes in the probability of arrival $p_a$ and the probability of departure $p_d$ values. The option used is a digital option, and PP2 is used, however, we do not care about the prices themselves, but rather how the prices change with respect to changes in the arrival-departure process. We find the same pattern for all option types and both price processes, which is to be expected as the arrival-departure process does not influence the price process.

We find that the value of the option changes the same way as the steady-state probability. This is not unsurprising, given that at least in the case of PP1, where we have the analytical solution to the options prices, the value of each of the four option types is directly multiplied by the steady-state probability, as shown in Section 4.3 on page 55. What we see is that if $p_a$ and $p_d$ are the same, the steady-state probability stays the same and auction prices stay constant. If $p_a$ increases but $p_d$ stays the same, more auctions take place, however, their duration is unchanged, so there is 'more' GOPACS, or a higher probability that a given hour is a GOPACS auction, and thus the value of the option increases. The opposite is the case when $p_d$ increases but $p_a$ stays the same, here we have the same number of auctions, but their duration gets shorter.

## 5.3 Case study

We will end this chapter with a case study on the valuation of a FA. We assume that his FA is one which follows the 'Consume less' strategy mentioned in Section 4.2.1 on page 49, as is typically the case for a DRA. To recapitulate, it means that we, the owner of this FA, typically consume energy at all times, and we buy this energy on the DA market for the price $S_{DA}$. By consuming this energy, we produce a good or service whose value we assume to be relative to the price of the energy we use for it. We denote the value or price we can sell this good or service for as $S_{DA} + x$. Now if there is a GOPACS auction, we assume that we can participate in this auction and achieve the price $S_{GO}$. We only participate if the price we can achieve is higher than the price we could have received for the product or services otherwise produced plus any additional costs $y$ occurring because of participation on GOPACS (and thus deviating from the normal schedule). We combine $x$ and $y$ into the strike price, that is, $K_s = x + y$. Now we are presented with the final equation which determines if it makes sense to participate in GOPACS, which is the case when the price for selling energy on GOPACS exceeds the costs of doing so, that is, $S_{GO} \geq S_{DA} + K_s$. Since the GOPACS model simulates $S$, defined as $S = S_{GO} - S_{DA}$, we can rewrite this as $S \geq K_s$. While this whole process was already explained extensively in Section 4.2 on page 49, we have summarised it here again to highlight that the value of the FA is the revenue of the asset, with everything taken care of. In order to get a profit, one only needs to deduct the investment cost of the asset.

In addition, in order to make this valuation as accurately as possible, we assume that the FA that we value is physically located in some Dutch province. This means that with this FA, we can only participate in auctions that take place in this province. In practice, this means that we need to use parameters for the underlying GOPACS process that are close to the parameters for this particular location. For the arrival-departure process, determined by $p_a$ and $p_d$, we use the parameters that we sampled from real data, seen in Table 3.4 on page 37. However since similar parameters are found for most regions, we will simply use the most common ones, $p_a = 0.014$ and $p_d = 0.133$. These are also the parameters found in Noord-Holland, which is potentially the best region to participate in GOPACS sell auctions as it appeared in the most sell auction announcements over all other provinces. Next, for the probability of a given auction being a buy or sell auction, we used $p_b = 0$, meaning all auctions are sell auctions. Doing this allows us to later discount the presented prices for each region by multiplying them with $1 - p_b$ with $p_b$ in this case their actual buy-sell ratio from Table 3.4 on page 37.

For the price process, we used PP2 with digital options as these give the most realistic results. For the other parameters of the price process, we mostly used those already presented in Section 5.1.3 on page 65 in which we argued which parameters can be used to most accurately reproduce GOPACS. To summarise these parameters we used $o_{\min,s} = 50$, $o_{\max,s} = 4000$[5], $v_o = 20$ and $n_o = V + X$ with $X$ a random integer between 1 and 20. For the volume, we

[5]We have not specified the parameters for the distribution of buy prices, as we specifically focus on the case of an asset that can only participate in sell auctions, hence we do not care about buy prices.

TABLE 5.1: Various case study results. All values are given in €/MW of capacity of the FA. The first column is the FA capacity, the second column the strike price in €/MWh. The row is the base case. For the base case we indicate the best bid to submit into the auction process. We also give various properties of the distribution. Rows S1, S2 and S3 are special scenarios explained below.

| FA Size | | 5 MW | | | 20 MW | | | 100 MW | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Strike | | 0 | 50 | 100 | 0 | 50 | 100 | 0 | 50 | 100 |
| B | Bid | 175 | 200 | 200 | 175 | 200 | 225 | 200 | 200 | 225 |
| | Mean | 61.5k | 38.7k | 23.5k | 21.9k | 14k | 9k | 5.37k | 3.28k | 1.87k |
| | SD | 7.2k | 4.49k | 3.03k | 2.65k | 1.73k | 1.01k | 561 | 363 | 225 |
| | VaR95 | 50.3k | 31.6k | 19.2k | 17.3k | 11.3k | 7.3k | 4.6k | 2.77k | 1.55k |
| | ES95 | 47.4k | 29.8k | 17.9k | 16.3k | 10.4k | 6.98k | 4.38k | 2.65k | 1.46k |
| S1 | Mean | 91k | 63.4k | 38.1k | 35.8k | 23.2k | 14k | 7.97k | 5.37k | 3.15k |
| | SD | 9.53k | 6.26k | 3.43k | 3.66k | 2.38k | 1.23k | 770 | 493 | 291 |
| | VaR95 | 77.2k | 54.2k | 32.8k | 29.5k | 19.3k | 11.9k | 6.63k | 4.68k | 2.7k |
| | ES95 | 73.5k | 50.9k | 31.3k | 27.7k | 18.5k | 11.3k | 6.45k | 4.31k | 2.55k |
| S2 | Mean | 114k | 78.8k | 47.4k | 45.1k | 29.9k | 17.7k | 10.2k | 6.95k | 4.04k |
| | SD | 9.04k | 6.54k | 4.53k | 3.38k | 3.04k | 1.77k | 905 | 570 | 367 |
| | VaR95 | 99.6k | 67.8k | 39.9k | 39.9k | 25.5k | 14.7k | 8.72k | 5.8k | 3.47k |
| | ES95 | 96.7k | 65k | 38.2k | 38.6k | 23.7k | 14k | 8.26k | 5.75k | 3.3k |
| S3 | Mean | 103k | 74.7k | 44.7k | 37.7k | 26.3k | 16.6k | 8.78k | 5.92k | 3.66k |
| | SD | 8.59k | 7.18k | 3.95k | 3.86k | 2.53k | 1.51k | 851 | 592 | 348 |
| | VaR95 | 89.8k | 64.2k | 38.2k | 30.9k | 21.7k | 13.9k | 7.49k | 4.97k | 3.13k |
| | ES95 | 86.1k | 61k | 35.5k | 29.7k | 20.9k | 13.3k | 7.17k | 4.75k | 2.99k |

used random values sampled from a beta distribution with $a = 0.623$, $b = 4.978$, $\text{loc} = 0$ and $\text{scale} = 1500$. We did not use a specific volume for each region, because our approach makes use of the CB volume, which is not given per region.

When it comes to our participation in GOPACS, in the case where we only participate in sell auctions, we have three parameters to specify: the bid price $B_s$, the strike price $K_s$ and the capacity with which we intend to participate, that is, $v_u$. With the intent of generalising the case study as much as possible, we chose three different strike prices, 0, 50 and 100 €/MWh, and also three different asset capacities, 5, 20 and 100 MW. We believe that these should cover most situations. Finally, for the bid price, we chose the bid price that produces the highest payoff. Since the asset capacity is bigger than one, we submit multiple bids, and we do so using the linear bidding style, with bids starting 50 €/MWh lower than the bid price and ending 50 €/MWh higher.

For the case study that we perform, we do 500 simulations, each simulation is the computation of the profit achievable with the FA over the time span of one year ($N = 8760$). For the discounting, we assume that the FA is ready to participate now, that is, $t = 0$, and the risk-free rate we assume is 1.5%, thus $r = 0.015$. For every simulation, we compute a distribution of valuations $A$ which we quantify by the mean, standard deviation (SD), 95%VaR ($\text{VaR}_{0.05}(A)$)[6] and 95% ES ($\text{ES}_{0.05}(A)$). The results of the investigation are given in Table 5.1

[6] Note that even though we refer to it as the '95%VaR' or 'VaR95', we actually compute the 5%VaR. The measure '95%VaR' is relatively well known, and typically refers to the 95-th percentile of a loss distribution (in which profits are negative and losses are positive). However our distribution is a profit (and loss) distribution (in which profits are positive and losses are negative), hence this 95th percentile actually becomes the 5th percentile.

In Table 5.1 on the preceding page we see the FA valuations $A$, for the nine combinations of strike price (in €/MWh) and FA size (in MW), and for four scenarios. The first scenario 'B' is the base case, it is simply the estimated value for one year of participation in GOPACS. We see that, across all strike prices an asset capacities, the mean prices per MW of capacity reach from 1,870 €/MW to 61,500 €/MW[7]. In this particular scenario, it would not make sense to invest in a large asset to participate in GOPACS, as the 5 MW asset achieves a considerably higher revenue per MW of capacity than the larger 100 MW asset. If the strike price is high, say 100 €/MWh[8], and our asset has 5 MW of capacity, we can expect a profit equal to or higher than 23,500 €/MW per year with a 95% chance, and in the worst 5% of outcomes, the average profit is still 17,900 €/MW.

This base scenario gives the profit that can be made with the FA. We do this valuation for one year, based on assumptions (for the price process) and historical data (for the volume and arrival–departure process). While we could technically extend this valuation to a longer period, this would most certainly result in inaccurate results, as it is highly unlikely that some dynamics will not change in longer periods. There is simply an element of uncertainty that grows as time moves on, which our model does not capture – on the other hand, the outcome distribution of our model shrinks as time increases as we approach the expected value of the FA. This limitation will be discussed in Section 6.2 on page 82.

There is however one way we can introduce more uncertainties into our model, which is by using non-stationary processes to do a scenario-based approach. Specifically, we investigate three scenarios, building on the base scenario: In S1, we explore the possibility of GOPACS auctions becoming more frequent, which we do by linearly increasing the probability of arrival $p_a$ from the original 0.014 to 0.3, so a halving of the time between auctions. In S2, we explore the possible scenario of the GOPACS auction volume increasing by changing the location parameter of the beta distribution from 0 to 100. In S3, we explore the possible scenario of liquidity drying up, simulated by reducing the total number of traders by 20 (although still making it relative to volume). These are three experimental scenarios, which we could imagine to happen, however additional scenarios can be easily investigated with our model as well.

The results for these scenarios are also indicated in Table 5.1 on the preceding page. We can see that in all three scenarios, the mean prices are higher than in the base scenario. The highest relative profit is achieved when strike prices are minimal, asset size is minimal, and in the case where the volume increases, with a value of 114,000 €/MW. The highest absolute profit in the same situation except when the asset size is maximal, we reach a total of 1,020,000 € with the 100 MW asset over one year. Also in all three of the scenarios, the risk measures, VaR and ES, are higher than in the corresponding situations of the base scenario, indicating that at least for these three scenarios, the risk profile stays proportional to the mean valuation.

The values presented by this case study may not the same ones as would be achieved by a real asset participating in GOPACS. All along with the development of our model and RO pricing strategy, various assumptions and simplifications have been made. These assumptions constitute a risk that can not be neglected and will be explored in the next chapter.

[7] We now give the prices in €/MW, as they represent the price per capacity of the asset, assuming that this asset participates continuously on GOPACS and always bids its full capacity during each auction.

[8] The strike price is still given in €/MWh as it only applies to the real energy that we actually deliver, not the capacity of the asset.

# Chapter 6

# Discussion

In this thesis, we have explored a Dutch electricity flexibility market, GOPACS and developed a model to reproduce this market, then used this model to value a FA participating in this market. Various types of financial risk are present in this valuation, which we will discuss in the following sections.

## 6.1 Model risk

There are three main parts of this work that each build on top of each other, which are (1) the GOPACS data investigation and analysis leading to an understanding of how the market functions and serving as the basis for (2) the development of a stochastic finite-difference model that is capable of reproducing simplified aspects of these dynamics which itself is used as the underlying in a (3) RO modelling framework that is used for the valuation of a flexible asset that is participating in GOPACS. These three parts of our work build on top of each other and influence each other in many ways. This influence may however not only have benefits, as inconsistencies, problematic or incorrect assumptions and errors from one part may propagate to others and grow in this process and become model risks. In this section, we want to investigate and discuss this model risk, in particular with respect to specific situations found or assumptions and decisions made in the three parts mentioned.

Model risk can be broadly defined as "the possibility that a financial institution suffers losses due to mistakes in the development and application of valuation models" [50], or more specifically as "the risk that the model is not a realistic/plausible representation of the factors affecting the derivative's value" [50]. In our context, we consider as the model risk all risks that someone using our model is exposed to that are not due to the implementation of the model, but rather due to faulty assumptions or simplifications. During the development of our model, we have tried to avoid or minimise the exposure to such risks, however, we can not fully avoid them, and thus we want to present them in this section.

### 6.1.1 GOPACS data

The first part of this thesis is the exploration of GOPACS dynamics. There are some problematic aspects, although most of them are out of our control. First, regarding the available data, there is a lack of clear documentation. One example is the four data sources as documentation for all sources but MA are nowhere to be found. In particular, the differences and similarities between the CB and MA data sources are not explained and all assumptions we made about them rely on our exploration of data. Next, there are also various inconsistencies in the data. The PM data source contains buy and sell prices, but not the prices for each DSO. The EX data source contains the prices of each DSO, but only the spread instead of the buy and sell prices. It is unclear to us why these are reported separately and not combined. However, among the worst inconsistencies are the relation between CB and MA. To give a recent example, in May 2022 there have been two MA announcements so far with total volumes of well over 1000 MWh. However, there has been zero volume cleared according to CB. It is unclear to us what happened to these announcements, why no volume was cleared, or whether the volume was cleared but just not reported. Similar situations are encountered all across the available data and complicate any investigation thereof. Another inconsistency is that MA volumes are sometimes negative in the machine-readable data source, which should be theoretically. Interestingly, those same values are reported as positive values on the human-readable front end of the website. Finally, one last problematic aspect of the GOPACS data is the lack of granular data. In particular, the price data, which is only published as a monthly aggregated mean, is a big obstacle to any model validation.

Most of the factors outlined in the previous paragraph however do not directly contribute to the model risk. Instead, they contribute indirectly by complicating the modelling that is based on these investigations. The lack of consistency, documentation and granularity of GOPACS were often the main reasons for the choice of certain assumptions that we were forced to make. It is these assumptions that contribute significantly to the overall model risk. We will outline now the key assumptions and also simplifications made in the development of the GOPACS model itself.

### 6.1.2 GOPACS model

The second part of our work was the development of a model to simulate GOPACS. This model relies on various assumptions and simplifications. First, there is the arrival-departure process. For this process, we assumed that it is stochastic, as we were unable to reliably predict the arrival or departure of an auction-based on other data. We modelled the arrivals and departures with two Bernoulli processes, simplifying the process by assuming a single rate of arrival and rate of departure. There is evidence to suggest that arrival and departure rates may change throughout the day (e.g. Figure 3.3 on page 23), which was one of the reasons the developed model has the ability to use non-stationary parameters which could be configured to reflect changing arrival or departure rates and a wide range of other configurations. Furthermore, we have also simplified the GOPACS process significantly. In our model, problem periods arrive and after some duration, they are over. In reality, auctions are first announced, then a bidding period is started, and once that ends, some time passes before the problem period starts. These were again simplifications made based on the assumption that this more complex process would not have a large impact on the model outcomes.

Besides modelling the arrival and departure times, we also model the volume and prices of GOPACS auctions which may be a greater contributor to the model risk. For the volume, we found that there may be seasonality and patterns in the volume data, however, we could not find a way to predict volume or establish a relation between the volume and other related data. For this reason, we assumed that the volume is entirely stochastic, an assumption that is likely not completely correct but the best we can do. Next, we also had difficulty establishing any relation between the volume that is announced (MA) and the volume that is to be cleared (CB). Being unable to determine their relation, we simplified the model again by making prices only dependent on one source of volume. We chose the cleared volume for multiple reasons – it is the volume that is used in determining the actual prices and is also included in the price data, and we also got a better stochastic fit for it, modelling it with a beta distribution. However, in some experiments, we treat the cleared volume like the announced volume (e.g. Section 5.1.3 on page 65), as we assume that participants make their decision of whether to participate or not based on volume. We add some randomness to the cleared volume to represent the uncertainty of the relation between announced and cleared volume.

The modelling of GOPACS prices was among the most challenging tasks encountered in this work and the various assumptions made are most likely contributing to the overall model risk. We have developed two different price processes. PP1 may give realistic prices as the distribution can be directly specified however it does not incorporate volume nor does it account for GOPACS being a sealed-bid pay-as-bid auction instead of a free market. PP1 relied on the assumption that participation in GOPACS yields the same payoff irrespective of the size of bids submitted, which is certainly incorrect. While we justified this simplification of real dynamics at first by discounting the achieved payoff based on its size, we have later resorted to developing a new price process that takes care of this issue more correctly.

PP2 was developed with the intent of more accurately modelling the auction aspect of GOPACS. However, this accuracy came at the expense of various assumptions. In particular, a key aspect of PP2 is the simulation of the bid distribution of other participants. All aspects of this distribution rely on assumptions instead of data as no data is available. Several of these assumptions may be incorrect, such as the minimum and maximum of bids that others submit, the number of other participants, whether their bids change significantly from auction to auction or stay more or less the same, and whether these bids are influenced by volume or not, or influenced by even other factors such as prices of other markets. None of these assumptions can be easily verified. Data for other flexibility markets is not trivial to find. Other Dutch markets function differently from GOPACS and the available data is both also limited and not really adequate for verification. While several of the assumptions that PP2 relies on are based on educated guesses, we are unable to validate the generated prices due to a lack of available GOPACS data.

Another critical assumption made in the development of PP2 is that there is a relation between the (announced) volume and the prices achieved in the GOPACS auction. This assumption was made based on the findings in Section 3.4.3 on page 39. These findings led us to outline some stylised facts that we hypothesize to be inherent to GOPACS prices, which in turn we used to validate our model. However, these stylised facts, mentioned in Section 5.1.3 on page 65, may not really exist in a larger data sample. These stylised facts contribute to the model risk.

### 6.1.3 Real option valuation

The last part of our work is the RO valuation of a flexible asset, which is done as an option pricing problem in which the underlying is the developed GOPACS model. This option pricing framework was inspired by literature in which similar approaches were used for the valuation of various energy assets. In all of the literature that we encountered, however, the underlying of these options was a free market. At first, we assumed the same for simplicity, which led us to the development of the European option. However, this European option relies on the assumption that the underlying is represented by some price $S$ that we as participant can actually achieve, which is not necessarily the case with the sealed-bid pay-as-bid auction model from GOPACS. Not being able to justify this assumption, we developed the digital option approach and focus our investigation on that.

Another simplification we do is that we assume that the DA price, the GOPACS price, and the strike price are not independent processes, but instead all deviations from the DA price. Already in our GOPACS model, we assumed that GOPACS prices are deviations from DA prices which is why we model the GOPACS price difference and not the price itself. The RO valuation builds on this assumption, as this GOPACS price difference is used in the formulation of the option problem. Specifically, we assume that the GOPACS price for a problem period, the DA price achieved the day before for the same time, and the cost of participation (i.e. the fuel consumed or goods or services produced), given by the strike price, are all related to each other on average. While this may certainly be the case on a large scale, we are not so sure if this assumption really holds up on a short scale, and this aspect of our option pricing framework may contribute to the model risk.

Another assumption we do is that our participation in GOPACS has no influence on the participation patterns of other participants. We do acknowledge that our participation decreases the prices achieved in the particular auction we participate in. However, we do assume that continuous participation by us, no matter the capacity of the asset respectively the quantity of volume we submit, does not influence the bidding behaviour of other participants. This assumption was taken to simplify the price dynamics, as otherwise, it would be necessary to connect the option pricing framework more deeply with the price process used. Moreover, even if we assume that our participation of others, another problem would be to define how it would influence their bidding behaviour, which is certainly no simple task to solve and would render our option pricing approach significantly more complex. However, we do have an idea of how this could be acknowledged through the use of an agent-based model (ABM) which we will cover in Section 6.3 on the next page.

## 6.2 Other risks and uncertainties

Besides the model risk, there are also other risks attached to our approach. One of them is market risk, which may be defined as "the risk of a change in the value of a financial position or portfolio due to changes in the value of the underlying components on which that portfolio depends" [46].

We have quantified the market risk of the FA by running simulations many times. Each such simulation represents a path that the underlying may take. The portfolio in our case is the FA. We quantify the market risk of this FA by looking at the severity of the worst 5% of outcomes, which we do using the VaR and ES risk measures. The risk numbers are shown in various plots in Chapter 5 on page 61 and also given in Table 5.1 on page 77. In addition, we can also explore the market risk by changing the parameters of our model and then analysing how the value and risk of the FA vary with respect to these changes.

In general, our model is well suited to compute and explore the market risk of various scenarios on a short-term investment horizon. However, it is less suited for long-term risk quantification. The reason for this shortcoming is that the model is stochastic and the processes are sampled from probability distributions. The longer we simulate the model, the more values we sample from these distributions. The statistics of these samples will eventually converge to some expected value as given by the law of large numbers. To give an example, the variance of outcomes of our model is higher for a one-year time horizon than for a ten-year time horizon. This behaviour does not align with the complexity of the real world, where uncertainty is typically found to increase with time. Our model can however be adapted to reproduce such real-world dynamics, by making some parameters non-stationary and making their value itself a function of some component whose uncertainty increases with time, such as e.g. a random walk where the variance increases with time.

Another type of risk that a potential owner of this FA would be exposed to is liquidity risk, which can be defined as "risk stemming from the lack of marketability of an investment that cannot be bought or sold quickly enough to prevent or minimize a loss" [46].

In the case of a FA that participates in GOPACS, this risk can not be neglected. It is related to the general uncertainty of the future of GOPACS which is higher than for other markets. GOPACS, and other flexibility markets, are a new development, and as such, they are not yet fully established. There is for example no guarantee about the future state of congestion. It may be that congestion increases or decreases in the future, and while some outcomes may be more likely than others, none can be ignored.

Assuming that congestion in the Dutch grid may not change in unforeseen ways in the future, there is still uncertainty about the evolution of GOPACS. It is possible that GOPACS may change – maybe the auction model is abandoned in favour of a limit order book or maybe more price data is published attracting a different type of participants. Various scenarios can be imagined.

The liquidity risk in face of these uncertainties is that a FA built specifically for participation in GOPACS may lose value if GOPACS becomes less liquid. It is possible to hedge against this risk by investigating the possibility of using this same FA to participate in other markets as well, such as DA, ID or ancillary services. Vice versa, the liquidity risk of any asset participating in those markets may also be hedged by considering the participation of said asset in GOPACS.

## 6.3 Future Work

### 6.3.1 ABM

As mentioned in Section 6.1 on page 79, a non-substantial amount of model risk is contributed by some of the assumptions taken in the development of the price process. In particular, the bidding behaviour of other agents and how, if at all, they are influenced by the prices they observe has been greatly simplified in our model. One idea for future work would be to introduce a third price process consisting of an agent-based model (ABM).

In this ABM, the price process, or particularly the bidding distribution, would be made up of bids of agents of different types and with different degrees of information. Various types of agents could be defined. Some agents would represent large participants that only participate if certain volume thresholds are surpassed. Other agents may be speculating, participating solely with the aim of maximising their profit. Some agents would look at existing data such as prices of other markets, while others would put more weight on the announced volumes. All of these agents then receive the announced volume and submit their bids to GOPACS. Once the auction is over, the bids are cleared and agents receive feedback, that is, they find out which of their bids, if any, were accepted, and what price they achieved. Agents may then either stop participating, change their behaviour for the next auctions or continue in the same way.

The advantages of using this ABM to generate prices are that it would allow to approach several drawbacks of our current approach. First, it would eliminate the need to specify any particular parameters for the distribution of the bids of other participants. Next, it would also make it possible to see how our participation in GOPACS may influence the behaviour of other participants. Moreover, it is also likely that the dynamics emerging from this ABM will be complex, such as e.g. price spikes arising as multiple agents left the market resulting in phase transitions. These complex dynamics could be interesting to investigate on their own. However, there are also some drawbacks to this approach that we can already outline. The ABM would introduce many new parameters that would require calibration and extensive sensitivity analysis may need to be performed. Additionally, unless additional GOPACS data were to be published, verification and validation of this ABM would also be difficult for the same reasons that it was difficult for our model.

### 6.3.2 Bidding optimisation

In Section 5.2.1 on page 68 we have presented two ways to submit multiple bids into the bidding process, which are submitting them all in one block, or submitting them linearly spaced in some range of values. However, besides these, there are many other alternatives. In particular, it would be possible to transform this bidding problem into an optimisation problem. In order to get the highest payoff, we submit some bids. Each bid can be thought of as a rectangle of width 1 and of length $B - K$, similar to the depiction in Figure 5.8 on page 71. In order to maximise the payoff, one would want to maximise the area of all submitted bids, which mathematically could be done by performing an optimisation on the integral of the function representing the bids. It would be a stochastic constrained optimisation problem, constrained by the total volume of bids to be accepted as well as the quantity and distribution of other bids, both of which are random variables in our model.

Another way to optimise the bidding process would be to consider strategies that rely on more information. We have two particular ideas. First, one could have a bidding approach in which the best value to be bid is not estimated once from historical data, but is instead continuously updated each month based on the price information published for this month by GOPACS. Another idea would be to participate in GOPACS with many small bids across a wide range of bidding values and then use the information on how many bids would be accepted to make more educated guesses on quantities of interest such as the mean and max of accepted bids. One last idea would be to always submit bids with some bidding price, and then increase this price if the last bid got accepted and decrease this price if it did not get accepted.

The advantages of both the mathematical bidding optimisation as well as the more complex strategies would be that one could extract potentially even more profit out of GOPACS. Moreover, some of the strategies may be more resilient to changes in the GOPACS market, in particular the complex strategies which rely on continuously adapting the bid value to market moves.

The drawback of diving deeper into this bidding optimisation is that all of the work would be done based on the GOPACS model and RO framework, both of which introduce some model risk. By introducing a more complex bidding approach to this model, we may amplify some of the risks already faced, and hence any potential profits achieved by a better bidding strategy may be eaten up by the introduction of additional model risk which may increase the uncertainty of actually seeing these bidding improvements in practice.

### 6.3.3 Other ideas

Apart from these two main ideas, we also have some smaller ideas that could lead to interesting future work. One of them would be the investigation of alternative ways of participating and profiting from GOPACS. First, there are the four types of flexible assets that we described in Section 4.2.1 on page 49. There are other possibilities. One would be to perform potentially direct arbitrage between GOPACS and the ID electricity market, for which one would however need more granular GOPACS auction price data to verify the validity of this approach. In addition, since this arbitrage would be relatively simple to do, not requiring any asset to participate, it is debatable if either it is not already being exploited, and if not, if there are any hurdles preventing participants from doing so. This could be interesting future work. A last alternative form of participation would be with storage assets. In particular, one can imagine a storage asset that is charged on DA and then discharged on GOPACS, or vice versa. We have given some reasons why we do not believe this strategy to work well, however, we have not explored this approach in-depth, and this may again be an area for future work.

Another interesting area for future work would be to go deeper into the forecasting of GOPACS. In Section 3.3 on page 30, we have shown our results of measuring the correlation of GOPACS with other data and also using said other data to forecast GOPACS. Our investigation of these aspects was rather limited, partially due to the constraints given by a computational science thesis. However some features looked promising, and their relation with GOPACS may be interesting to study in more depth.

Finally, one last area of future work would be a thorough validation of our approach. While we did some validation of our model, much more could be done if better data on GOPACS is published. Two specific improvements of data that could be of significant impact would be more granular price data (e.g. daily data, per auction, or even the whole bid ladder) and more consistent volume data (e.g. combining the announced and cleared volumes in one data source). Any such improvement of the data would allow a more refined validation of the developed model. Moreover, we were also unable to validate our FA valuation based on the RO approach as we have had no access to alternative valuations that one could compare against or even real numbers of the prices achievable with some asset in GOPACS. If one were to have access to such data, one could validate our approach.

## 6.4 Conclusion

A lot of different topics and areas were explored in the context of this thesis. Looking back at our work, we can identify two main contributions.

The first such contribution is an in-depth exploration, investigation, and quantification of GOPACS data revealing insights that are of relevance to anyone interested in GOPACS. This work was conducted with the aim of answering our first research question: *Who / What / When / Where / Why is GOPACS?*. We devoted an entire chapter of our work to answering these questions in as much detail as possible, with the intention of giving any reader a good overview of GOPACS. We expect this work to be relevant to anyone concerned with GOPACS or the state of congestion and flexibility in the Netherlands in general.

The second contribution of our work is the development of a RO-based valuation methodology aimed at quantifying the value of a FA participating in GOPACS. This contribution emerged from our second research question: *What is the value of a flexible asset that is participating in GOPACS?*. We do not only compute the value, but we also show how this value depends on various aspects – external factors influencing the behaviour of GOPACS, and internal decisions that one faces when participating in GOPACS auctions.

However, a valuation is not complete if one does not also present the risks associated with this approach. Still in the context of our second contribution is where we address our third research question: *What are the risks and uncertainties in the participation in GOPACS and our valuation thereof?*. We discuss the various risks faced when participating in GOPACS, ranging from model risk based on assumptions taken in the development of our model, to the general uncertainties inherent present in this newly established market. Both the asset valuation and risk quantification thereof is what we consider the second contribution of our work, and we see it as relevant again to anyone interested in participating in GOPACS, but also more generally, anyone wanting to participate in any flexibility market and looking for a valuation method for their investment.

We hope that our work can shine some light on the opportunities that GOPACS has to offer and attract new participants into this market that can help reduce congestion in, and reinforce the stability of the Dutch electricity grid.

# Bibliography


[1] René Aïd. *Electricity derivatives*. Springer, 2015.

[2] Mohamed H Albadi and Ehab F El-Saadany. "A summary of demand response in electricity markets". In: *Electric power systems research* 78.11 (2008), pp. 1989–1996.

[3] Kenneth Van den Bergh, Jonas Boury, and Erik Delarue. "The flow-based market coupling in central western europe: Concepts and definitions". In: *The Electricity Journal* 29.1 (2016), pp. 24–29.

[4] Fischer Black and Myron Scholes. "The Pricing of Options and Corporate Liabilities". In: *Journal of Political Economy* 81.3 (1973), pp. 637–654.

[5] Steven W Blume. *Electric power system basics for the nonelectrical professional*. John Wiley & Sons, 2016.

[6] Herman Bontius and John Hodemaekers. "USEF DSO Workstream report". In: *Universal Smart Energy Framework (USEF)* (2018).

[7] Alexander Boogert and Cyriel De Jong. "Gas storage valuation using a Monte Carlo method". In: *The journal of derivatives* 15.3 (2008), pp. 81–98.

[8] Alexander Boogert and Dominique Dupont. "On the effectiveness of the anti-gaming policy between the day-ahead and real-time electricity markets in The Netherlands". In: *Energy Economics* 27.5 (2005), pp. 752–770.

[9] Phelim P Boyle. "Options: A monte carlo approach". In: *Journal of financial economics* 4.3 (1977), pp. 323–338.

[10] Gert Brunekreeft. "Empirics of intraday and real-time markets in Europe: The Netherlands". In: (2015).

[11] John Burkardt. "The truncated normal distribution". In: *Department of Scientific Computing Website, Florida State University* 1 (2014), p. 35.

[12] *Capacity Allocation Congestion Management*. entso-e, 2021. URL: https://www.entsoe.eu/network_codes/cacm/ (visited on 02/03/2022).

[13] Álvaro Cartea and Carlos González-Pedraz. "How much should we pay for interconnecting electricity markets? A real options approach". In: *Energy Economics* 34.1 (2012), pp. 14–30.

[14] TenneT Customer Care Center. private email. Jan. 2022.

[15] *Challenges in the regions*. Liander. URL: https://2020.jaarverslag.alliander.com/verslagen/annual-report-2020/thevaluewecreate2/onetwohireliabililcos2/challengesintheregions2 (visited on 02/06/2022).

[16] H De Heer and W van den Reek. "USEF white paper flexibility platforms". In: *Universal Smart Energy Framework (USEF)* (2018), pp. 1–30.

[17] Shi-Jie Deng, Blake Johnson, and Aram Sogomonian. "Exotic electricity options and the valuation of electricity generation and transmission assets". In: *Decision support systems* 30.3 (2001), pp. 383–392.

[18] Theo Dronne, Fabien Roques, and Marcelo Saguan. "Local Flexibility Markets for Distribution Network Congestion-Management in Center-Western Europe: Which Design for Which Needs?" In: *Energies* 14.14 (2021), p. 4113.

[19] Abbas El Gamal. *Random Processes Lecture 7*. URL: http://isl.stanford.edu/~abbas/ee178/lect07-2.pdf (visited on 05/17/2022).

[20] *Enera website*. Enera. URL: https://projekt-enera.de (visited on 02/05/2022).

[21] *EPEX SPOT Trading*. EPEX SPOT. 2021. URL: https://www.epexspot.com/sites/default/files/download_center_files/21-03-09_Trading%20Brochure.pdf (visited on 02/03/2022).

[22] Ayman Esmat, Julio Usaola, and Mª Moreno. "A decentralized local flexibility market considering the uncertainty of demand". In: *Energies* 11.8 (2018), p. 2078.

[23] *ETPA website*. ETPA, 2021. URL: https://etpa.nl/wp-content/uploads/2016/09/Frequently_Asked_Questions.pdf (visited on 02/03/2022).

[24] Julia Frayer and Nazli Z Uludere. "What is it worth? Application of real options theory to the valuation of generation assets". In: *The Electricity Journal* 14.8 (2001), pp. 40–51.

[25] Thomas Webster Gedra and Pravin Pratap Varaiya. "Markets and pricing for interruptible electric power". In: *IEEE Transactions on power systems* 8.1 (1993), pp. 122–128.

[26] Douglas C Giancoli. *Physics: principles with applications*. Boston: Pearson, 2016.

[27] Paul Glasserman. *Monte Carlo methods in financial engineering*. Vol. 53. Springer, 2004.

[28] *GOPACS website*. GOPACS. URL: https://www.gopacs.eu (visited on 02/05/2022).

[29] Charles R Harris et al. "Array programming with NumPy". In: *Nature* 585.7825 (2020), pp. 357–362.

[30] Julia Hentschel, Hartmut Spliethoff, et al. "A parametric approach for the valuation of power plant flexibility options". In: *Energy Reports* 2 (2016), pp. 40–47.

[31] Jaroslava Hlouskova et al. "Real options and the value of generation capacity in the German electricity market". In: *Review of Financial Economics* 14.3-4 (2005), pp. 297–310.

[32] Michael Hsu. "Spark spread options are hot!" In: *The Electricity Journal* 11.2 (1998), pp. 28–39.

[33] *IDCONS Productvoorwaarden*. GOPACS. Feb. 2021. URL: https://www.gopacs.eu/documentatie/ (visited on 01/18/2022).

[34] Xiaolong Jin, Qiuwei Wu, and Hongjie Jia. "Local flexibility markets: Literature review on concepts, models and clearing methods". In: *Applied Energy* 261 (2020), p. 114387.

[35] B Johnson and G Barz. *Selecting stochastic processes for modelling electricity prices, chapter 1*. 1999.

[36] Guolin Ke et al. "Lightgbm: A highly efficient gradient boosting decision tree". In: *Advances in neural information processing systems* 30 (2017), pp. 3146–3154.

[37] Jussi Keppo and Hao Lu. "Real options and a large producer: the case of electricity markets". In: *Energy Economics* 25.5 (2003), pp. 459–472.

[38] Hendrik Kondziella and Thomas Bruckner. "Flexibility requirements of renewable energy based electricity systems–a review of research results and methodologies". In: *Renewable and Sustainable Energy Reviews* 53 (2016), pp. 10–22.

[39] Kalimuthu Krishnamoorthy. *Handbook of statistical distributions with applications*. Chapman and Hall/CRC, 2006.

[40] Dominik Kryzia, Michał Kopacz, and Katarzyna Kryzia. "The valuation of the operational flexibility of the energy investment project based on a gas-fired power plant". In: *Energies* 13.7 (2020), p. 1567.

[41] Ioannis Lampropoulos et al. "Analysis of the market-based service provision for operating reserves in the Netherlands". In: *2012 9th International Conference on the European Energy Market*. IEEE. 2012, pp. 1–8.

[42] Francis A Longstaff and Eduardo S Schwartz. "Valuing American options by simulation: a simple least-squares approach". In: *The review of financial studies* 14.1 (2001), pp. 113–147.

[43] *Machine Announcements Implementation*. GOPACS. Aug. 2021. URL: https://www.gopacs.eu/documentatie/ (visited on 01/18/2022).

[44] Eric Martinot. “Grid integration of renewable energy: flexibility, innovation, and experience”. In: *Annual Review of Environment and Resources* 41 (2016), pp. 223–251.

[45] Scott Mathews, Vinay Datar, and Blake Johnson. “A practical method for valuing real options: The Boeing approach”. In: *Journal of Applied Corporate Finance* 19.2 (2007), pp. 95–104.

[46] Alexander J McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative risk management: concepts, techniques and tools-revised edition*. Princeton university press, 2015.

[47] N Metropolis. “The beginning”. In: *Los Alamos Science* 15 (1987), pp. 125–130.

[48] Hamed Mohsenian-Rad. “Optimal bidding, scheduling, and deployment of battery systems in California day-ahead energy market”. In: *IEEE Transactions on Power Systems* 31.1 (2015), pp. 442–453.

[49] John Moriarty and Jan Palczewski. “Real option valuation for reserve capacity”. In: *European Journal of Operational Research* 257.1 (2017), pp. 251–260.

[50] Massimo Morini. *Understanding and Managing Model Risk: A practical guide for quants, traders and validators*. John Wiley & Sons, 2011.

[51] Johnathan Mun. *Modeling risk: Applying Monte Carlo simulation, real options analysis, forecasting, and optimization techniques*. Vol. 347. John Wiley & Sons, 2006.

[52] Stewart C Myers. “Determinants of corporate borrowing”. In: *Journal of financial economics* 5.2 (1977), pp. 147–175.

[53] Selvaprabu Nadarajah, François Margot, and Nicola Secomandi. “Comparison of least squares Monte Carlo methods with applications to energy real options”. In: *European Journal of Operational Research* 256.1 (2017), pp. 196–204.

[54] *NODES website*. NODES. URL: https://nodesmarket.com (visited on 02/05/2022).

[55] Pol Olivella-Rosell et al. “Local flexibility market design for aggregators providing multiple flexibility services at distribution network level”. In: *Energies* 11.4 (2018), p. 822.

[56] Shmuel S Oren. “Integrating real and financial options in demand-side electricity contracts”. In: *Decision Support Systems* 30.3 (2001), pp. 279–288.

[57] *Piclo Flex website*. Piclo Flex. URL: https://picloflex.com (visited on 02/05/2022).

[58] Svetlozar Rachev. *Average value-at-risk Lecture 7*. URL: https://statistik.ets.kit.edu/download/doc_secure1/7_StochModels.pdf (visited on 05/24/2022).

[59] Ariana Ramos et al. “Realizing the smart grid’s potential: Defining local markets for flexibility”. In: *Utilities Policy* 40 (2016), pp. 26–35.

[60] Rijksoverheid. *Hoofdstuk 9 Netcode elektriciteit*. Dec. 2021. URL: https://wetten.overheid.nl/jci1.3:c:BWBR0037940&hoofdstuk=9&z=2021-12-14&g=2021-12-14 (visited on 01/30/2022).

[61] Apurba Sakti et al. “Enhanced representations of lithium-ion batteries in power systems models and their effect on the valuation of energy arbitrage applications”. In: *Journal of Power Sources* 342 (2017), pp. 279–291.

[62] Jonathan A Schachter and Pierluigi Mancarella. “Demand response contracts as real options: a probabilistic evaluation framework under short-term and long-term uncertainties”. In: *IEEE Transactions on Smart Grid* 7.2 (2015), pp. 868–878.

[63] Jonathan Alexandre Schachter and Pierluigi Mancarella. “A critical review of Real Options thinking for valuing investment flexibility in Smart Grids and low carbon energy systems”. In: *Renewable and Sustainable Energy Reviews* 56 (2016), pp. 261–271.

[64] Hans Schermeyer, Claudio Vergara, and Wolf Fichtner. “Renewable energy curtailment: A case study on today’s and tomorrow’s congestion management”. In: *Energy Policy* 112 (2018), pp. 427–436.

[65] Tim Schittekatte and Leonardo Meeus. “Flexibility markets: Q&A with project pioneers”. In: *Utilities policy* 63 (2020), p. 101017.

[66] Osman Sezgen, CA Goldman, and P Krishnarao. “Option value of electricity demand response”. In: *Energy* 32.2 (2007), pp. 108–119.

[67] Anna Stawska et al. "Demand response: For congestion management or for grid balancing?" In: *Energy Policy* 148 (2021), p. 111920.

[68] Bjarne Steffen and Christoph Weber. "Optimal operation of pumped-hydro storage plants with continuous time-varying power prices". In: *European Journal of Operational Research* 252.1 (2016), pp. 308–321.

[69] Matt Thompson, Matt Davison, and Henning Rasmussen. "Natural gas storage valuation and optimization: A real options application". In: *Naval Research Logistics (NRL)* 56.3 (2009), pp. 226–238.

[70] Lenos Trigeorgis. *Real options: Managerial flexibility and strategy in resource allocation*. MIT press, 1996.

[71] Chung-Li Tseng and Graydon Barz. "Short-term generation asset valuation: a real options approach". In: *Operations Research* 50.2 (2002), pp. 297–310.

[72] Orlando Valarezo et al. "Analysis of new flexibility market models in Europe". In: *Energies* 14.12 (2021), p. 3521.

[73] José Villar, Ricardo Bessa, and Manuel Matos. "Flexibility products and markets: Literature review". In: *Electric Power Systems Research* 154 (2018), pp. 329–340.

[74] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[75] *What is Generation Capacity*. U.S. Department of Energy. URL: https://www.energy.gov/ne/articles/what-generation-capacity (visited on 02/05/2022).

[76] Dimitrios Zafirakis et al. "The value of arbitrage for energy storage: Evidence from European electricity markets". In: *Applied energy* 184 (2016), pp. 971–986.

# Appendix A

# GOPACS prediction feature importances

The model we developed for GOPACS prediction, presented in Section 3.3 on page 30, uses LightGBM. LightGBM provides an estimate of feature importance, which is defined as the number of times a particular feature is used by the model, i.e. how many times it appears in the decision trees that constitute the model. A higher number means that a feature was used a lot, which means the feature plays an important role in the achievement of the model output. We have plotted this feature importances of the four models that we developed, which are the MA TP problem period classification model (Figure A.1 on the following page), the CB problem period classification model (Figure A.2 on page 92), the MA TP volume regression model (Figure A.3 on page 93) and the CB volume regression model (Figure A.4 on page 94).

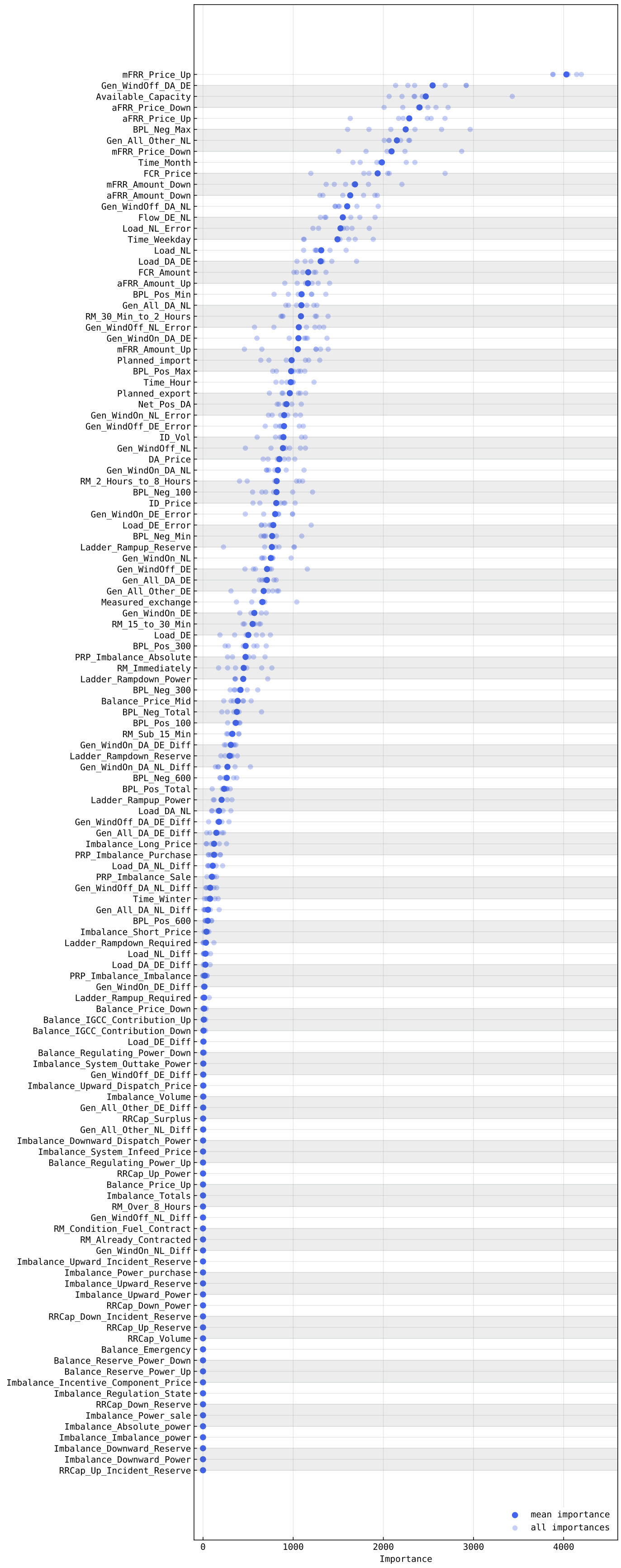



FIGURE A.1: The feature importances of the MA TP problem period classification model as given by LightGBM

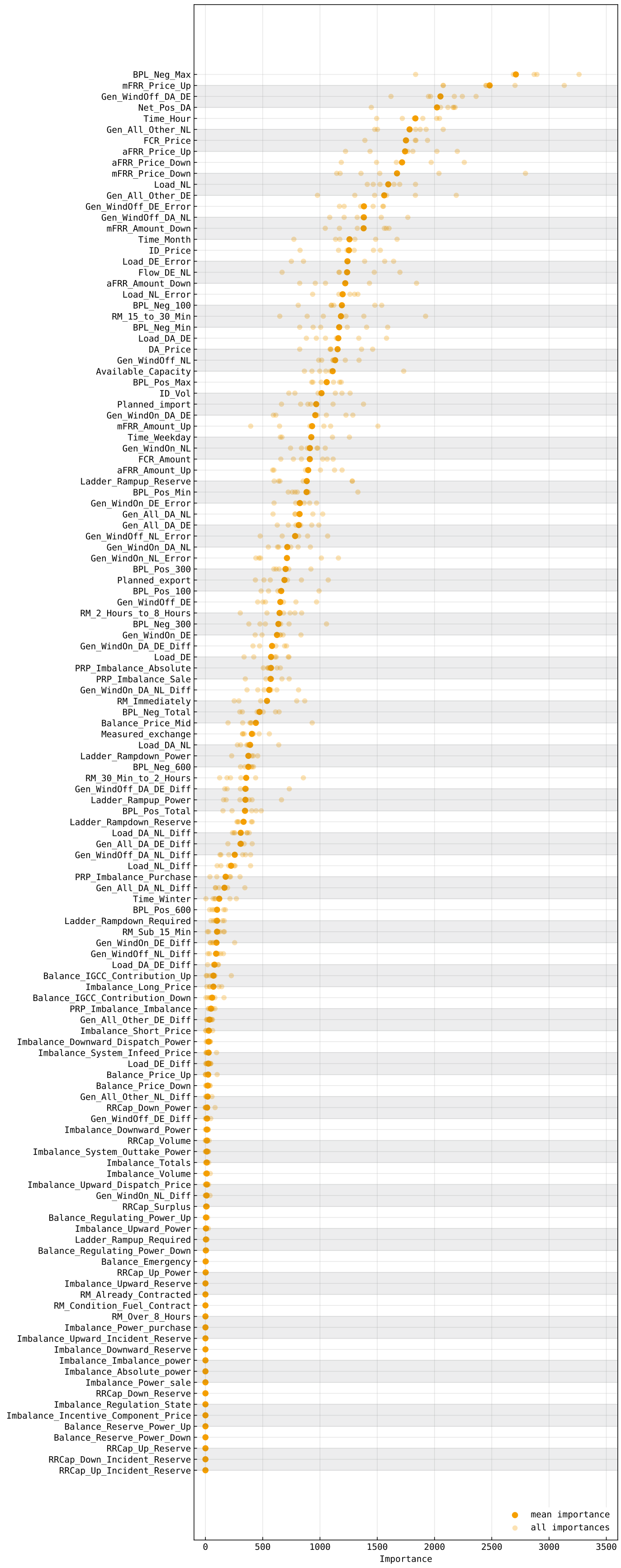



FIGURE A.2: The feature importances of the CB problem period classification model as given by LightGBM

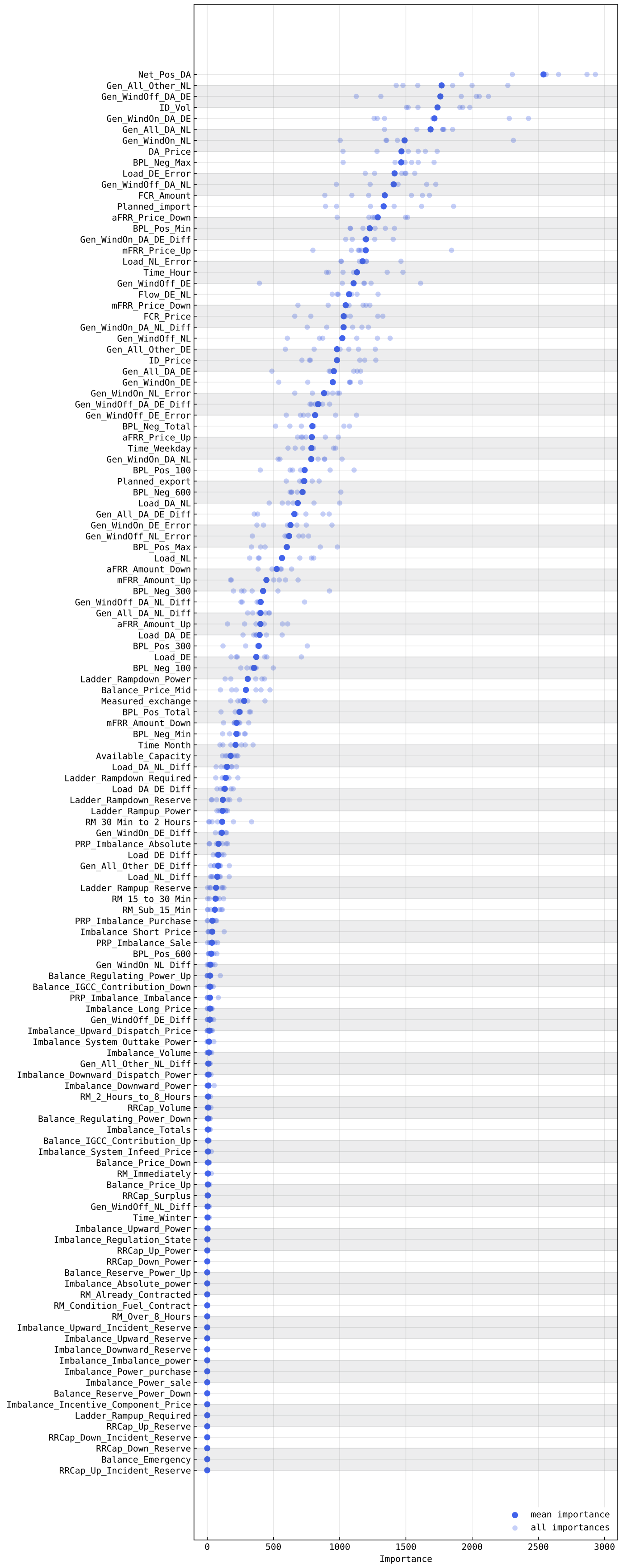



Figure A.3: The feature importances of the MA TP volume regression model as given by LightGBM

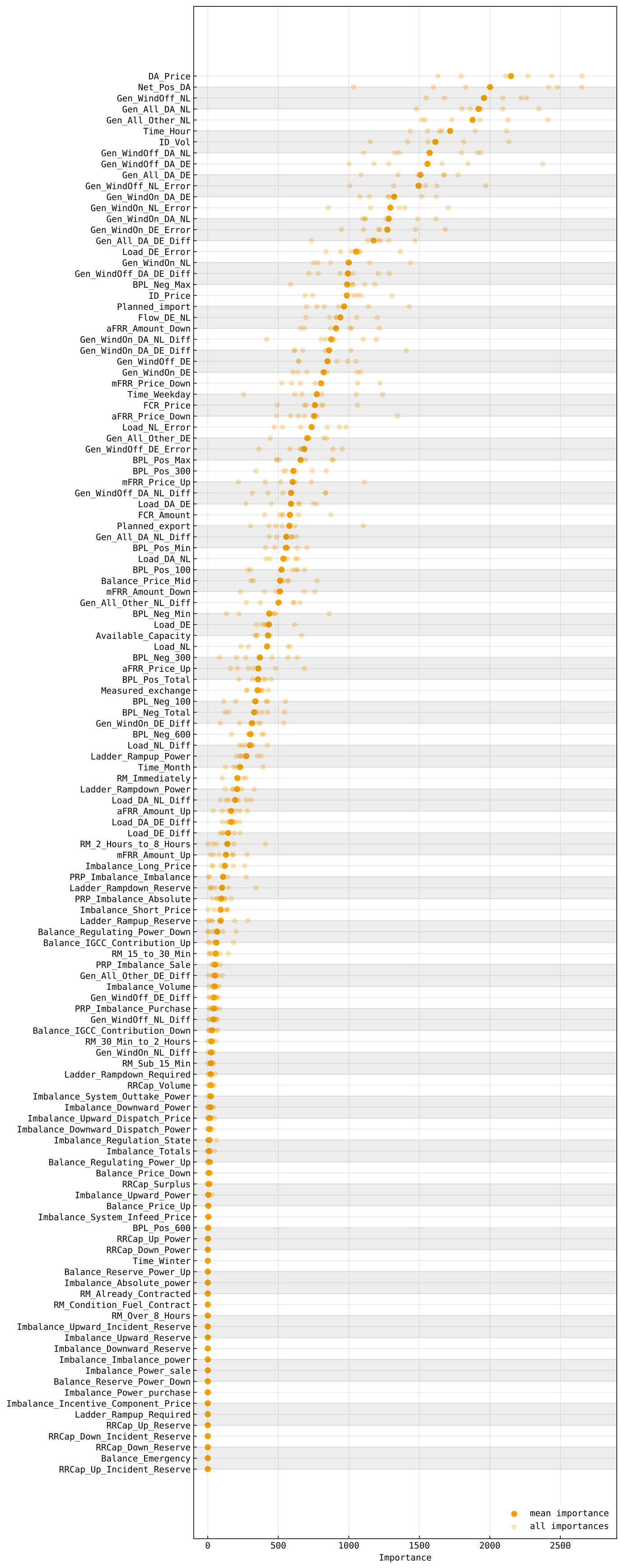


Figure A.4: The feature importances of the CB volume regression model as given by LightGBM