energies

Operation, Regulation and Planning of Power and Natural Gas Systems

Edited by
Javier Reneses and Antonio Bello

Printed Edition of the Special Issue Published in *Energies*

www.mdpi.com/journal/energies

MDPI

Operation, Regulation and Planning of Power and Natural Gas Systems

Operation, Regulation and Planning of Power and Natural Gas Systems

Editors

Javier Reneses
Antonio Bello

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade • Manchester • Tokyo • Cluj • Tianjin

Editors

Javier Reneses
Institute for Research in Technology, Universidad Pontificia Comillas
Spain

Antonio Bello
Institute for Research in Technology, Universidad Pontificia Comillas
Spain

Editorial Office
MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Energies* (ISSN 1996-1073) (available at: https://www.mdpi.com/journal/energies/special_issues/Power_Natural_Gas_Systems).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

LastName, A.A.; LastName, B.B.; LastName, C.C. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-03943-821-1 (Hbk)
ISBN 978-3-03943-822-8 (PDF)

Contents

About the Editors

Javier Reneses is Senior Research Associate Professor in the Institute for Research in Technology (IIT) at the Universidad Pontificia Comillas in Madrid. He was a visiting scholar at the Berkeley Lab from 2016 to 2019. He has headed more than 80 research and consultancy projects and has participated in more than 150. He has worked and lectured extensively on the operation, planning, and regulation of energy systems, and particularly on medium-term operation of electricity markets, regulation of electric distribution business, tariff design, and natural gas markets. He has served as a consultant on these topics for governments, international institutions, industrial associations, and utilities in numerous countries. Dr. Reneses has published more than 100 papers and book chapters published in national and international journals and conference proceedings. He also has been teaching Electric Distribution Business at the Universidad Pontificia Comillas for the master's degree in the Electric Power Industry as well as Statistics and Energy and Sustainability at the School of Engineering for several years. He obtained his Electrical Engineering degree in 1996 from the School of Engineering (ICAI) of the Universidad Pontificia de Comillas, Madrid, Spain. He received his doctoral degree in May 2004 for the thesis "Mid-Term Operation of Electricity Markets". He obtained his degree in Mathematics in 2005 from the Universidad Nacional de Educación a Distancia (UNED), Madrid.

Antonio Bello is Research Assistant Professor at the Institute for Research in Technology (IIT) of ICAI School of Engineering at Comillas Pontifical University. Antonio holds an Industrial Engineering degree, an MSc degree in Electric Power Systems, and a Ph.D. degree in Electrical Engineering from the Comillas Pontifical University-ICAI, with a dissertation about probabilistic forecasting of electricity prices, which he defended in 2016. Throughout his career, he has participated in over 70 research and consultancy projects dealing with operation, simulation models, planning of electricity and gas markets, energy forecasting, and risk management support. Besides his research activities, he teaches time series analysis and forecasting in the energy sector. He has published numerous papers in national and international journals and conference proceedings on these topics. Antonio Bello has been a visiting researcher at London Business School (London, United Kingdom).

energies

MDPI

Article

Expansion Planning Method of the Industrial Park Integrated Energy System Considering Regret Aversion

Haokai Xie [1,*], Pu Zhao [2], Xudong Ji [2], Qun Lin [2] and Lianguang Liu [1]

[1] School of Electrical and Electronic Engineering, North China Electric Power University, Beijing 102206, China
[2] State Grid Wenzhou Power Supply Company, Wenzhou 325000, China
* Correspondence: xiehaokai@ncepu.edu.cn; Tel.: +86-132-6120-2892

Received: 3 October 2019; Accepted: 23 October 2019; Published: 27 October 2019

Abstract: Industrial parks have various sources and conversion forms of energy. The many uncertainties in the planning of industrial park integrated energy systems (IPIES) pose a great risk of regret in planning schemes; thus, an expansion planning method for an IPIES, considering regret aversion, is proposed. Based on comprehensive regret value consisting of min–max regret aversion and the min average regret value, the method optimizes the comprehensive cost of the expansion planning scheme in IPIES under different natural gas price fluctuation scenarios, including costs of construction, operation and maintenance, and environmental protection. A multi-stage expansion planning scheme and typical daily operation plans under multiple natural gas price fluctuation scenarios of the IPIES in an economic and technological development zone in southeast China are used to demonstrate the validity of the method. The results show that, compared with a traditional planning method based on expectation, the proposed expansion planning method could reduce the maximum regret value by 14% on average, and greatly reduces the risk of decision-making regret by up to 18%. At the same time, the influence of natural gas price on expansion planning of the IPIES is discussed.

Keywords: industrial park integrated energy system; expansion planning; natural gas price uncertainty; regret aversion; min–max regret value

1. Introduction

The integrated energy system (IES) integrates energy production, conversion, storage, and consumption [1,2]. It is an important development trend of energy technology to achieve coordinated and complementary optimization of multiple energy sources in the future [3–5]. Industrial parks with intensive demand for electricity, steam, cold, and heat energy are typical application scenarios for IES. How to plan industrial park integrated energy systems (IPIES) is an important issue in current research [6,7]. In order to optimize the structure and capacity of the IES, domestic and foreign scholars have proposed various models, algorithms, and planning objectives. Geidl et al. [8] proposed the concept of an energy hub (EH) and established a planning model for an electric power and natural gas system with the objective function of minimum energy loss in the EH. Zhang et al. [9] considered a variety of combined heat and power (CHP) generation units, designing CHP units with the objective of system economics and environmental performance. Zhang et al. [10] decomposed the planning model into two aspects: Investment and operation feasibility of a power system or natural gas system. Zhao et al. [11] proposed a three-level collaborative global optimization method for a combined cooling, heating, and power (CCHP) system. These research and planning methods were mainly designed for a system plan in a specific state.

Based on EH, Zhou et al. [12] proposed a collaborative expansion optimization configuration method for a renewable power system and a natural gas system. Considering the topological constraints of grid and gas networks, a regional integrated energy system expansion planning model based on CCHP was proposed in [13], and the economic scheduling strategy of the system was analyzed by the scenario method. Considering the value chain of natural gas, the long-term, multi-regional, and multi-stage expansion planning of a gas-electric coupling system was studied in reference [14]. The authors of [15,16] proved that flexible expansion planning can better handle uncertainties, compared to traditional lowest-cost planning methods. Adaptation cost is defined as the additional investments required for a proposed plan, if an unexpected load growth happens. Qiu et al. [17] established a joint expansion planning model for natural gas networks and power grids with the objective of maximizing social benefits. The Monte Carlo simulation was applied to create scenarios that simulate random system characteristics in [18]. An extended planning model based on two-stage stochastic optimization was established to realize the construction planning of natural gas and power facilities under uncertainty of demand growth in [19].To deal with the uncertainty of renewable energy, the scenario method was used to deal with the wind power and energy storage and load, and the capacity of the internal device in the EH was configured in [20]. Robust optimization was used to obtain an optimized solution that is immune to the effects of all possible wind power realizations within the uncertainty interval in [21,22] Cesena et al. [23] proposed a unified planning and scheduling method to assess the flexibility of system investment and operation under long-term uncertainty. The proposed approach in [24] reflected real-options thinking borrowed from finance, and had been cast as a stochastic mixed-integer linear program. These integrated energy expansion planning studies did not consider the need for policymakers to avoid the risk of regret in planning schemes together with the lowest cost.

Regret is an emotion that affects decision-making behavior. When decision-makers face a variety of schemes, one scheme is selected and compared with other unselected schemes. When the uncertainty causes the actual situation to be different from the expected, resulting in the profit of the selected scheme being smaller than one or more of the unselected schemes, a feeling of regret in the decision-maker is generated [24,25]. Savage et al. [26] proposed a model based on minimum-maximization regret, which shows that decision-makers will choose a decision with a minimal regret value from the decision plan that maximizes regret. In [27], regret is considered as one of the objectives in a multi-objective optimization framework. By applying the min-max regret criterion, model obtained a solution that minimizes the worst-case regret over all possible scenarios while ensuring system robustness [28]. The theory of regret has been applied in the study of consumer behavior in economics, travel path planning, etc. [29,30]. In [31], the min-max regret criterion is considered for the unit commitment problem. Compared with the min-max cost criterion, it is concluded that min-max regret outperforms min-max cost for certain unit commitment problems. Min-max cost and min-max regret have been proposed to address wind power generation uncertainties in [32]. Both criteria provide good upper bounds for the total costs under scenarios contained in an uncertainty set. The min-max cost criterion provides a smaller upper bound while min-max regret has higher variability. Depending on the characteristics of uncertainty sets and the preference of decision-makers, different models outperform each other under different situations. In the IPIES, the prediction error caused by uncertainty will cause decision-makers to pay excessive construction and operation costs for the system, which will also cause the decision-makers regret. In this paper, there are many types of comprehensive energy sources and conversion forms in industrial parks, such as natural gas, electric, heat, cold, and steam. In view of the various sources and conversion forms of energy in industrial parks as well as the large amount of uncertainty posing a great risk of regret in planning schemes, an expansion planning method of the IPIES considering regret aversion is proposed. The method deals with uncertainty using the scenario method, sets the regret value as the main indicator, and the comprehensive regret value, which considers the min-max regret value with its distribution and average regret value together, as the objective function.

2. IPIES Structure and Expansion Planning Method Model

Referring to [8,11,14,23,33,34], the structure and energy flow of the IPIES studied in this paper is shown in Figure 1. The system is connected to a steam network provided by an external large thermal power plant. The IPIES includes four parts: (1) Supply part: Power grid, natural gas network, steam network, photovoltaic (PV); (2) conversion part: Micro turbine (MT), heat recovery device (HR), electric boiler (EB), gas boiler (GB), heat pump (HP), electric chiller (EC), heat exchanger (HC), absorption chiller (AC); (3) storage part: Battery (BAT), steam heat storage (HS), cold energy storage (CS); and 4) load part: Electrical loads, steam loads, heat loads, cold loads, gas loads.

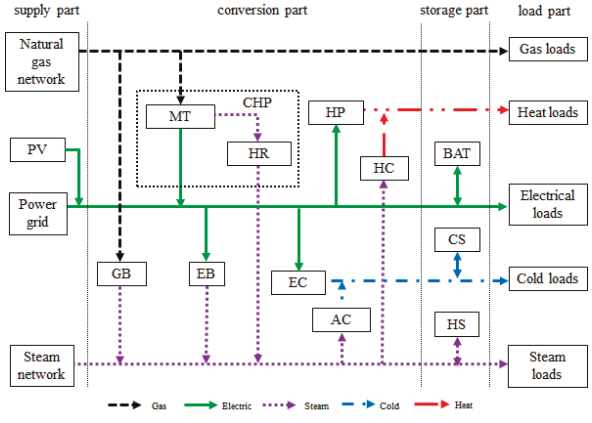


Figure 1. Structure and energy flow of industrial park integrated energy system (IPIES).

Based on the structure and energy flow of the IPIES, the expansion planning method of the IPIES considering regret aversion proposed in this paper is shown in Figure 2. It includes three layers: The stage scenario analysis layer, the expansion planning layer, and the regret aversion layer.

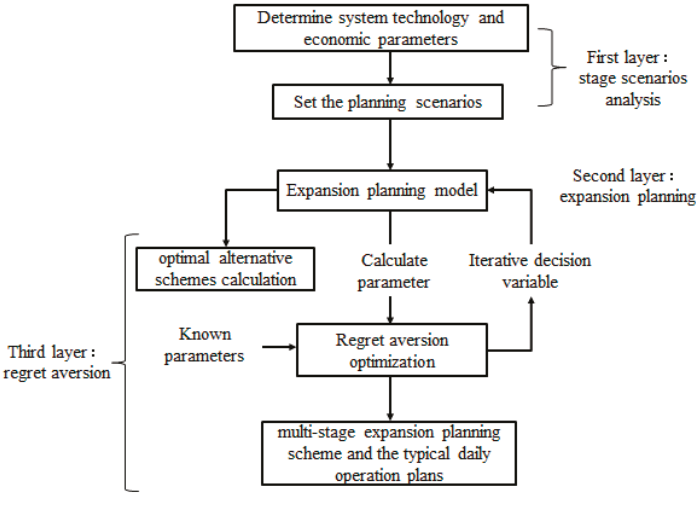


Figure 2. Expansion planning method of IPIES considering regret aversion.

In the stage scenarios analysis layer, the method mainly studies the typical daily load scenarios of the system and the natural gas price fluctuation scenarios.

In the expansion planning layer, the method sets the capacity and the typical daily operating power corresponding to device capacity as the variables. The method establishes an expansion planning model by taking the net present value of the comprehensive cost, including costs of construction, operation and maintenance, and environmental protection, as the objective function.

In the regret aversion layer, the method firstly calculates the optimal alternative schemes under different natural gas price fluctuation scenarios with the lowest comprehensive cost. Then, the method sets a min–max regret value and the lowest average regret value between the final planning scheme and the optimal alternative schemes as the objective function to optimize the device capacity within the final multi-stage expansion planning scheme and the typical daily operation plans under multiple natural gas price fluctuation scenarios.

2.1. Stage Scenarios Analysis Layer

Firstly, the scenario analysis method in [35] is used to deal with the volatility and randomness of each energy load and photovoltaic (PV) unit output power. The number of typical day scenarios after the load scenes and the lighting scenes are reduced is M, and each typical day scenario, m, has D_m days in a whole year. In order to meet the normal operation of the system at all times, an additional daily limit scenario with a constant zero PV output power is added. In view of the growth of energy load during the planning stages, the paper refers to the multi-stage planning method of a power system [36], introduces the continuous load curve to describe the medium and long-term load growth expectations, and divides the load level in the planning period into several horizontal sections, though the simplification will affect the accuracy of the model to some extent. A typical day scenario's load characteristic curves at different load levels can be obtained by equal ratio changes.

In addition, considering the price of the system device will decrease during the planning stages with the development of science and technology, and will stabilize after the technology matures [37], a piecewise exponential function is used to represent the dynamic change in device prices [38]:

$$c^{I,k,y} = \begin{cases} c_0^{I,k}(1+g_k)^y & 1 < y < Y_k^c \\ c_0^{I,k}(1+g_k^c) & Y_k^c < y < Y \end{cases}, \quad (1)$$

where $c^{I,k,y}$ is the construction price of the device k in year y; g_k is the price correction coefficient of the device k; g_k^c is the critical price reduction factor; Y_k^c is the time for the device k to reach the critical price; and Y is the operating period of the IPIES.

The electricity market is still not perfect in China, but electricity prices are relatively stable due to policy decisions; natural gas prices, however, will be affected by changes in global trade prices and domestic supply and demand factors, and there will be greater uncertainty in prices over time and space. In the past, natural gas was originally developed as a replacement for traditional fuel, and its pricing was linked to other energy such as oil and the oil-indexed gas imports in China accounted for the majority [39–41]. However, with the changes in the international natural gas supply and demand pattern and the continuous reform of China oil and gas market, natural gas, especially liquefied natural gas (LNG), is gradually becoming an independent energy product. In 2018, China LNG imports accounted for 60% of total natural gas imports [42]. LNG breaks the restrictions on natural gas transmission and trade between regions, greatly enhancing the transmission and impact range of natural gas prices. The 2019 Wholesale Gas Price Survey shows that Henry-Hub priced US LNG exports continued rising and there is more gas price convergence amongst countries since the global gas market and market-related pricing [43]. Figure 3 shows the China LNG ex-factory price index given by the Shanghai Petroleum and Natural Gas Exchange, reflecting the price trend of LNG in the domestic market [42]. The Shanghai Petroleum and Natural Gas Exchange, which was officially launched in 2015, opened the situation that China natural gas prices are determined by competition between supply and demand. It can be seen from Figure 3 that China's LNG spot price also has large fluctuations in different periods.

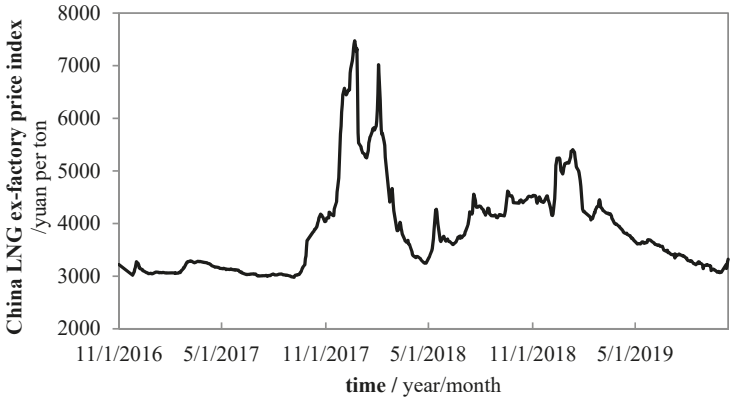


Figure 3. China liquefied natural gas ex-factory price index given by the Shanghai Petroleum and Natural Gas Exchange.

Changes in natural gas prices will directly affect the operating costs of the IPIES, which in turn will affect the economic operation of the system and the expansion of the plan; the greater fluctuation, the higher the impact on the plan. Seasonal or yearly consideration of natural gas price fluctuations during the planning cycle will lead to excessive calculation of the model. The method uses stage scenario analysis techniques to analyze the uncertainty of natural gas prices during expansion planning. The method takes the natural gas price in the initial planning stage as the benchmark price, and considers that there may be fluctuations, η, in the price of the subsequent stage compared to the benchmark price; that is, it may rise, fall, or remain unchanged from the benchmark price. In the resulting natural gas price fluctuation scenario set, S, a planning stages corresponds to 3^{a-1} natural gas price fluctuation scenarios, and the scenario probability corresponding to the natural gas price fluctuation scenario, s, is π^s.

2.2. Expansion Planning Layer

2.2.1. Objective Function

The expansion planning layer takes the net present value of the comprehensive cost, C_s^{COM}, of multi-stage planning in a natural gas price fluctuation scenario, $s \in$ S, as the objective function, including the system construction cost, C_s^I, operation cost, C_s^O, maintenance cost, C_s^M, and environmental protection cost, C_s^{ENV}:

$$C_s^{COM} = C_s^I + C_s^O + C_s^M + C_s^{ENV}, \tag{2}$$

$$C_s^I = \sum_a \frac{1}{(1+\lambda)^{T(a-1)}} \cdot \left[\mathbf{c_a^{I,k}} \cdot \left(\mathbf{W_{s,a}^k} - \mathbf{W_{s,a\text{-}1}^k}\right)^{\mathbf{T}} + \mathbf{c}_a^{I,T} \cdot \left(I_a^{des} - I_{a-1}^{des}\right)\right], \tag{3}$$

$$C_s^O = \sum_y \delta_y \cdot \sum_m D_m \cdot \sum_t \left(c_t^E \cdot P_{s,a,m,t}^{SYS} + c_{s,a}^G \cdot G_{s,a,m,t}^{SYS} + c^S \cdot S_{s,a,m,t}^{SYS}\right), \tag{4}$$

$$\begin{aligned} C_s^O = \sum_y \delta_y \cdot \sum_m D_m \cdot \sum_t \Big(& c^{M,BAT} \cdot \left|P_{s,a,m,t}^{BAT}\right| + c^{M,HS} \cdot \left|S_{s,a,m,t}^{HS}\right| + c^{M,CS} \cdot \left|C_{s,a,m,t}^{CS}\right| \\ & + \mathbf{c^{M,k1}} \cdot \mathbf{Q_{s,a,m,t}^{k1}}^{\mathbf{T}}\Big), \end{aligned} \tag{5}$$

$$C_s^{ENV} = \sum_y \delta_y \cdot \sum_m D_m \cdot \sum_t \left(\gamma^E \cdot P_{s,a,m,t}^{SYS} + \gamma^G \cdot G_{s,a,m,t}^{SYS} + \gamma^S \cdot S_{s,a,m,t}^{SYS}\right), \tag{6}$$

$$\delta_y = \frac{1}{(1+\lambda)^{y-1}}, \tag{7}$$

where a is the planning stage, and each stage has T years; y is the operation year of IPIES; t is 24 intraday hours; δ_y is the year discount rate; λ is the annual discount rate; $\mathbf{c_a^{I,k}}$ is the unit construction cost matrix of each device, including BAT, HS, CS, PV, CHP, GB, EB, HP, EC, HC, and AC; $\mathbf{W_{s,a}^{k}}$ is the total capacity matrix of each device in stage a under scenario s; $c_a^{I,T}$ is the power transmission capacity expansion cost; I_a^{des} is the 0–1 mark for the power transmission expansion status, 1 after expansion, 0 before expansion; c_t^E is the electricity price at time t; c_a^G is the price for a unit kWh of energy natural gas in stage a under scenario s; c^S is the price for a unit kW·h of energy steam, and the unit kW·h energy price can be calculated by the unit cubic meter price or the unit steaming price and the low calorific value of the energy; $P_{s,a,m,t}^{SYS}$, $G_{s,a,m,t}^{SYS}$, and $S_{s,a,m,t}^{SYS}$ are the electricity, gas, and steam power, respectively, that the system interacts with in the external network at typical day m, time t in stage a under scenario s; $c^{M,BAT}$, $c^{M,HS}$, and $c^{M,CS}$ are the unit maintenance cost of BAT, HS, and CS, respectively; $\mathbf{c^{M,k1}}$ is the unit maintenance cost matrix of the device except for energy storage in IPIES; $P_{s,a,m,t}^{BAT}$, $S_{s,a,m,t}^{HS}$, and $C_{s,a,m,t}^{CS}$ are the power exchange of the energy storage device for BAT, HS, and CS, respectively; $\mathbf{Q_{s,a,m,t}^{k1}}$ is the operating power matrix of each device except for energy storage in IPIES; and γ^E, γ^G, and γ^S are the environmental cost of emissions from unit electricity, gas, and steam power, respectively, and can be calculated from the environmental value of the pollutants discharged per kW·h of energy.

2.2.2. Constraints

The constraints in the expansion planning layer include expansion constraints, load part constraints, supply part constraints, conversion part constraints and storage part constraints.

(1) Expansion constraints

Each stage can expand the capacity of each device in the IPIES or maintain the configuration of the previous stage. Decommissioning is required for the life of the device to expire:

$$W_{s,a}^{k} \geq W_{s,a-1}^{k} - W_{s,a}^{k,out}, \tag{8}$$

$$I_a^{des} \geq I_{a-1}^{des}, \tag{9}$$

$$W_{s,a}^{k,out} = W_{s,a-n_k}^{k} - W_{s,a-n_k-1}^{k}, \tag{10}$$

where $W_{s,a}^{k}$ is the capacity of the device k in stage a under scenario s; $W_{s,a}^{k,out}$ is the capacity of the device k to be decommissioned in stage a under scenario s; and n_k is the number of planned stages that device k can serve; when $a \leq 0$, $W_{s,a}^{k}$ is 0.

(2) Load part constraints:

$$P_{s,a,m,t}^{PV} + P_{s,a,m,t}^{CHP} + P_{s,a,m,t}^{SYS} = P_{s,a,m,t}^{LD} + P_{s,a,m,t}^{EB} + P_{s,a,m,t}^{EC} + P_{s,a,m,t}^{EH} + P_{s,a,m,t}^{BAT}, \tag{11}$$

$$S_{s,a,m,t}^{CHP} + S_{s,a,m,t}^{EB} + S_{s,a,m,t}^{GB} + S_{s,a,m,t}^{SYS} = S_{s,a,m,t}^{LD} + S_{s,a,m,t}^{HC} + S_{s,a,m,t}^{AC} + S_{s,a,m,t}^{HS}, \tag{12}$$

$$H_{s,a,m,t}^{HC} + H_{s,a,m,t}^{EH} = H_{s,a,m,t}^{LD}, \tag{13}$$

$$C_{s,a,m,t}^{AC} + C_{s,a,m,t}^{EC} = C_{s,a,m,t}^{LD} + C_{s,a,m,t}^{CS}, \tag{14}$$

$$G_{s,a,m,t}^{SYS} = G_{s,a,m,t}^{LD} + G_{s,a,m,t}^{CHP} + G_{s,a,m,t}^{GB}, \tag{15}$$

where $P_{s,a,m,t}^{PV}$, $P_{s,a,m,t}^{CHP}$, $P_{s,a,m,t}^{LD}$, $P_{s,a,m,t}^{EB}$, $P_{s,a,m,t}^{EC}$, and $P_{s,a,m,t}^{EH}$ are the output power of the PV, CHP, the electric load power, and the electric power consumed by the EB, EC, and EH, respectively; $S_{s,a,m,t}^{CHP}$, $S_{s,a,m,t}^{EB}$, $S_{s,a,m,t}^{GB}$, $S_{s,a,m,t}^{LD}$, $S_{s,a,m,t}^{HC}$, and $S_{s,a,m,t}^{AC}$ are the output steam power of CHP, EB, GB, and the steam load power, the steam power consumed by the HC, and AC, respectively; $H_{s,a,m,t}^{HC}$, $H_{s,a,m,t}^{EH}$, and $H_{s,a,m,t}^{LD}$ are the output heat power of the HC, EH, and the heat load power, respectively; $C_{s,a,m,t}^{AC}$, $C_{s,a,m,t}^{EC}$, and $C_{s,a,m,t}^{LD}$ are the output cold power of the AC, EC, and the cold load power, respectively; $G_{s,a,m,t}^{LD}$, $G_{s,a,m,t}^{CHP}$, and $G_{s,a,m,t}^{GB}$ are

the natural gas load power, the natural gas power consumed by CHP, and GB, respectively. Considering that the accuracy requirements of the planning are not as high as the actual running, in order to improve the efficiency of the model solving, the transmission loss of power grid, gas network and steam network were neglected.

(3) Supply part constraints:

$$P_{\min}^{SYS} \le P_{s,a,m,t}^{SYS} \le P_{\max}^{SYS} + P_0 \cdot I_a^{des}, \tag{16}$$

$$G_{\min}^{SYS} \le G_{s,a,m,t}^{SYS} \le G_{\max}^{SYS}, \tag{17}$$

$$S_{\min}^{SYS} \le S_{s,a,m,t}^{SYS} \le S_{\max}^{SYS}, \tag{18}$$

where $P_{\max}^{SYS}$, $G_{\max}^{SYS}$, and $S_{\max}^{SYS}$ are the upper limit of the interaction power between the system and the external electricity, gas, and steam networks, respectively; P_0 is the capacity for power transmission expansion; $P_{\min}^{SYS}$, $G_{\min}^{SYS}$, and $S_{\min}^{SYS}$ are the lower limit of the interaction power between the system and the external electricity, gas, and steam networks, respectively;

(4) Conversion part constraints

In order to simplify the analysis, the operating efficiency of each energy conversion device is constant, and the variable operating characteristics are neglected. The constraints of GB, EB, AC, EC, HP, and HC are uniformly stated as:

$$Q_{s,a,m,t}^{k1,out} = \eta^{k1} \cdot Q_{s,a,m,t}^{k1,in}, \tag{19}$$

$$\varepsilon_{\min}^{k1} \cdot W_{s,a}^{k1} \le Q_{s,a,m,t}^{k1,in} \le W_{s,a}^{k1}, \tag{20}$$

where $Q_{s,a,m,t}^{k,in}$ and $Q_{s,a,m,t}^{k1,out}$ are the input and output power, respectively, of the above device, k_1, at typical day m, time t in stage a under scenario s; $W_{s,a}^{k1}$ is the total configuration capacity of the device in stage a under scenario s; η^{k1} is the operating efficiency of the device; and $\varepsilon_{\min}^{k1}$ is the lowest power factor of the device.

The CHP is coupled to both electricity and steam with the following constraints:

$$P_{s,a,m,t}^{CHP} = G_{s,a,m,t}^{CHP} \cdot \eta_{ge}^{CHP}, \tag{21}$$

$$S_{s,a,m,t}^{CHP} = G_{s,a,m,t}^{CHP} \cdot \eta_{gh}^{CHP}, \tag{22}$$

$$\varepsilon_{\min}^{CHP} \cdot W_{s,a}^{CHP} \le P_{s,a,m,t}^{CHP} \le W_{s,a}^{CHP}. \tag{23}$$

(5) Storage part constraints

The three types of energy storage device, including BAT, HS, and CS, have similar operating characteristics:

$$W_{s,a,m,t}^{k_2} = W_{s,a,m,t-1}^{k_2}\left(1 - \mu_{loss}^{k_2}\right) + \left(\eta_{ch}^{k_2} \cdot \max\left(P_{s,a,m,t}^{k_2}, 0\right) + \frac{\min\left(P_{s,a,m,t}^{k_2}, 0\right)}{\eta_{dis}^{k_2}}\right) \cdot \Delta t, \tag{24}$$

$$\varphi_{\min}^{k_2} \cdot W_{s,a}^{k_2} \le W_{s,m,a,t}^{k_2} \le \varphi_{\max}^{k_2} \cdot W_{s,a}^{k_2}, \tag{25}$$

$$-P_{\max}^{k_2} \le P_{s,m,a,t}^{k_2} \le P_{\max}^{k_2}, \tag{26}$$

$$W_{s,a,m,24}^{k_2} = W_{s,a,m,0}^{k_2}, \tag{27}$$

where k_2 is the type of energy storage device in the IPIES; $W_{s,a,m,t}^{k_2}$ is the stored energy of energy storage device k_2 at typical day m, time t in stage a under scenario s; $\mu_{loss}^{k_2}$ is the self-consumption rate of the energy storage device, k_2; $\eta_{ch}^{k_2}$ and $\eta_{dis}^{k_2}$ are the charging efficiency and discharging efficiency of the

energy storage device k_2, respectively; Δt is the unit scheduling time; $\varphi_{\max}^{k_2}$ and $\varphi_{\min}^{k_2}$ are the upper and lower limit coefficients, respectively, of the energy storage device, k_2,energy stored; $P_{\max}^{k_2}$ is the upper limit of the switching power of energy storage device k_2, related to the converter device. In order to achieve continuous scheduling, constrained energy storage stores the same energy at the beginning and end of the day.

2.3. Regret Aversion Layer

2.3.1. Optimal Alternative Schemes

The regret aversion layer can be divided into two parts: Optimal alternative schemes calculation and regret aversion optimization. The optimal alternative schemes part uses the expansion planning layer model, with the comprehensive cost in Equation (2) being lowest as the objective function, and Equations (3)–(27) as the constraint. The comprehensive costs C_s^{COM} under all natural gas price fluctuation scenarios, $s \in S$, are calculated as well as the corresponding optimal alternative schemes, ω_s, and operational plans, $\tau_s^{\omega_s}$. The calculated C_s^{COM} will then be substituted as a known parameter into the evasive optimization part.

2.3.2. Regret Aversion Optimization

In response to the regret resulting from the fact that decision-makers did not choose a better expansion planning scheme, the additional aggregate comprehensive cost was chosen as the regret value: In view of the regret aversion optimization part, the decision-maker selects the additional comprehensive cost as the regret value:

$$C_s^{REG}(\omega, \tau_s^{\omega}) = C_s^{COM}(\omega, \tau_s^{\omega}) - C_s^{COM}(\omega_s, \tau_s^{\omega_s}), \tag{28}$$

where $C_s^{REG}(\omega, \tau_s^{\omega})$ is the regret value of the expansion planning scheme, ω,under scenario s; τ_s^{ω} is the typical daily operation plans based on scheme ω under scenario s; $C_s^{COM}(\omega, \tau_s^{\omega})$ is the comprehensive cost of the scheme, ω, and operation plans, τ_s^{ω}, which is obtained during the expansion planning layer; and $C_s^{COM}(\omega_s, \tau_s^{\omega_s})$ is the lowest comprehensive cost under scenario s, which is obtained in the optimal alternative schemes part.

We used the minimum–maximum regret value under all natural gas price fluctuation scenarios considering the distribution of the scenarios as an objective to control the regret risk of the IPIES expansion planning scheme, ω, under different natural gas price fluctuation scenarios:

$$\min_{s \in \mathbf{S}} \max \pi^s \cdot C_s^{REG}(\omega, \tau_s^{\omega}). \tag{29}$$

Further, the minimum–maximum regret aversion objective can be considered together with the minimum objective of the average comprehensive cost. As the optimal comprehensive cost in each scenario is known, the average regret value is equivalent to the objective of the minimum expected comprehensive cost. The objective function of the expansion planning method of the IPIES is finally constructed as:

$$\min \quad C^{CRE}(\omega, \tau_s^{\omega}) \ , \tag{30}$$

$$C^{CRE}(\omega, \tau_s^{\omega}) = \alpha \cdot \max_{s \in \mathbf{S}} \pi^s \cdot C_s^{REG}(\omega, \tau_s^{\omega}) + \beta \cdot \sum_{s \in \mathbf{S}} \pi^s \cdot C_s^{REG}(\omega, \tau_s^{\omega}), \tag{31}$$

where $C^{CRE}(\omega, \tau_s^{\omega})$ is the comprehensive regret value; α and β is the weight coefficient of the minimum–maximum regret aversion objective and the minimum average regret objective, and $\alpha + \beta = 1$. When α is taken as 1, Equation (30) is equivalent to Equation (29).

The proposed method of regret aversion optimization uses Equation (30) as the objective function and Equations (2)–(28) as the constraint condition to optimize the multi-stage expansion

planning scheme of IPIES and the typical daily operation plans under multiple natural gas price fluctuation scenarios.

2.4. Method Solution

The expansion planning method of IPIES considering regret aversion proposed in this paper is a mixed integer nonlinear programming model. Considering the variables and constraints of the model, mathematical modeling was performed on the MATLAB platform through the YALMIP toolbox, and the commercial optimization solver GUROBI was used to solve the model. The model solution environment for this article was: Intel Core 2 Duo P8600 CPU; 6 GB memory; software version: MATLAB R2014A; YALMIP R20180612; GUROBI 8.1.

3. Case Study

3.1. Basic Data

Taking an economic and technological development zone that includes more than 3000 industrial enterprises in Zhejiang province in southeastern China as the case study, the planned total operating life of the IPIES is 15 years, and every 5 years is an expansion planning stage. The typical days of the industrial park can be divided into summer, winter, and ordinary days. The number of days in the whole year is 100, 60, and 200 days. There is cold load demand in summer and heat load demand in winter. There is no gas load demand. The garment industry and beverage processing industry in the park have steam demand. The typical daily energy load curve and PV output curve of the park are shown in Figures 4 and 5, respectively. In addition, consider an extreme scenario where PV units do not output power in a typical summer day. According to the high growth forecast of the park, the sustained load of the three extended planning stages is 150%, 200%, and 250% of the current load. The current electric, steam, and natural gas prices in the park are shown in Table 1. Considering the continuous advancement of China oil and gas marketization reform, referring to Figure 3, it is assumed that the future natural gas price may fluctuate by up to 50% compared with the current price, which has uniform distribution characteristics, corresponding to nine natural gas price fluctuation scenarios with the same probability: (S1) 100%-150%-150%; (S2) 100%-150%-100%; (S3) 100%-150%-50%; (S4) 100%-100%-150%; (S5) 100%-100%-100%; (S6) 100%-100%-50%; (S7) 100%-50%-150%; (S8) 100%-50%-100%; and (S9) 100%-50%-50%.

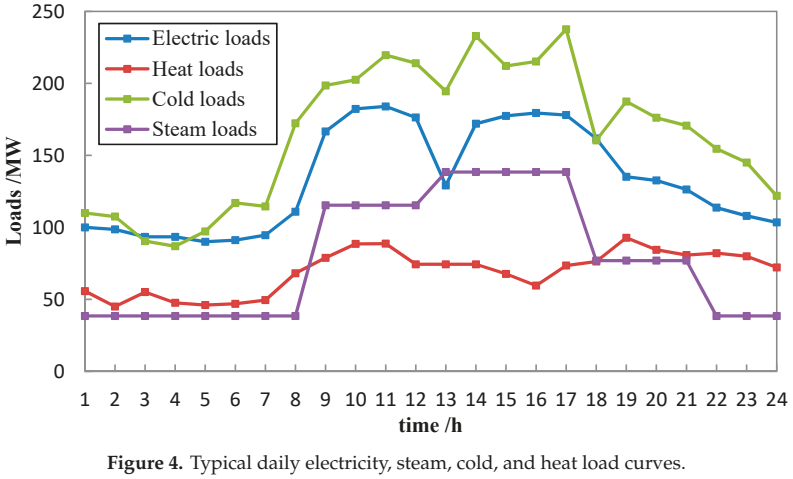


Figure 4. Typical daily electricity, steam, cold, and heat load curves.

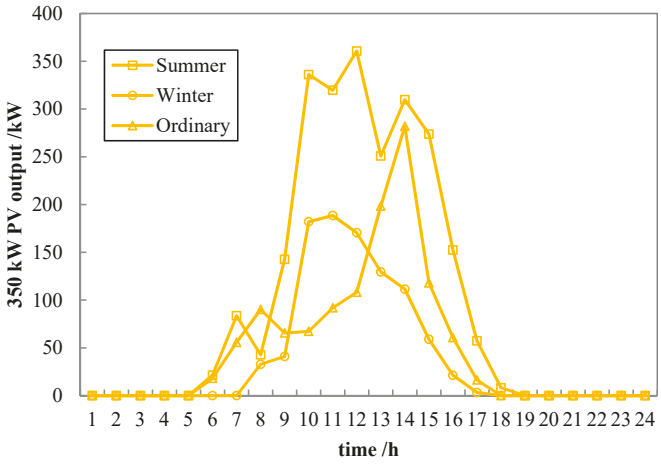


Figure 5. Typical daily 350-kW photovoltaic output curve.

Table 1. Energy peak, normal, and valley prices.

Energy Style	Unit	Peak Time	Normal Time	Vally Time
		19:00–21:00	8:00–11:00 13:00–19:00	00:00–8:00 11:00–13:00 22:00–00:00
electric	yuan/kWh	1.3097	1.0047	0.4817
gas	yuan/kWh	0.353	0.353	0.353
steam	yuan/kWh	0.312	0.312	0.312

The approximate IPIES device operating parameters, unit construction costs, and operation and maintenance costs are shown in Table 2 [44–47]. The relevant parameters of the energy storage device are shown in Table 3 [33,44,47]. The upper limits of the interaction power between the park and the grid and steam network are related to the network capacity, and they are 280 and 192.25 MW (250 T/h), respectively. The system does not output energy to the external network; that is, the lower limit of interaction between each network is 0. The minimum power factor of each device is 0. The power system in the park can be expanded to 35 MW, and the expansion cost is 5 million yuan. The environmental protection costs of electricity, natural gas, and steam using kWh energy per unit in the park are 0.07, 0.01, and 0.04 yuan/kWh, respectively [44,45]. The annual discount rate of the park is 5%, the low heat value of steam is 769 kWh/T, and the low heat value of natural gas is 9.9 kWh/m^3.

Table 2. Device parameters of IPIES.

Device Style	Operating Efficiency	Unit Construction Cost (Yuan/kW)	Unit Maintenance Cost (Yuan/kWh)	Price Correction Coefficient	Critical Price Reduction Factor	Serve Years (A)
PV	-	9000	0.04	−5%	−50%	>15
CHP	Gas-electric: 0.3 Gas-steam: 0.45	6500	0.05	−5%	−50%	>15
GB	0.85	700	0.02	−3%	−20%	>15
EB	0.95	900	0.01	−3%	−20%	>15
HP	3	1000	0.01	−3%	−20%	>15
EC	4.5	1000	0.01	−3%	−20%	>15
HC	0.89	150	0.01	−2%	−10%	>15
AC	1.2	1100	0.01	−3%	−20%	>15
BAT	-	-	0.01	−10%	−80%	10
HS	-	-	0.02	−5%	−50%	>15
CS	-	-	0.02	−5%	−50%	>15

Table 3. Parameters of energy storage.

Energy Storage Style	Unit Construction Cost/yuan/kWh	Self-Consumption Rate	Charging/Discharging Efficiency	Upper/Lower Limit Coefficients	Upper Limit of the Switching Power/MW
BAT	1000	0.001	0.95/0.9	0.95/0.2	10
HS	100	0.01	0.9/0.9	1/0	10
CS	150	0.01	0.85/0.85	1/0	10

3.2. Results and Analysis

The optimized IPIES expansion planning scheme, which takes α as 0.5, is shown in Table 4. The regret value and cost of the scheme under different natural gas price fluctuation scenarios are shown in Tables 5 and 6, respectively.

Table 4. The IPIES expansion planning scheme considering regret aversion.

Capacity	PV	CHP	GB	EB	HP	EC	HC	AC	BAT	HS	CS	I
Capacity in stage 1/MW	426.15	129.43	0	39.96	77.63	126.47	68.79	37.93	44.11	8.50	426.15	0
Capacity in stage 2/MW	536.87	205.09	6.13	59.00	103.51	184.43	143.84	104.32	61.14	11.88	536.87	1
Capacity in stage 3/MW	536.87	224.06	6.13	61.97	122.01	191.41	143.84	104.32	61.14	11.88	536.87	1

Table 5. Regrets under multiple natural gas price fluctuation scenarios.

Values	S1	S2	S3	S4	S5	S6	S7	S8	S9
Regret value/10^6	252.19	161.07	185.70	65.47	13.97	80.32	69.49	50.09	347.41
Comprehensive regret value/10^6					87.40				
Maximum regret value with distribution/10^6					38.60				
Average regret value/10^6					136.19				

Table 6. Costs under multiple natural gas price fluctuation scenarios.

Case1	S1	S2	S3	S4	S5	S6	S7	S8	S9
construction cost/10^9 yuan					5.81				
operation cost/10^9 yuan	19.49	18.49	16.81	18.64	17.65	15.97	16.97	15.98	14.30
maintenance cost/10^9 yuan	0.57	0.62	0.70	0.64	0.70	0.77	0.72	0.78	0.85
environmental protection cost/10^9 yuan	1.54	1.44	1.32	1.40	1.30	1.18	1.28	1.18	1.06
comprehensive cost/10^9 yuan	27.41	26.36	24.64	26.50	25.46	23.74	24.80	23.75	22.03
Average comprehensive cost/10^9 yuan					24.97				

As seen from the expansion planning scheme in Table 4, with an increase in the load level of the park, the capacity of most energy conversion devices in the system expands. Due to the dynamic changes in the price of the system device, and the unpredictable fluctuations in the system's natural gas price from the second phase, the device capacity increase between stages 1 and 2 is greater than that between stages 2 and 3.

The PV in the system uses renewable energy and has excellent economic benefits with the most installed capacity. The CHP is the main natural gas drive device in the IPIES. At the current natural gas price, the CHP has a certain economic advantage over the grid peak hour electricity price and normal electricity price. However, in the case of natural gas price fluctuations, this situation will change. Figure 6 shows the system energy purchases of stage 2 under the scenarios of different natural gas price fluctuations. It can be seen that, when the price of natural gas rises by 50%, the purchase of natural gas in the system decreases, and the purchase of electric and steam increases. When the price of natural gas drops by 50%, the purchase of natural gas in the system increases, and the purchase of electric and steam declined, with the purchase of steam almost falling to zero.

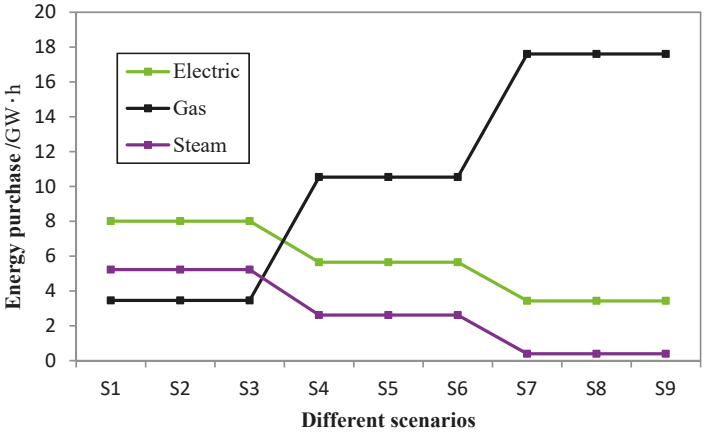


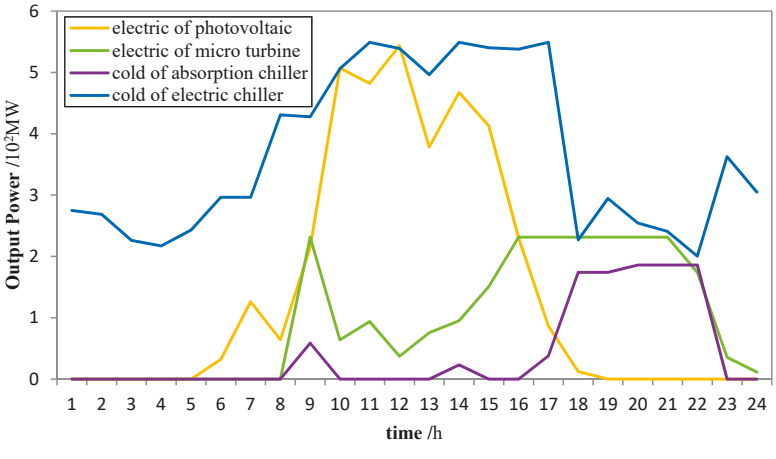
Figure 6. Stage 2 energy purchase situation under different scenarios.

As the system is connected to an external steam network, the role of the GB in supplementing the electro-thermal ratio is replaced to some extent, with fewer configurations in the system. The EB works only at the peak of the PV output, making full use of the electric energy generated by the PV and making up for the lack of steam energy caused by the reduction of the power of the CHP. Heat storage stores energy when the load demand is low and discharges it when the load demand is high. The BAT can store energy in the electricity price valley, and the energy can be released at the peak and normal time to achieve the peaking and filling of the load and the economic improvement of the system.

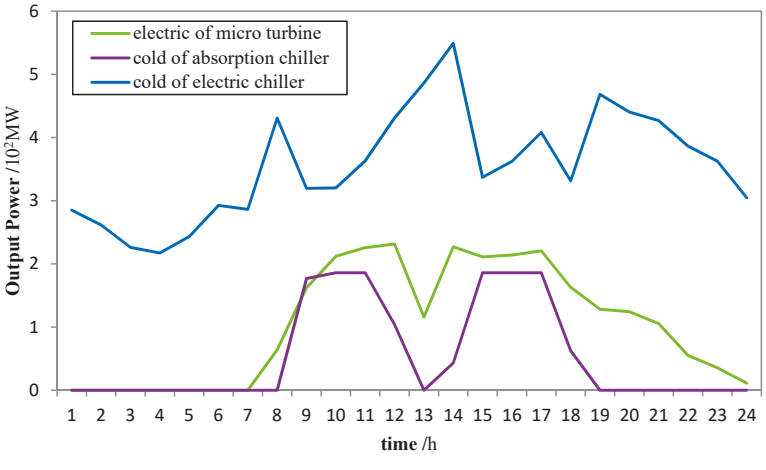
The absorption capacity of the AC and HC in the IPIES is matched with the CHP. It can be seen from Figure 7a that, in the normal output power of the PV, the CHP reduces the output power to make full use of the renewable energy, and the cooling power in the system is mainly provided by the EC. As the PV output power decreases after 14:00, the output power of the CHP increases, and the output cooling power of the AC increases. When entering the electricity price valley at 22:00, the system energy supply turns to the grid and the output power of the EC rises again. In the case of unexpected failure of the photovoltaic unit, the daily load needs to be supplemented by the CHP. At the same time, the output of the EC will rise when the electricity price valley is between 11:00 and 13:00, and the output of the CHP and the AC will decrease.

Comparing Figure 7a,b, it can be seen that in order to ensure the user's energy demand, the system often needs to configure other backup devices to prevent the renewable sources from fluctuating or even zero output, causing the load shedding. Thus, although photovoltaic units have high economic benefits, there are certain restrictions on the permeability of renewable energy in the system. It is necessary to consider the extreme output scenarios of some renewable energy units in the planning process to optimize the capacity of the units and other related units to ensure the safe and economic operation of the system.

Tables 5 and 6 show the regret value and cost of the plan in different natural gas price scenarios, respectively. The planning scheme has a large regret value under scenario 1 and scenario 9, and the regret value of scenario 5 is the smallest. Comparing Table 4 and optimal alternative schemes' typical device planning in stage 3 under scenarios 1, 5, and 9 in Table 7, it can be seen that the high or low natural gas price in scenario 1 and scenario 9 leads to the significant difference in the configuration capacity of the equipment between the schemes, and further causes the actual planning scheme not to match the optimal alternative, resulting in an increase in regret value.



(**a**)



(**b**)

Figure 7. Summer typical day partial devices output power in stage 3 under scenario 5: (**a**) Typical device output power; (**b**) Typical device output power when the PV output is 0.

Table 7. Optimal alternative schemes' typical device planning in stage 3 under scenarios 1, 5, and 9.

Scenario	PV/MW	CHP/MW	EB/MW	HC/MW	AC/MW
1	692.21	205.80	253.37	93.29	183.70
5	536.87	220.51	93.31	181.54	128.52
9	355.05	296.35	33.86	242.28	246.23

The cost difference between the different scenarios of the planning scheme is mainly due to the difference in operating costs. The overall energy consumption of scenario 1 is high, and the running cost is high. The overall energy consumption of scenario 9 is low, and the running cost is low. Because the energy demand of phase 3 is high, the high energy cost of phase 3 has a greater impact on the overall operating cost than phase 2, and the cost of scenario 7 is higher than that of scenario 3.

4. Discussion

To verify the effectiveness of the method, we considered three planning methods for comparison:

Case1: Expansion planning method that considers regret aversion proposed in this paper;
Case2: Expansion planning method based on the lowest expected cost; and
Case3: Expansion planning method that does not consider gas price fluctuations.

The regret value of the schemes obtained by different planning methods is shown in Figure 8. The typical device expansion planning of each scheme is shown in Table 8. The regret values under multiple scenarios with different planning methods are shown in Table 9.

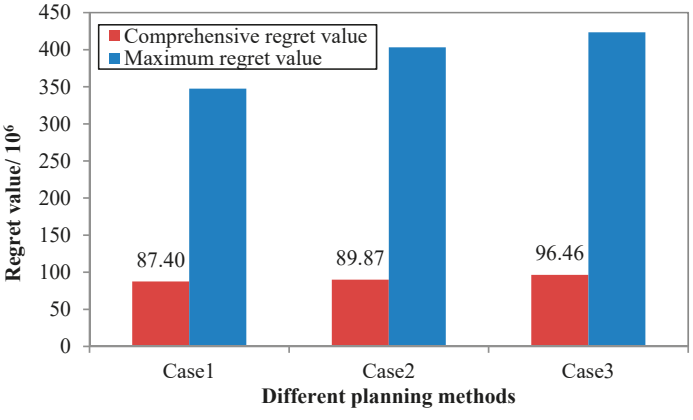


Figure 8. Regret value under different planning methods.

Table 8. Typical device expansion planning under different planning methods.

Cases	Stage	PV/MW	CHP/MW	GB/MW	EB/MW	HC/MW	AC/MW
case1	1	426.15	129.43	0	39.96	126.47	68.79
	2	536.87	205.09	6.13	59.00	184.43	143.84
	3	536.87	224.06	6.13	61.97	191.41	143.84
case2	1	426.15	129.43	0	91.13	126.47	68.79
	2	553.52	194.88	6.21	110.46	184.43	128.52
	3	553.52	220.51	6.21	110.46	184.43	128.52
case3	1	426.15	129.43	0	91.13	126.47	68.79
	2	536.87	184.56	0	93.31	181.54	113.04
	3	536.87	220.51	0	93.31	181.54	128.52

Table 9. Regret value under multiple scenarios with different planning methods.

Values	Cases	S1	S2	S3	S4	S5	S6	S7	S8	S9
Regret value/10^6	Case 1	252.19	161.07	185.70	65.47	13.97	80.32	69.49	50.09	347.41
	Case 2	200.11	130.54	181.91	36.53	6.60	99.68	98.50	57.54	403.15
Comprehensive regret value/10^6	Case 1					87.40				
	Case 2					89.87				
Maximum regret value with distribution/10^6	Case 1					38.60				
	Case 2					44.80				
Average regret value/10^6	Case 1					136.19				
	Case 2					134.95				

More PV units are deployed in the expansion plan with the lowest expected cost, showing the economic benefits of renewable energy in the system. However, in case 2 and case 3, the plan of

natural gas-related devices, such as combined heat and power (CHP), is insufficient, resulting in a large increase in regret when the price of natural gas is low. Compared to the extended planning method without consideration of the fluctuation of gas prices, the proposed method reduces the maximum regret value by 17.8% and reduces the comprehensive regret value by 9.4%.

Compared with the extended planning method based on the lowest expected cost, the method proposed in this paper has a lower regret value. The maximum regret value in case 1 is effectively constrained by the objective function. By introducing more natural gas equipment such as CHP, plan scheme in case 1 has better performance in scenario 6 to 9, especially in the worst scenario 9, but worse performance in scenario 1 to 5. The proposed method reduces the maximum regret value by 13.8% and reduces the comprehensive regret value by 2.7%. Although the average regret value increases by 0.92%, the reduction in the comprehensive regret value indicates that the benefit of controlling the maximum regret value exceeds the control of the average regret value under the decision-maker's risk control requirement.

Further considering the influence of the minimum maximum regret aversion weight coefficient, which represents the risk control requirement of the decision maker, the reduction of comprehensive regret value between case 1 and case 2 can be calculated as:

$$\frac{C^{CRE}\left(\omega_2, \tau_s^{\omega_2}\right) - C^{CRE}\left(\omega_1, \tau_s^{\omega_1}\right)}{C^{CRE}\left(\omega_2, \tau_s^{\omega_2}\right)}, \tag{32}$$

where ω_1, ω_2 is the plan scheme in case 1 and case 2, respectively

The comprehensive regret reduction between case 1 and case 2 under different minimum–maximum regret aversion objective weights, α, are shown in Figure 9. It can be seen from Figure 9 that when the range of is changed from 0.1 to 0.9, the comprehensive regret reduction rises, which indicates that with the increase of the decision-makers' requirement for maximum regret risk control, the planning method proposed is better than traditional method based on the lowest expected cost, making the plan more adaptive when faced with uncertain natural gas prices. If the decision makers have low demand for risk control, the planning method proposed is similar to the traditional method but still provides a little reduction in the comprehensive regret value. It shows that in the industrial park integrated energy system expansion plan, due consideration is given to the regret aversion factor, which can effectively control the regret risk of system planning decisions, and make the plan more adaptive when faced with uncertain natural gas prices.

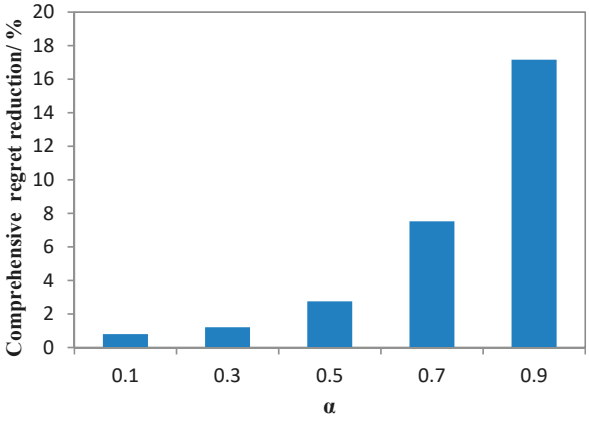


Figure 9. Comprehensive regret reduction between case1 and case 2 under different minimum–maximum regret aversion objective weights, α.

However, the model is relatively simple while the transmission loss of the power grid, gas network, and steam network were entirely neglected in the paper. The theory of how the regret value and expansion plan is affected by load growth expectation was also not put forward in this paper.

5. Conclusions

This paper proposed an expansion planning method for the industrial park integrated energy systems considering regret aversion. Based on the min–max regret aversion and the lowest average regret value, the method optimized the comprehensive cost of an expansion planning scheme in an IPIES under different natural gas price fluctuation scenarios, including costs of construction, operation and maintenance, and environmental protection. The example verifies the rationality and effectiveness of the proposed method. The optimized industrial park integrated energy system expansion plan greatly reduces the degree of decision-making regret and reduces the system cost compared with the traditional expansion plan, which does not consider natural gas price fluctuation. Compared with the expansion plan based on the lowest expected cost, it also effectively controls the system's decision-making regret risk. At the same time, the simulation results show that natural gas price fluctuations have a greater impact on system planning and operation.

With the deepening of the national power system reform, multi-regional integrated energy system collaborative planning and multi-subject integrated energy system planning and operation game theory will be the focus of future research.

Author Contributions: H.X. conceived the main idea and wrote the manuscript with guidance from Q.L., and L.L., who reviewed the work and gave helpful improvement suggestions. P.Z. and X.J. managed the project and provided case data.

Funding: This work was supported by the Science and Technology Project of the State Grid Corporation of China (No. SGZJWZ00FZJS1901007).

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

Acronyms	
IPIES	Industrial park integrated energy system
EH	Energy hub
CHP	Combined heat and power
CCHP	Combined cooling, heat and power
PV	Photovoltaic
EB	Electric boiler
GB	Gas boiler
HP	Heat pump
EC	Electric chiller
HC	Heat exchanger
AC	Absorption chiller
BAT	Battery
HS	Steam heat storage
CS	Cold energy storage
Symbols and matrix	
s	Type of natural gas price fluctuation scenario
m	Type of typical day scenario
k	Type of device in IPIES
k_1	Type of device except for energy storage in IPIES
k_2	Type of energy storage device in IPIES
a	Planning stage
y	year
t	Time

ω	Expansion planning scheme
τ_s^{ω}	Operational plans based on ω under scenario s
ω_{S}	Optimal alternative scheme under scenario s
$\tau_s^{\omega_s}$	Operational plans based on ω_{S} under scenario s
$\mathbf{S}$	Matrix of the natural gas price fluctuation scenario
$\mathbf{c_a^{I,k}}$	unit construction cost matrix of device at stage a
$\mathbf{W_{s,a}^{k}}$	Capacity matrix of device at stage a under scenario s
$\mathbf{c^{M,k1}}$	Unit maintenance cost matrix of the device except for energy storage in IPIES
$\mathbf{Q_{s,a,m,t}^{k1}}$	Operating power matrix of device except for energy storage in IPIES
Variables	
Y	Operating period of the IPIES
T	The years in a planning stage
M	Total number of typical day scenario
D_m	Days of typical day scenario a whole year
g_k	Price correction coefficient of the device k
g_k^c	Critical price reduction factor of the device k
$c_y^{I,k}$	Construction price of the device k in year y
η	Amplitude of the natural gas price fluctuation
π^s	Scenario probability of the natural gas price fluctuation
C_s^{COM}	Comprehensive cost of IPIES under natural gas price fluctuation s
C_s^I	Construction cost
C_s^O	Operation cost
C_s^M	Maintenance cost
C_s^{ENV}	Environmental protection cost
δ_y	Discount rate of year y
λ	Annual discount rate
I_a^{des}	0–1 mark for power transmission expansion status in stage a
c_t^E	Electricity price at time t
c_a^G	Price for unit kW·h energy natural gas in stage a under scenario s
c^S	Price for unit kW·h energy steam
$P_{s,a,m,t}^{SYS}$	Electric power interact with the power grid at typical day m, time t in stage a under scenario s
$G_{s,a,m,t}^{SYS}$	Gas power interact with the gas network at typical day m , time t in stage a under scenario s
$S_{s,a,m,t}^{SYS}$	Steam power interact with the gas network at typical day m , time t in stage a under scenario s
c^{M,k_2}	Unit maintenance cost of energy storage device k_2
$P_{s,a,m,t}^{BAT}$	Power exchange of BAT at typical day m , time t in stage a under scenario s
$S_{s,a,m,t}^{HS}$	Power exchange of HS at typical day m , time t in stage a under scenario s
$C_{s,a,m,t}^{CS}$	Power exchange of CS at typical day m , time t in stage a under scenario s
γ^E , γ^G , γ^S	Environmental cost of emissions from unit electricity, gas, and steam power, respectively
$W_{s,a}^k$	Capacity of the device k in stage a under scenario s
$W_{s,a}^{k,out}$	Capacity of device k to be decommissioned at stage a under scenario s
n_k	Number of planned stages that device k can serve
$P_{s,a,m,t}^{k_1}$	Electric power output or consumed by device k_1 at typical day m , time t in stage a under scenario s
$S_{s,a,m,t}^{k_1}$	Steam power output or consumed by device k_1 at typical day m , time t in stage a under scenario s
$H_{s,a,m,t}^{k_1}$	Heat power output by device k_1 at typical day m , time t in stage a under scenario s
$C_{s,a,m,t}^{k_1}$	Cold power output by device k_1 at typical day m , time t in stage a under scenario s
$G_{s,a,m,t}^{k_1}$	Gas power consumed by device k_1 at typical day m , time t in stage a under scenario s
$P_{s,a,m,t}^{LD}$	Electric power loads at typical day m , time t in stage a under scenario s
$S_{s,a,m,t}^{LD}$	Steam power loads at typical day m , time t in stage a under scenario s
$H_{s,a,m,t}^{LD}$	Heat power loads at typical day m , time t in stage a under scenario s
$C_{s,a,m,t}^{LD}$	Cold power loads at typical day m , time t in stage a under scenario s
$G_{s,a,m,t}^{LD}$	Gas power loads at typical day m , time t in stage a under scenario s
$P_{\max}^{SYS}$	Upper limit of the interaction power between the IPIES and power grid

$P^{SYS}_{\min}$	Lower limit of the interaction power between the IPIES and power grid
$G^{SYS}_{\max}$	Upper limit of the interaction power between the IPIES and gas network
$G^{SYS}_{\min}$	Lower limit of the interaction power between the IPIES and gas network
$S^{SYS}_{\max}$	Upper limit of the interaction power between the IPIES and steam network
$S^{SYS}_{\min}$	Lower limit of the interaction power between the IPIES and steam network
P_0	Capacity for power transmission expansion
$Q^{k,in}_{s,a,m,t}$	Input power of device k_1 at typical day m , time t in stage a under scenario s
$Q^{k1,out}_{s,a,m,t}$	Output power of device k_1 at typical day m , time t in stage a under scenario s
η^{k1}	Operating efficiency of the device k_1
$\varepsilon^{k1}_{\min}$	Lowest power factor of the device k_1
$W^{k_2}_{s,a,m,t}$	Stored energy of energy storage device k_2 at typical day m , time t in stage a under scenario s
$\mu^{k_2}_{loss}$	Self-consumption rate of energy storage device k_2
$\eta^{k_2}_{ch}$, $\eta^{k_2}_{dis}$	Charging efficiency and discharging efficiency of energy storage device k_2, respectively
Δt	Unit scheduling time
$\varphi^{k_2}_{\max}$, $\varphi^{k_2}_{\min}$	Upper and lower limit coefficients of energy storage device k_2, respectively
$P^{k_2}_{\max}$	Upper limit of the switching power of energy storage device k_2
$C^{REG}_s(\omega, \tau^{\omega}_s)$	Regret value of the based on ω and τ^{ω}_s under scenario s
$C^{COM}_s(\omega, \tau^{\omega}_s)$	Comprehensive cost based on ω and τ^{ω}_s under scenario s
$C^{COM}_s(\omega_s, \tau^{\omega_s}_s)$	Lowest comprehensive cost based on ω_s and $\tau^{\omega_s}_s$ under scenario s
$C^{REG}(\omega, \tau^{\omega}_s)$	Comprehensive regret value of the based on ω and τ^{ω}_s under scenario s
α	Weight coefficient of the minimum–maximum regret aversion objective
β	Weight coefficient of the minimum average regret objective

References

1. Quelhas, A.; Gil, E.; Mccalley, J.D.; Ryan, S.M. A Multiperiod generalized network flow model of the U.S.integrated energy system: Part I—Model description. *IEEE Trans. Power Syst.* **2007**, *22*, 829–836. [CrossRef]
2. Wang, Y.; Huang, Y.; Wang, Y.; Yu, H.; Li, R.; Song, S. Energy Management for Smart Multi-Energy Complementary Micro-Grid in the Presence of Demand Response. *Energies* **2018**, *11*, 974. [CrossRef]
3. Wu, J.; Yan, J.; Jia, H.; Hatziargyriou, N.; Djilali, N.; Sun, H. Integrated Energy Systems. *Appl. Energy* **2016**, *167*, 155–157. [CrossRef]
4. Tang, B.; Gao, G.; Xia, X.; Yang, X.; Bo, T.; Gangfeng, G.; Xiangwu, X.; Xiu, Y. Integrated Energy System Configuration Optimization for Multi-Zone Heat-Supply Network Interaction. *Energies* **2018**, *11*, 3052.
5. Voropai, N.; Stennikov, V.; Senderov, S.; Barakhtenko, E.; Voitov, O.; Ustinov, A. Modeling of Integrated Energy Supply Systems: Main Principles, Model, and Applications. *Energy Eng.* **2017**, *143*, 04017011. [CrossRef]
6. National Energy Administration. *Implementation Opinions on Promoting the Construction of Multi-energy Complementary Integration Optimization Demonstration Project*; National Energy Administration: Beijing, China, 2016.
7. Chen, F.; Liang, H.; Gao, Y.; Yang, Y.; Chen, Y. Research on Double-Layer Optimal Scheduling Model of Integrated Energy Park Based on Non-Cooperative Game. *Energies* **2019**, *12*, 3164. [CrossRef]
8. Geidl, M.; Andersson, G. Optimal Power Flow of Multiple Energy Carriers. *IEEE Trans. Power Syst.* **2007**, *22*, 145–155. [CrossRef]
9. Zhang, D.; Evangelisti, S.; Lettieri, P.; Papageorgiou, L.G. Optimal design of CHP-based microgrids: Multiobjective optimisation and life cycle assessment. *Energy* **2015**, *85*, 181–193. [CrossRef]
10. Zhang, X.; Shahidehpour, M.; Alabdulwahab, A.; Abusorrah, A. Security-constrained co-optimization planning of electricity and natural gas transportation infrastructures. *IEEE Trans. Power Syst.* **2015**, *30*, 2984–2993. [CrossRef]
11. Zhao, F.; Zhang, C.; Sun, B.; Wei, D. Three-stage collaborative global optimization design method of combined cooling heating and power. *Proc. CSEE* **2015**, *35*, 3785–3793. [CrossRef]
12. Zhou, X.; Guo, C.; Wang, Y.; Li, W. Optimal Expansion Co-planning of Reconfigurable Electricity and Natural Gas Distribution Systems Incorporating Energy Hubs. *Energies* **2017**, *10*, 124. [CrossRef]

13. Dong, X.; Quan, C.; Jiang, T. Optimal Planning of Integrated Energy Systems Based on Coupled CCHP. *Energies* **2018**, *11*, 2621. [CrossRef]
14. Unsihuayvila, C.; Marangonlima, J.W.; Souza, A.C.Z.D.; Perezarriaga, I.J.; Balestrassi, P.P. A model to long-term, multiarea, multistage, and integrated expansion planning of electricity and natural gas systems. *IEEE Trans. Power Syst.* **2010**, *25*, 1154–1168. [CrossRef]
15. Zhao, J.H.; Dong, Z.Y.; Lindsay, P.; Wong, K.P. Flexible transmission expansion planning with uncertainties in an electricity market. *IEEE Trans. Power Syst.* **2009**, *24*, 479–488. [CrossRef]
16. Zhao, J.H.; Foster, J.; Dong, Z.Y.; Wong, K.P. Flexible transmission network planning considering distributed generation impacts. *IEEE Trans. Power Syst.* **2011**, *26*, 1434–1443. [CrossRef]
17. Qiu, J.; Dong, Z.Y.; Zhao, J.H.; Xu, Y.; Zheng, Y.; Li, C.; Wong, K.P. Multi-stage Flexible Expansion Co-planning Under Uncertainties in a Combined Electricity and Gas Market. *IEEE Trans. Power Syst.* **2015**, *30*, 2119–2129. [CrossRef]
18. Huang, J.; Xue, Y.; Dong, Z.Y.; Wong, K.P. An efficient probabilistic assessment method for electricity market risk management. *IEEE Trans. Power Syst.* **2012**, *27*, 1485–1493. [CrossRef]
19. Zhao, B.; Conejo, A.J.; Sioshansi, R. Coordinated Expansion Planning of Natural Gas and Electric Power Systems. *IEEE Trans. Power Syst.* **2017**, *33*, 3064–3075. [CrossRef]
20. Pazouki, S.; Haghifam, M.R. Optimal planning and scheduling of energy hub in presence of wind, storage and demand response under uncertainty. *Int. J. Electr. Power* **2016**, *80*, 219–239. [CrossRef]
21. Jabr, R.A. Robust transmission network expansion planning with uncertain renewable generation and loads. *IEEE Trans. Power Syst.* **2013**, *28*, 4558–4567. [CrossRef]
22. Liang, Z.; Chen, H.; Wang, X.; Ibn Idris, I.; Tan, B.; Zhang, C. An Extreme Scenario Method for Robust Transmission Expansion Planning with Wind Power Uncertainty. *Energies* **2018**, *11*, 2116. [CrossRef]
23. Ceseña, E.A.M.; Capuder, T.; Mancarella, P. Flexible Distributed Multienergy Generation System Expansion Planning Under Uncertainty. *IEEE Trans. Smart Grid* **2015**, *7*, 348–357. [CrossRef]
24. Dhami, M.K.; Mandel, D.R.; Souza, K.A. Escape from reality: Prisoners' counterfactual thinking about crime, justice, and punishment. In *Routledge Research International Series in Social Psychology. The Psychology of Counterfactual Thinking*; Mandel, D.R., Hilton, D.J., Catellani, P., Eds.; Routledge: New York, NY, USA, 2005.
25. Dijk, E.V.; Zeelenberg, M. On the psychology of if only: Regret and the comparison between factual and counterfactual outcomes. *Organ. Behav. Hum. Dec.* **2005**, *97*, 152–160. [CrossRef]
26. Savage, L.J. The theory of statistical decision. *Am. Stat. Assoc.* **1951**, *46*, 55–67. [CrossRef]
27. Maghouli, P.; Hosseini, S.H.; Buygi, M.O.; Shahidehpour, M. A scenario-based multi-objective model for multi-stage transmission expansion planning. *IEEE Trans. Power Syst.* **2011**, *26*, 470–478. [CrossRef]
28. Bazgan, C.; Vanderpooten, D. Min-max and min-max regret versions of combinatorial optimization problems: A survey. *Eur. J. Oper. Res.* **2009**, *197*, 427–438.
29. Higashikawa, Y.; Augustine, J.; Cheng, S.; Golin, M.J.; Katoh, N.; Ni, G.; Su, B.; Xu, Y. Minimax regret 1 sink location problem in dynamic path networks. *Theor. Comput. Sci.* **2015**, *85*, 24–36. [CrossRef]
30. Halpern, J.Y.; Leung, S. Minimizing regret in dynamic decision problems. *Theor. Decis.* **2016**, *81*, 123–151. [CrossRef]
31. Jiang, R.; Wang, J.; Zhang, M.; Guan, Y. Two-stage minimax regret robust unit commitment. *IEEE Trans. Power Syst.* **2013**, *28*, 2271–2282. [CrossRef]
32. Chen, B.; Wang, J.; Wang, L.; He, Y.; Wang, Z. Robust optimization for transmission expansion planning: Minimax cost vs. minimax regret. *IEEE Trans. Power Syst.* **2014**, *29*, 3069–3077. [CrossRef]
33. Lyu, Q.; Jiang, H.; Chen, T.; Wang, H.; Lyu, Y.; Li, W. Wind power accommodation by combined heat and power plant with electric boiler and its national economic evaluation. *Autom. Electr. Power Syst.* **2014**, *38*, 6–12. [CrossRef]
34. Zhao, D.; Xia, X.; Tao, R. Optimal Configuration of Electric-Gas-Thermal Multi-Energy Storage System for Regional Integrated Energy System. *Energies* **2019**, *12*, 2586. [CrossRef]
35. Li, Z.; Wang, C.; Liang, J.; Zhao, P.; Zhang, Z. Expansion Planning Method of Integrated Energy System Considering Uncertainty of Wind Power. *Power Syst. Technol.* **2018**, *42*, 3477–3487. [CrossRef]
36. Ding, T.; Li, C.; Hu, Y.; Bie, Z. Multi-Stage Stochastic Programming for Power System Planning Considering Nonanticipative Constraints. *Power Syst. Technol.* **2017**, *41*, 3566–3573. [CrossRef]
37. Jeon, C.; Shin, J. Long-term renewable energy technology valuation using system dynamics and Monte Carlo simulation: Photovoltaic technology case. *Energy* **2014**, *66*, 447–457. [CrossRef]

38. Cao, X.; Wang, J.; Zhang, Z.; Cheng, H. Dynamic Assessment of Stand-Alone Microgrid Planning Using Long-Term Simulation. *Trans. Chin. Electrotech. Soc.* **2016**, *31*, 46–56. [CrossRef]
39. Geng, J.; Ji, Q.; Fan, Y. How regional natural gas markets have reacted to oil price shocks before and since the shale gas revolution: A multi-scale perspective. *J. Nat. Gas Sci. Eng.* **2016**, *36*, 734–746. [CrossRef]
40. Serletis, A.; Xu, L. Volatility and a century of energy markets dynamics. *Energy Econ.* **2016**, *55*, 1–9. [CrossRef]
41. Cansado-Bravo, P.; Rodríguez-Monroy, C. Persistence of Oil Prices in Gas Import Prices and the Resilience of the Oil-Indexation Mechanism. The Case of Spanish Gas Import Prices. *Energies* **2018**, *11*, 3486. [CrossRef]
42. Shanghai Petroleum and Natural Gas Exchange (SHPGX). Available online: https://www.shpgx.com (accessed on 20 October 2019).
43. IGU (International Gas Union). *Wholesale Gas Price Survey*; IGU: Barcelona, Spain, 2019.
44. Wei, F.; Wu, Q.H.; Jing, Z.X.; Chen, J.; Zhou, X.X. Optimal unit sizing for small-scale integrated energy systems using multi-objective interval optimization and evidential reasoning approach. *Energy* **2016**, *111*, 933–946. [CrossRef]
45. Zhou, C.; Zheng, J.; Jing, Z.; Wu, Q.; Zhou, X. Multi-Objective Optimal Design of Integrated Energy System for Park-Level Microgrid. *Power Syst. Technol.* **2018**, *42*, 1687–1697. [CrossRef]
46. Liu, D.; Qin, G.; Li, Q. Research on Optimized Operation of Integrated Energy Microgrid Considering Multi-Type Energy Conversion and Storage. *South. Power Syst. Technol.* **2018**, *12*, 105–115. [CrossRef]
47. Lei, J.; Yu, L.; Guo, X.; Li, P.; Li, C.; Wu, Y. Equipment Selection and Capacity Planning for Combined Power Heat and Gas Integrated Energy System. *Proc. CSU EPSA* **2019**, *31*, 19–24. [CrossRef]

energies

MDPI

Article

Financial Impacts of Net-Metered Distributed PV on a Prototypical Western Utility's Shareholders and Ratepayers

Peter Cappers [1,*], Andrew Satchwell [1], Will Gorman [1] and Javier Reneses [2]

[1] Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA; asatchwell@lbl.gov (A.S.); wgorman@lbl.gov (W.G.)
[2] Institute for Research in Technology, Technical School of Engineering (ICAI), Universidad Pontificia Comillas, 28015 Madrid, Spain; javier.reneses@iit.comillas.edu
* Correspondence: pacappers@lbl.gov; Tel.: +1-315-637-0513

Received: 5 November 2019; Accepted: 13 December 2019; Published: 16 December 2019

Abstract: Distributed solar photovoltaic (DPV) under net-energy metering with volumetric retail electricity pricing has raised concerns among utilities and regulators about adverse financial impacts for shareholders and ratepayers. Using a pro forma financial model, we estimate the financial impacts of different DPV deployment levels on a prototypical Western U.S. investor-owned utility under a varied set of operating conditions that would be expected to affect the value of DPV. Our results show that the financial impacts on shareholders and ratepayers increase as the level of DPV deployment increases, though the magnitude is small even at high DPV penetration levels. Even rather dramatic changes in DPV value result in modest changes to shareholder and ratepayer impacts, but the impacts on the former are greater than the latter (in percentage terms). The range of financial impacts are driven by differences in the amount of incremental capital investment that is deferred, as well as the amount of incremental distribution operating expenses that are incurred. While many of the impacts appear relatively small (on a percentage basis), they demonstrate how the magnitude of impacts depend critically on utility physical, financial, and operating characteristics.

Keywords: distributed solar PV; financial analysis; net-energy metering; investor-owned utility; earnings; return on equity; retail rates; ratepayer bills

1. Introduction

Residential solar power is rapidly expanding in the United States (U.S.). In 2018, there was a 7% increase in residential distributed solar photovoltaic (DPV) deployment [1]. Such large increases in the deployment of DPV in the U.S. over the previous 5–7 years has been attributed to significant declines in equipment costs [2], state and federal tax credits, and electric utility net-energy metering (NEM) compensation programs [3]. NEM is a billing mechanism that credits customers with distributed generation systems for any electricity they export to the grid [4]. Use of NEM in conjunction with volumetric retail electricity pricing (i.e., uniform compensation of generation in excess of consumption, regardless of its characteristics such as time of generation), however, has also raised concerns among utilities and regulators of higher retail electric power rates and shifting of costs from DPV to non-DPV customers [5]. While current amounts of DPV in many jurisdictions are small and thus any retail rate and cost-shifting concerns may be anticipatory in nature, NEM reforms have been proposed and, in certain cases, adopted by state public utility commissions [6]. Importantly, most reforms change the DPV system payback periods and have the potential to reduce distributed solar PV deployment [7]. Barbose et al. [8], for example, modeled effects of a reduction in NEM compensation for grid exports

from retail to wholesale electric power rates. They found that this reduction in NEM compensation would reduce residential 2050 solar PV deployment by approximately 20%.

Electric investor-owned utilities (IOUs), particularly those in the United States, are concerned about the effects of DPV on future earnings opportunities from deferred or avoided capital investments under existing regulatory and business models [9]. IOUs increase their earnings base by investing in capital, which may be growth-related (e.g., new distribution system investments and generating plants to serve increasing load) [10]. Thus, stagnant or declining sales as a result of DPV, as well as energy efficiency [11] and other forms of other forms of distributed energy resources (DERs), may reduce these future earnings opportunities [12]. Future growth in electric vehicle penetration, among other sources of electrification which increase electricity consumption, may counter the prevailing trend of declining sales.

Furthermore, the decrease in DPV and other forms of DER costs (e.g., battery storage) has led to increased financial pressure on the utility from customer self-supply [13]. Many utilities around the world typically allocate a significant portion of their fixed costs to volumetric energy charges. As a result, any reduction in electricity sales without a similar reduction in fixed costs erodes a utility's net revenues. Such impacts are especially disconcerting to utility shareholders due to the reduction in achieved earnings and return on equity (ROE) [14].

At the same time, utility regulators are increasingly concerned about possible increases in retail rates and cost-shifting from customers with DPV (i.e., participants) to non-DPV customers (i.e., non-participants) [15]. In instances where costs increase faster than sales, there is upward pressure on retail rates. In addition, customers who invest in DPV and can significantly reduce or even eliminate the volumetric portion of their bills via NEM may not adequately pay their full share of the utility's fixed costs, which places an increased cost burden on non-participating customers.

Regulators of any utility in such a situation must weigh utility and ratepayer concerns as they consider changes to NEM and retail rate design that directly affects DPV, battery storage, and other forms of DER. Ultimately, they must make a determination that they believe serves the broader public interest based on the information they have available to them.

While these concerns are qualitatively understood, there is a lack of empirical and quantitative analysis to bound the magnitude of these concerns and the efficacy of alternative utility regulatory and business models. Instead, most quantitative analyses evaluating DPV impacts on an electric utility exclusively focus on a simplified cost and/or benefit analysis without considering the financial implications on a utility [16]. Those studies which focus on DPV costs find incremental electric system costs which range from $0–25/MWh [17–24]. On the other hand, those studies which focus on DPV benefits find electric system benefits which range from $0–53/MWh [25–32]. In general, these analyses focus on system costs and benefits without considering the role of a regulated utility and their accompanying business practices.

Another subset of the literature evaluates how NEM and DPV adoption impacts utility ratepayers but notably does not incorporate a fully-integrated pro forma financial model. Poulilkkas [33] studied the effect of NEM on DPV adoption at one representative household while Christoforidis et al. [34] performed a similar analysis across a broader set of 31 customers. Neither study evaluated the impact on electric utility collected revenues or earnings. Eid et al. [35] and Picciariello et al. [36] expanded on this work by evaluating the cross-subsidization due to NEM between PV owners and non-owners. Furthermore, Johnson et al. [37] added an analysis of the DPV impact on utility load shape into the analysis of NEM cross-subsidization between PV owners and non-owners. However, none of the above quantitative studies: (1) calculate utility shareholder impacts, (2) take into account regulatory lag and other artifacts of utility regulation (e.g., test years), or (3) integrate a long-term analysis horizon that incorporates feedback effects between PV hourly loads and utility costs that accumulate over time to impact electricity rates.

Many of the concerns expressed by utilities and regulators, though, depend critically on the specific changes in costs and revenues that are a function of utility cost compositions (e.g., proportion

of fixed versus variable costs), physical system attributes (e.g., hourly loads), and fixed cost recovery mechanisms, among others, which interact with PV adoption and utility regulatory considerations over time. Most of the prior studies reviewed above do not perform a comprehensive financial analysis using these key utility characteristics and do not incorporate a robust review of the costs and benefits of DPV in retail electricity rates.

To fill this research gap, this study quantifies the financial impacts of net-metered DPV on a prototypical Western U.S. IOU and identifies the key sensitivities driving lesser or greater magnitude of impacts. While we integrate the above cost-benefit studies into our financial modeling, we use a financial framework that better assesses the implications for a regulated IOU. Furthermore, we build on prior quantitative analysis of the financial impacts of net-metered PV [38,39] by assessing a wider range of sensitivities specific to the ability of DPV to avoid or defer utility costs (i.e., "DPV value"). Although the costs, revenue, and regulatory accounting assumptions are based on the U.S. context, we believe the results are generalizable and relevant for other utility circumstances around the world.

2. Materials and Methods

We quantify the shareholder and ratepayer impacts for a Western U.S., vertically-integrated IOU (i.e., that owns generation, transmission, and distribution assets) at two DPV deployment levels (i.e., 4%, and 8% of 2027 retail sales) representing the range of forecasted DPV deployment among Western states [40] using a proprietary pro forma financial model.

The FINancial impacts of Distributed Energy Resources (FINDER) pro forma financial model quantifies the utility's annual costs and revenues over a pre-defined analysis period. See Appendix A for a more detailed description of the pro forma financial model used in this analysis. The model performs all cost calculations at the total utility level but has the ability to allocate those costs to different rate classes in order to assess differential revenue impacts. Key outputs include achieved ROE and earnings, average all-in retail rates and customer bills, which can be used by utilities, policymakers, customer groups, and other stakeholders when assessing the impacts and implications of policy proposals and decisions.

Because the model derives customer class level ratepayer metrics, a more comprehensive bill impact analysis across different customer types, which assumes different hourly consumption profiles for customers and different DPV production profiles for participants, was not possible. A number of other studies have sought to quantify participant and/or non-participant bill impacts in substantial detail, but without any associated feedback effects on rates [41–43]. However, our estimates of the percentage change in average all-in retail rates can be used as a proxy for the percentage change in a non-participants' bill assuming there are no associated changes in their electricity consumption.

Results of our analysis using this pro forma financial model are compared against a case without DPV, incremental energy efficiency, or other DERs in order to isolate the DPV impacts. The DPV is installed linearly over ten years to reach its terminal deployment level (as a percentage of retail electricity sales) with impacts measured over 20 years to capture utility system economic end-effects (i.e., cost avoidance or deferral). We limit our analysis to 20 years despite DPV system lifetimes in excess of 20 years due to the effects of discounting costs and revenues, in addition to increasing uncertainty in utility cost and load forecasts. We also assess the sensitivity of impacts to different assumptions about the "DPV value" (i.e., ability of DPV to avoid or defer utility costs). Each of these different cases are discussed in more detail below.

2.1. Prototypical Western Utility Characterization—Base Case

We developed a 20-year cost and load forecast for a prototypical Western utility by starting with the 2014 general rate case filing of Public Service of Colorado. The general rate case filing was the most recent for the utility that included a cost-of-service study and provided reasonable starting year cost levels, starting year class-level retail sales, peak demand, and customer counts, and class-level cost allocation and rate design. Growth in retail sales, peak demand, and customers are based on Public

Service of Colorado's 2016 integrated resource plan (IRP), which was the most recent one available. Growth in utility cost categories, specifically generation capital expenditures (CapEx), distribution CapEx, and operations and maintenance (O & M), are based on historical 5-year average annual growth rates among Western utilities derived from their FERC Form 1 filings. Last, the Western utility's hourly load shape is based on Public Service of Colorado's 2017 hourly load (reported in EIA Form 930) and we used a simple average of load in hours before and after missing values to derive a complete 8760-h load shape. Importantly, while many of the input assumptions were seeded with a single utility's data, the utility in this analysis is not meant to represent Public Service of Colorado specifically.

We refer to four key assumptions in the Western utility characterization when describing financial impacts. First, non-fuel costs (inclusive return of (i.e., depreciation) and on capital investments, fixed O & M, interest expense, and taxes) grow by 3.3% per year over the 20-year analysis period (2018–2037) (see Figure 1). Western utilities have seen median generation, transmission, and distribution CapEx costs increase by 6.4%, 3.6%, and 3.8% per year from 2012 to 2016, respectively (calculated based on utility FERC Form 1 data), and we assume similar, rounded CapEx cost escalations (i.e., 6%, 4%, and 4% annual growth in generation, transmission and distribution CapEx costs, respectively). Utility fuel and purchased power (energy and capacity) costs grow by 2.6% per year over the same 20-year analysis period. Our fuel and purchased power costs for non-renewable generation technologies are based on EIA fuel, heat rate, and variable O & M cost assumptions [44]. Wind and solar PPA costs are based on NREL's Annual Technology Baseline LCOEs in the "Wind TRG 4" and "Solar CF 20%" forecasts [45].

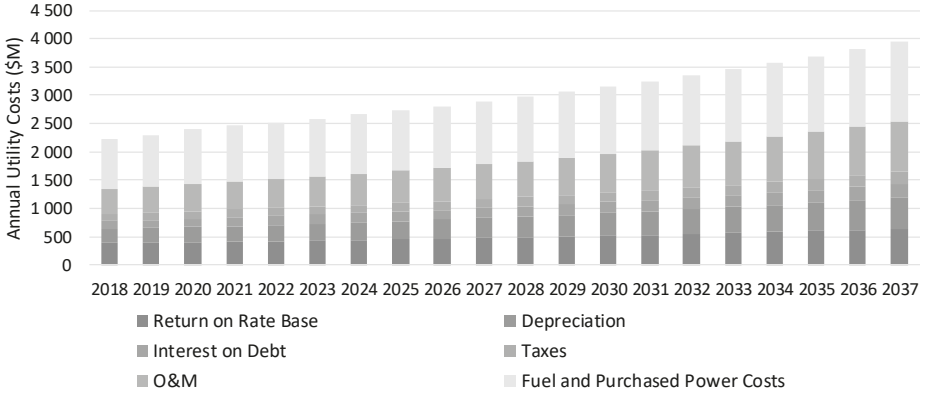


Figure 1. Forecasted Western utility costs (without DPV).

Second, the Western utility's retail sales grow by about 1.0% per year and annual peak demand grows by about 0.9% per year. Our load growth assumptions are based on a specific utility's IRP forecasts in order to match any incremental generation or purchased power, though the retail sales forecasts are higher than historical, median sales growth among Western utilities. From 2012 to 2016, Western utility median sales slightly declined by about 0.3% per year based on EIA-861 data.

Third, we assume no incremental generation additions in the Base case analysis, as Public Service of Colorado is forecasting PPAs to meet incremental load in its 2016 IRP. We make this assumption about no incremental generation additions in order to maintain consistency between our load and cost assumptions. Given that Western utilities averaged flat, or declining, load growth over the last 5-years, we believe our assumption is reasonable. Nevertheless, we consider the case of incremental generation additions in our sensitivity analysis. We also assume retirement of three generating units during the 20-year analysis period to maintain consistency with Public Service of Colorado's IRP loads and resources table. These are all input assumptions used to develop a pro forma revenue requirement and were not derived as part of a system planning module within FINDER. As such, the impacts of DPV

on utility capital costs are based on a coarser set of assumptions than might be possible with planning models representing the utility's generation dispatch and power flows on distribution feeders.

Fourth, we assume a flat retail rate design for all customer classes (as opposed to inclining block or time-of-use rates). Residential customer rates and bills collect 90% of revenues via volumetric energy rates with the remainder of revenues (10%) collected via a monthly, fixed customer charge. Commercial and industrial (C & I) customer rates and bills collect about 40% of revenues via an energy charge, 55% of revenues via a demand charge (based on class monthly non-coincident peak), and the remaining via a fixed, monthly customer charge.

2.2. Alternative Assumptions in Utility Characterization—Sensitivity Cases

We developed a set of cases to better understand the sensitivity of shareholder and ratepayer impacts from DPV to assumptions related to its capacity value and avoided generation, transmission, and distribution costs. These sensitivity cases involved modifying a number of parameters from the Base case (see Figure 2), based on ranges that exist in either the literature or Western utility historical data.

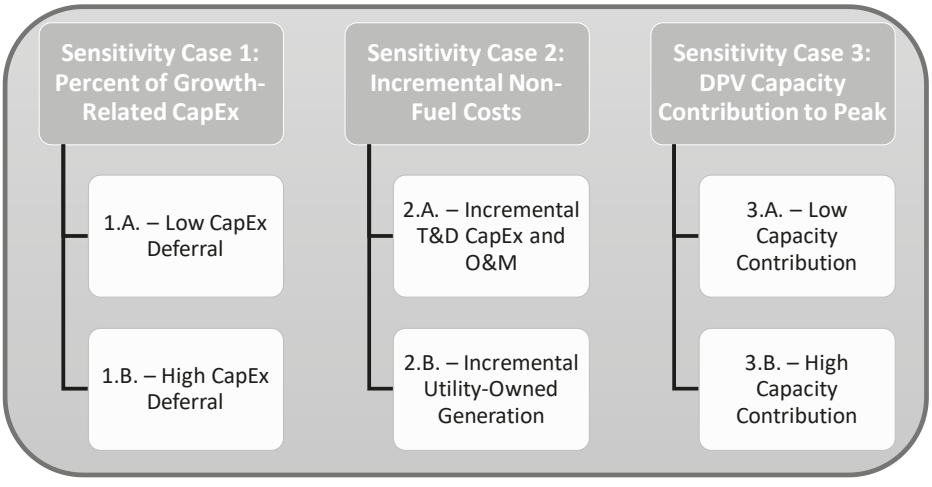


Figure 2. Modeled sensitivity cases among three key assumptions related to DPV value.

In our first sensitivity case, we change the assumed percent of transmission and distribution (T & D) CapEx that are growth-related. As previously discussed, we model two categories of T & D CapEx: load growth-related and non-load growth-related. The Base Case assumes a portion (33%) of investments are growth-related CapEx and the addition of DPV reduces this growth-related CapEx proportional to reductions in annual peak demand. The 33% assumption is consistent with assumptions in prior studies [38,39,46]. This results in corresponding reductions in returns on ratebase, depreciation expenses, interest, and taxes For the sensitivity case, we bound the assumption with a low value of 10% growth-related T & D CapEx (i.e., Sensitivity Case 1.A.) and high value of 90% growth-related T & D CapEx (i.e., Sensitivity Case 1.B.). Appendix B shows the sensitivity case analysis results as change in DPV value, change in achieved earnings, and change in average all-in retail rates.

In our second sensitivity case, we change assumptions related to incremental CapEx. In one case, we increase distribution CapEx and O & M costs in conjunction with DPV to represent the possibility that integration costs for DPV could result in a net increase in distribution system expenditures (i.e., Sensitivity Case 2.A.). For the purposes of our study, system integration costs are borne by the utility (and all ratepayers via retail rates) and are different from interconnection costs that are paid by DPV owners. We assume incremental distribution O&M cost of $15/kW-year installed DPV [47]

and incremental distribution CapEx of $30/kW installed DPV [48]. Alternatively, we add incremental utility generation to meet future capacity needs motivated by the fact that some Western utility DPV value studies assumed deferred generation (i.e., Sensitivity Case 2.B.). As discussed in Section 2.1, the Base case utility characterization assumes the Western utility meets future capacity and energy through PPAs and short-term market purchases (as is consistent with the IRP data we used as the basis for our Western utility characterization). We base this sensitivity case on a simple loads and resources table and add mid-merit and natural gas generating plants in 100MW and 250MW increments to meet forecasted peak demand in the Base case without DPV. Capital and O & M costs of the incremental generation are based on EIA overnight capital cost estimates for generating plants and inflated at 2% per year.

In our third sensitivity case, we change the amount of DPV coincident with the Western utility's annual peak demand (i.e., capacity contribution to peak). The Base case assumes a 22% DPV capacity contribution to peak (discussed in the next section). We bound this assumption with a lower value of 12% (i.e., Sensitivity Case 3.A.) and higher value of 32% (i.e., Sensitivity Case 3.B.). The DPV capacity contribution to peak drives the reduction in annual system peak demand upon which capacity and T & D CapEx costs are based. Thus, an increase in DPV capacity contribution to peak would result in greater avoided capacity-driven costs at the same DPV deployment level.

2.3. DPV Characterization

Our pro forma financial model derives the impacts of DPV through several key static and dynamic interrelationships. DPV impacts utility billing determinants; specifically retail sales and peak demand, which has an effect on utility costs and subsequently retail rates. DPV reduces volumetric sales based on a direct relationship between the assumed annual DPV penetration, expressed as a percentage of annual sales on a customer-class basis, and the utility's class-level sales. The model derives reductions in the utility's peak demand through dynamic relationships of several variables that take into account: (1) the specific timing of DPV relative to the utility's hourly load, and (2) potential differences in alignment between when the DPV causes reductions in the utility's load and the utility system annual peak demand.

The timing of DPV production (savings) and the utility annual peak demand is a key driver in the analysis. DPV reduces energy only in hours when the PV system operates (i.e., during the daylight hours) and may also reduce utility system demand depending on whether the reduction in energy is coincident with the utility's system peak. Thus, the timing of DPV energy and demand savings in relation to customer class and utility peak demands (monthly and annual) drives projections of future costs, retail rates, and revenues collected on a volumetric basis through energy and demand charges.

To calculate the DPV shape, we relied on the National Renewable Energy Laboratory's (NREL) System Advisor Model (SAM) [49]. We simulated five solar profiles with typical meteorological year (TMY) weather data for different locations in Colorado's main metropolitan areas (i.e., Boulder, Aurora, Denver, Colorado Springs, and Pueblo). These simulations relied on PV Watts default assumptions (i.e., azimuth of 180 degrees, DC to AC ratio of 1.2, and tilt of 40 degrees). To estimate a single input profile for our pro forma financial model, we applied a population-weighted average solar output of the five metro area's solar shapes.

We further analyzed DPV's capacity contribution to peak load reduction by simulating DPV profiles using 2017 weather data. While the TMY weather data described above provides an ideal average profile, it does not provide an understanding of DPV's contribution to peak load reduction for our underlying load year of 2017. To determine this value, we simulated the additional solar profiles and sampled the capacity factor of our DPV simulation in the top-20 load hours of 2017 for Public Service of Colorado. The resulting capacity contribution to the top-20 load hours was 22%. DPV capacity factors are often calculated based on probabilistic simulation and modeling methods such as Effective Load Carrying Capability (ELCC), however, carrying out such a simulation is not the focus of this study; alternatively, we quantify DPV capacity contribution as the percentage of

the nameplate capacity that is available during top-20 load hours; this factor provides a simple and intuitive measure of the value provided by DPV in terms of capacity and can be represented in our pro forma financial model. This value became our Base Case capacity contribution to peak load reduction for DPV. We performed this analysis for a number of other Western utilities and found that this capacity value in the top-20 load hours ranged from 7% to 26%, which we use to inform our sensitivity cases discussed above.

The DPV portfolios reduce the Western utility's annual retail sales and peak demand. Retail sales grow by 1.0% per year in the case without PV, but the annual growth rate declines to 0.6%, and 0.2% in the 4% and 8% DPV deployment cases, respectively, from 2018 to 2027 as DPV systems are installed. Because the DPV penetration levels are specified in terms of a percent reduction of retail sales, they each reduce the Western utility's sales on a one-for-one basis. As shown in Figure 3, the reduction in retail sales increases as the DPV deployment level increases.

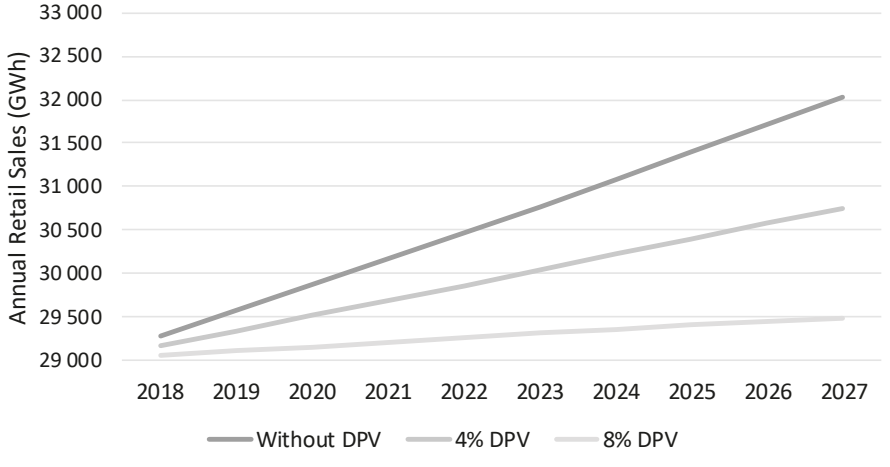


Figure 3. Forecasted Western utility annual retail sales without DPV and at increasing DPV deployment levels (4% and 8% of 2027 retail sales).

The impacts of DPV on the Western utility's annual peak demand depends on the timing and coincidence of DPV relative to the utility's annual peak demand. Figure 4 shows the forecasted annual peak demand for the Western utility from 2018 to 2027 and the coincident peak demand savings attributable to the DPV deployment cases. The prototypical Western utility modeled in this study has peak loads that occur in July generally between 2 p.m. and 6 p.m. prior to the addition of DPV systems. The coincident peak demand impact of DPV in our study is less than the retail sales impacts on a percentage basis (e.g., 0.8% per year reduction in retail sales and 0.6% per year reduction in peak demand in the 8% DPV deployment case) because the timing of maximum PV output does not coincide perfectly with the utility's annual peak demand. This is particularly the case for Public Service of Colorado that serves load near the Rocky Mountains, which results in lower DPV production in afternoon hours relative to other geographic locations due to the effect of mountain shadows [50]. We consider lower and higher contribution of DPV savings to peak in the DPV value sensitivities.

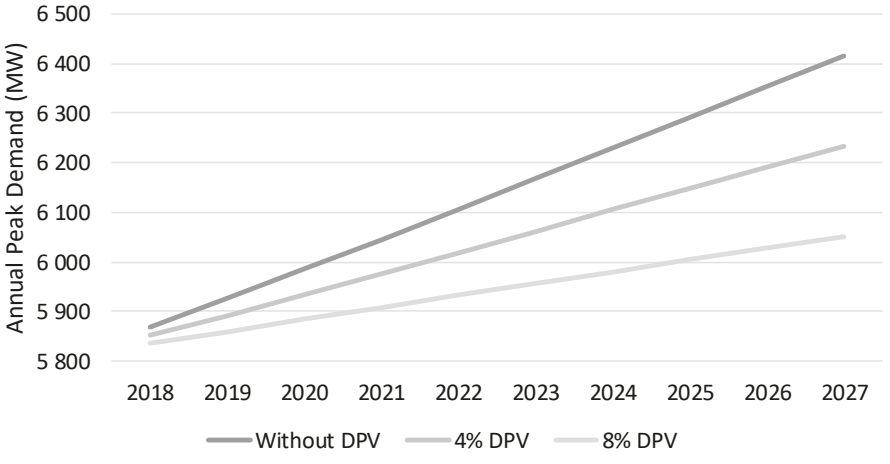


Figure 4. Forecasted Western utility annual peak demand without DPV and at increasing DPV deployment levels (4% and 8% of 2027 retail sales).

3. Results

3.1. Sensitivity of DPV Value to Alternative Utility Assumptions

Figure 5 shows the change in DPV value with Base and alternate assumptions of the proportion of T & D CapEx that is growth-related (i.e., Sensitivity Case 1.A. and 1.B.). A lower proportion of growth-related T & D CapEx results in a lower DPV value, and vice-versa, where the change in value occurs exclusively among non-fuel cost categories. In the 4% DPV deployment case, the DPV value ranges from $51/MWh to $57/MWh and, in the 8% DPV deployment case, the DPV value ranges from $50/MWh to $55/MWh. The modeled DPV value results are not particularly sensitive to this assumption, ranging from −4% to 6% relative to the Base Case assumption.

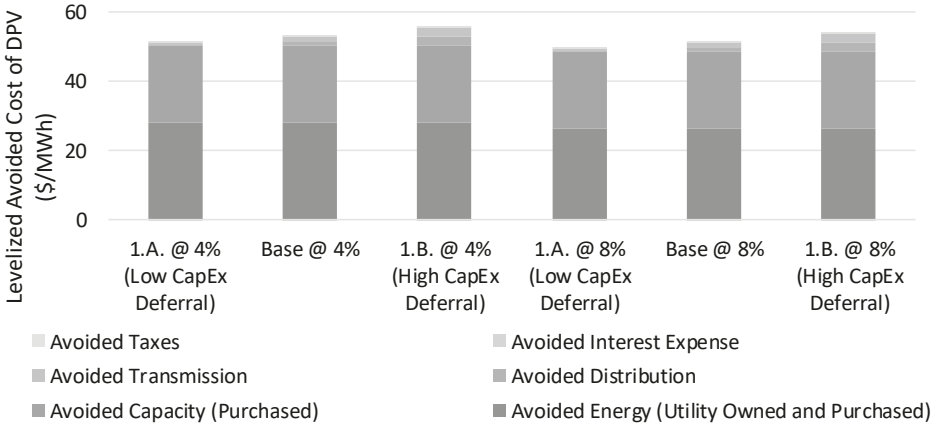


Figure 5. Sensitivity of DPV value to assumed proportion of growth-related CapEx.

Figure 6 shows the change in DPV value assuming incremental distribution and generation CapEx and distribution O & M costs to the Base Case (i.e., Sensitivity Case 2.A. and 2.B.). The incremental distribution CapEx and O & M costs (i.e., Sensitivity Case 2.A.) do not result in additional value, as the costs are added incrementally with the installed DPV and counteract many of the avoided costs.

Thus, the DPV value declines significantly in both the 4% DPV and 8% DPV deployment cases. In fact, DPV value at 8% deployment and assuming incremental distribution CapEx and O & M is roughly half of the Base Case DPV value ($27/MWh compared to $52/MWh). The incremental utility generation assumption (i.e., Sensitivity Case 2.B.) reduces the avoided purchased capacity value, as would be expected due to the incremental generation installed in lieu of the capacity purchases. Also, as to be expected, the proportion of CapEx deferral value increases as the DPV defers or avoids some of the incremental generation. In total, however, the DPV value in the incremental utility generation case is about 10% lower than the Base Case assumption because the cost of utility-owned generation is lower relative to meeting the same capacity needs through PPAs and short-term market purchases. Thus, the incremental utility generation sensitivity case produces a lower total DPV value.

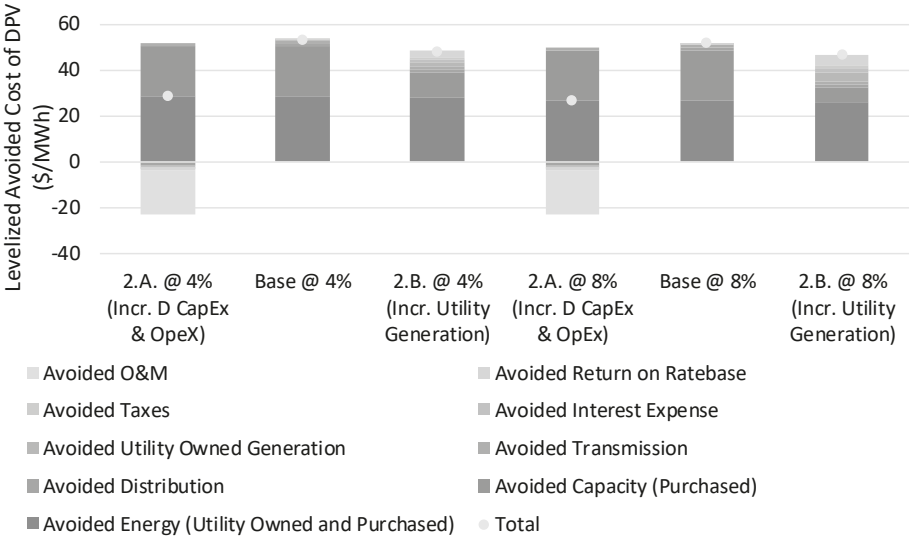


Figure 6. Sensitivity of DPV value to incremental non-fuel costs.

Figure 7 shows the change in DPV value with lower or higher assumed DPV capacity contribution to peak relative to the Base Case assumptions (i.e., Sensitivity Cases 3.A. and 3.B.). As expected, a lower capacity contribution to peak (i.e., Sensitivity Cases 3.A.) results in lower DPV value, and vice-versa, with the largest change in avoided capacity purchases. In the 4% DPV deployment case, the DPV value ranges from $42/MWh to $59/MWh and, in the 8% DPV deployment case, the DPV value ranges from $40/MWh to $54/MWh. The modeled DPV value results are quite sensitive to this assumption, ranging from −27% to 11% relative to the Base Case assumption. The results for all sensitivity cases show that DPV value is sensitive to alternate assumptions, but the degree depends on the specific assumption. For example, the assumed proportion of growth-related CapEx (i.e., Sensitivity Case 1.A. and 1.B.) has a small range whereas the DPV capacity contribution to peak (i.e., Sensitivity Case 3.A. and 3.B.) exhibits a much larger range of results.

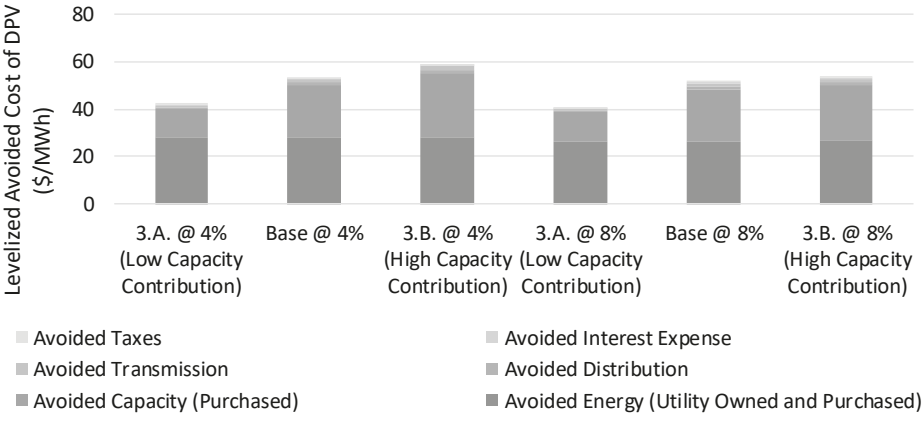


Figure 7. Sensitivity of DPV value to assumed DPV capacity contribution to peak.

Importantly, we did not combine changes in more than one key assumption, which would likely drive greater change in DPV value than is observed in isolated cases (e.g., combine higher DPV capacity contribution to peak with higher CapEx deferral which would likely leader to greater DPV value).

3.2. Sensitivity of Shareholder and Ratepayer Metrics to Alternative Utility Assumptions

As shown in Figures 8 and 9, the impacts of DPV on shareholder earnings and ROE vary under these different assumptions related to the penetration of DPV and the value of DPV to the utility. In the Base Case, the after-tax earnings achieved by the Western utility decline as the DPV deployment level increases (1.5% reduction at 4% DPV and 3.1% reduction at 8% DPV) while its achieved average ROE declines as the DPV deployment level increases (1.6% reduction at 4% DPV and 3.2% reduction at 8% DPV) compared to the case without DPV. Adding incremental distribution CapEx and distribution O&M costs (i.e., Scenario 2.A.), or substituting PPAs with utility generation (i.e., Scenario 2.B.) alters the utility's non-fuel cost assumptions directly and, therefore, produce the most significant change in shareholder impacts. Across the range of sensitivity cases at 8% DPV, earnings impacts range from a 3.0% reduction (i.e., Scenario 1.B) to a 4.8% reduction (i.e., Scenario 2.A.), and ROE impacts range from a 2.8% reduction (i.e., Scenario 2.B) to a 5.4% reduction (i.e., Scenario 2.A) compared to the case without DPV (on a percentage, not absolute, basis).

Importantly, these percentage reductions are against small reductions in earnings and ROE on an absolute basis. For example, achieved earnings decline by $123M (20-year present value) in the 8% DPV Base case out of a total earnings basis of $3.96B (20-year present value). Even with DPV value assumptions driving the most impactful change in earnings that assume incremental distribution CapEx and O & M (i.e., Scenario 1.A.), the absolute reduction in earnings is $190M (20-year present value). The same is true for achieved average ROE impacts that are a 25 basis-point reduction at 8% DPV in the Base Case. The reduction in achieved ROE assuming incremental distribution CapEx and O & M (i.e., Scenario 1.A.) is 42 basis points.

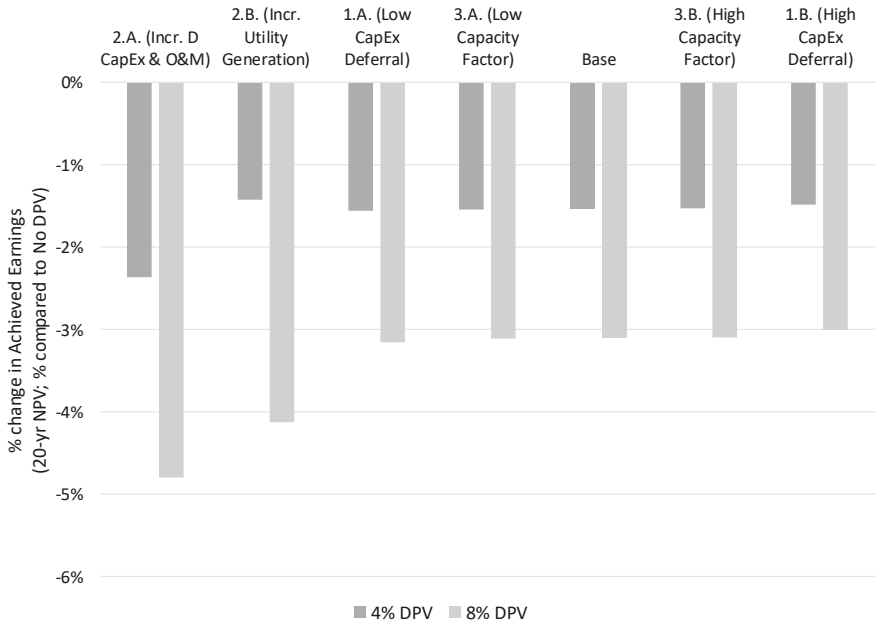


Figure 8. Sensitivity of Western utility achieved earnings to assumptions related to DPV value.

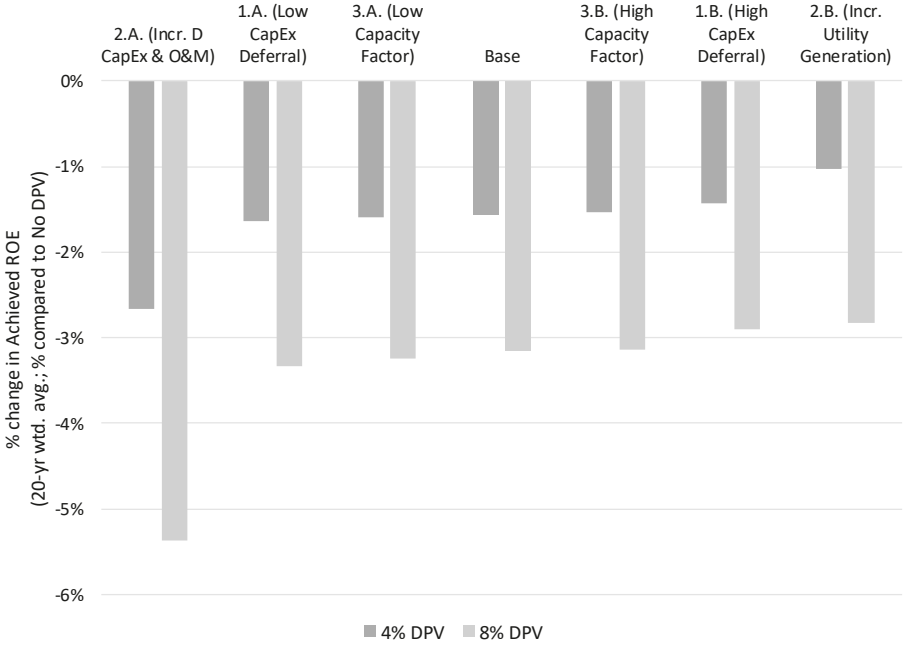


Figure 9. Sensitivity of Western utility achieved average ROE to assumptions related to DPV value.

As shown in Figures 10 and 11, the impacts of DPV on average all-in retail rates and customer bills vary under these different assumptions related to the penetration of DPV as well as the value of

DPV. In the Base Case, the Western utility's all-in average retail rate increases as the DPV deployment level increases (1.1% increase at 4% DPV and a 2.4% increase at 8% DPV); however the reduction in sales associated with DPV exceeds the impact from the rate increase, resulting in a decline in total customer bills (1.8% reduction at 4% DPV and a 3.6% reduction at 8% DPV). Relative to the case without DPV, assumptions driving greater cost deferral (regardless of fuel and purchased power costs or CapEx-related costs), result in lower ratepayer impacts (i.e., lower average rate impacts and greater total customer bill reductions). Across the range of sensitivity cases at 8% DPV, average retail rate impacts range from a 2.3% increase (i.e., Scenario 1.B.) to a 3.4% increase (i.e., Scenario 2.A.) and total customer bill impacts range from a 2.6% reduction (i.e., Scenario 2.A.) to a 3.7% reduction (i.e., Scenario 1.B.) compared to the Base case without DPV. Importantly, these bill savings reflect the aggregate impact across all customers and do not reflect the distribution of customer bill impacts among participating and non-participating customers. However, for non-participants who are not assumed to change their electricity consumption across scenarios, the reported percentage changes in average all-in retail rates is a proxy for their associated bill impacts.

Like the shareholder impacts, though, the percentage reductions are against small changes in ratepayer metrics on an absolute basis. Specifically, average all-in retail rates increase by 0.22 cents/kWh in 8% DPV Base Case and by 0.32 cents/kWh with DPV value assumptions with the most impactful change in average all-in retail rates (i.e., Scenario 2.A.). These changes compared to an average all-in retail rate of 9.20 cents/kWh (20-year present value) without DPV in the Base case. Similarly, there is a $1.31B decrease in *total* customer bills (out of ~$36B basis) at 8% DPV. The total customer bill savings in the high CapEx deferral DPV value sensitivity (i.e., Scenario 1.B.) are $1.35B at 8% DPV.

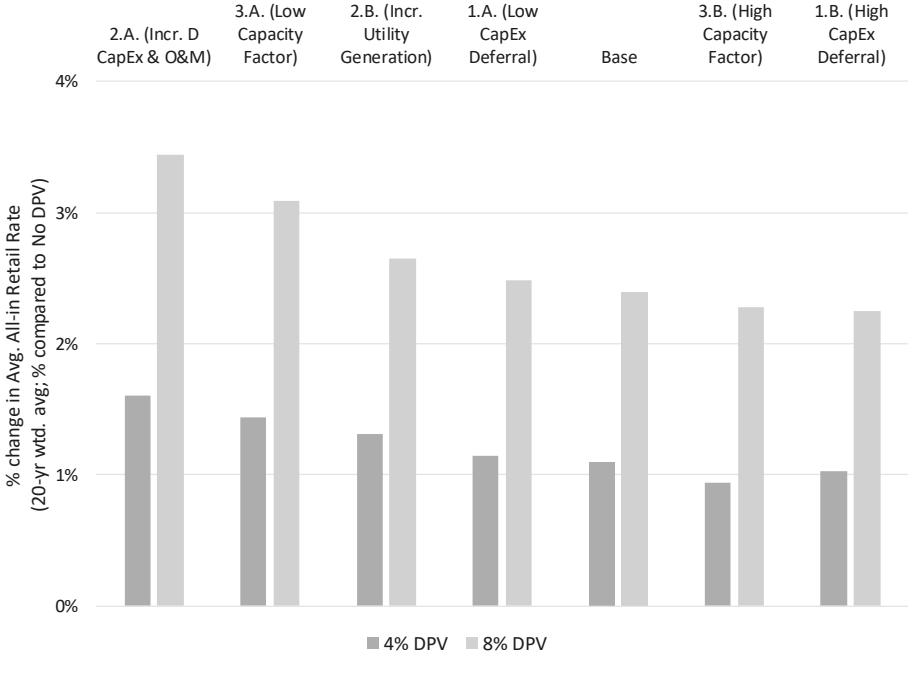


Figure 10. Sensitivity of Western utility average all-in retail rate to assumptions related to DPV value.

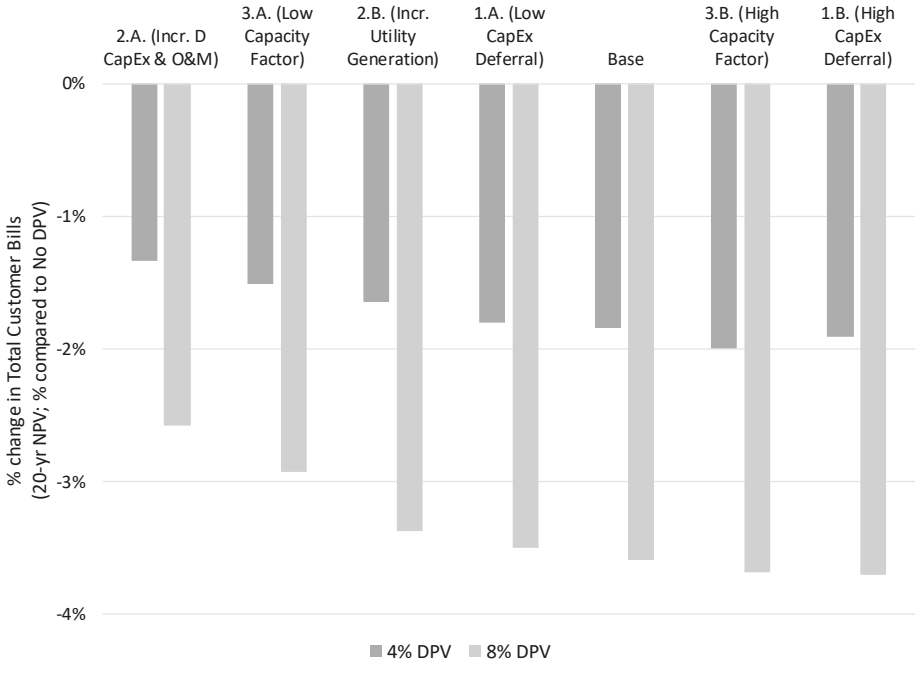


Figure 11. Sensitivity of Western utility total customer bills to assumptions related to DPV value.

4. Discussion and Conclusions

This analysis quantified the financial impacts of different levels of net-metered DPV deployment on a prototypical Western U.S. utility over 20 years and estimated changes in the utility's costs, revenues, achieved shareholder earnings and ROE, average all-in retail rates, and customer bills. We also quantified the sensitivity of results to different assumptions about the ability of DPV to avoid, defer, or increase utility fixed and variable costs.

Our analysis shows that DPV does in fact reduce utility achieved earnings, which occurs through two separate means. First, if DPV reduces utility revenues more than utility costs (as is likely under rate structures that recover the majority of utility costs via volumetric energy charges), then net revenues or earnings are likewise reduced (i.e., the "revenue erosion effect"). Second, and separately, DPV savings may also diminish future earnings opportunities by reducing the rate of growth or deferring capital investments (T & D CapEx in our Base Case assumptions, specifically) that would otherwise contribute to the utility's ratebase (i.e., the "lost earnings opportunity effect"). Although our analysis does illustrate that the financial impacts on shareholders increase as the level of DPV deployment increases, the magnitude is small even at high DPV penetration levels (e.g., 2 to 4% change in financial metrics at 8% DPV deployment).

Our analysis also shows that net-metered DPV increases average all-in retail rates, which occurs in two, interrelated ways. First, DPV affects the retail rates set within each general rate case (GRC) through the net result of reductions in the test-year utility costs and billing determinants used to establish rates. DPV generally tends to reduce utility sales more than costs and, as a result, average retail rates established through each GRC increase with the addition of DPV in order to ensure the utility's rates collect sufficient revenue to recover authorized costs. Second, DPV affects average all-in retail rates in the years between GRCs, though this effect is simply a mathematical artifact. Average all-in rates are, by definition, equal to total collected revenues divided by total retail sales in any given year. Retail sales (i.e., the denominator) are reduced due to the incremental DPV. Because

of these lower volumetric energy billing determinants, the revenues (i.e., the numerator) collected on an annual basis will also be reduced. However, the reduction in revenues are necessarily smaller than the reduction in retail sales, given that some portion of revenues are derived from fixed customer charges (which are unaffected by DPV) and demand charges (which are only marginally affected by DPV). Thus, DPV tends to increase average all-in retail rates in between GRCs as well. As with shareholder metrics, our analysis illustrates that ratepayer financial impacts increase as the level of DPV deployment increases, but the magnitude is small even at 8% DPV penetration levels.

We know that utilities in the Western U.S. are varied, and exist within dramatically different operating environments, face substantially different cost structures, and provide service to vastly different customer bases. To better understand the likely impacts of DPV on utilities in the West, and potentially extend the application of the results more broadly to other utilities around the world, this study also explicitly links different estimates of DPV value to shareholder and ratepayer impacts. Our analysis finds that even rather dramatic changes in DPV value result in modest changes to shareholder and ratepayer impacts. Also, the range of financial impacts under alternative DPV value assumptions are greater for shareholders than ratepayers on a percentage basis and driven by differences in the amount of incremental CapEx that is deferred, as well as the amount of incremental distribution O&M that is incurred. The sensitivity cases reflect the key drivers of our results, but are not a complete list of all the sources of uncertainty and variation in modeled assumptions. There are other utility characteristics that might also change the magnitude and, in more extreme instances, the direction of impacts (e.g., higher or lower assumed load growth, higher or lower proportion of revenues from fixed charges, current or future test years). See [46] for the results of a number of sensitivity cases beyond the value of DPV. To be sure, the shareholder and ratepayer impacts presume no change in the underlying regulatory model or ratemaking approaches. More fundamental changes in the way electric utilities price energy services and recover fixed and variable costs may suggest different impacts than reported herein (e.g., see [51]). As such, our purpose here is not to bound the full range of impacts, but rather to illustrate some key themes and considerations related specifically to DPV value.

It is worth noting two particular feedback effects that our pro forma financial model does not account for in the present study and that would have potential implications for our results. First, we do not represent the feedback effects between retail rate impacts and DPV adoption rates. An increase in retail rates will improve the economics of DPV investment to customers (i.e., lower payback time for PV system) which, all else being equal, would be expected to drive greater PV adoption and thus lead to increased reductions in the utility's future load. Though these effects have been found to be small [52]. Darghouth et al. [52] also addressed a separate feedback mechanism between increasing PV penetration and the timing of time-of-use (TOU) periods; their analysis found that greater PV penetration causes TOU peak periods to shift into the evening hours, which in turn dampens further adoption. Second, we do not represent the feedback effects of changes in retail rate designs or NEM alternatives on overall customer energy consumption (e.g., fixed customer charge may reduce financial incentive to invest in energy efficiency or net billing may encourage DPV system design to maximize exports), all else being equal. Instead, we construct an annual load and PV penetration forecast which is adhered to regardless of these factors. Incorporating such changes into the model will be pursued as a future research effort.

Most Western U.S. utilities, with the exception of some of those in California, currently have distributed generation deployments equivalent to less than 1% of annual retail sales. It will take them several years to see the kinds of impacts depicted in this analysis, at even the 4% penetration level, let alone the 8% level. Accordingly, policymakers and regulators likely have time to study and deliberate changes to NEM, as well as other policy and regulatory actions, before observing material financial impacts. While many of the impacts appear relatively small (on a percentage basis), they demonstrate how underlying ratemaking and regulatory mechanisms can change utility support

for or customer interest in DERs, and the magnitude of impacts depends critically on utility physical, financial, and operating characteristics.

Author Contributions: P.C. wrote the original draft manuscript, contributed to the project's conceptualization, performed a subset of the formal analysis, and reviewed and edited the final manuscript. A.S. provided supervision of the research activity, led the development of the project's conceptualization, performed a subset of the formal analysis, and reviewed and edited the manuscript. W.G. and J.R. contributed to the project's conceptualization, performed a subset of the formal analysis, and reviewed and edited the manuscript. All authors read and approved the final manuscript.

Funding: This work was supported by the U.S. Department of Energy Solar Energy Technologies Office under Lawrence Berkeley National Laboratory Contract No. DE-AC02-05CH11231. This manuscript has been authored by an author at Lawrence Berkeley National Laboratory under Contract No. DE-AC02-05CH11231 with the U.S. Department of Energy. The U.S. Government retains, and the publisher, by accepting the article for publication, acknowledges, that the U.S. Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for U.S. Government purposes.

Conflicts of Interest: The authors acknowledge that one of them is also an editor for this journal.

Appendix A FINDER Model Overview

The FINDER pro forma financial model was developed to quantify the financial impacts on ratepayers, utilities, and shareholders from the deployment of DERs as well as the introduction of alternative regulatory/business models. The basic structure of the model is depicted in Figure A1. What follows is a more detailed description of the different modules within the model.

Utility costs are based on model inputs that characterize current and projected utility costs over the analysis period. The model represents major cost categories of the utility's physical, financial, and operating environment, including fuel and purchased power, operations and maintenance, and capital investments in generation and non-generation assets (i.e., transmission and distribution investments). Some costs are projected using stipulated first year values and compound annual growth rates (CAGRs); other costs are based on schedules of specific investments (e.g., generation expansion plans). The model calculates the utility's ratebase over the analysis period, accounting for increases due to additional capital investments as well as decreases due to depreciation of existing assets. The model also estimates interest payments for debt and returns for equity shareholders based on an authorized amount used to finance capital investments and includes taxes on earnings.

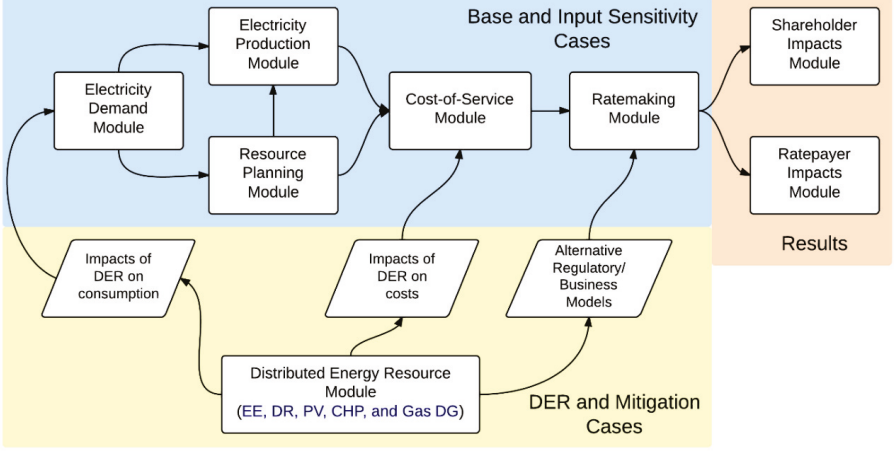


Figure A1. FINDER Model Overview.

The utility's collected revenues are based on retail rates that are set in periodic general rate cases (throughout the analysis period. By default, the model assumes that a GRC occurs at some specified

frequency (e.g., every three years); the model also allows the utility to file a GRC that may be triggered by a significant capital investment (e.g., new power plant, proposal to install advanced metering infrastructure).

GRCs are used to establish new rates for customers based on the revenue requirement set in a test year, including an authorized ROE for capital investments, the test year billing determinants (i.e., retail sales, peak demand, and number of customers), and assumptions about how the test year revenue requirement is allocated to customer classes and among the billing determinants. The model allows for different types of test years (i.e., historical, current, and future test years). Many states allow the utility to file an adjustment to its historical test-year costs during a GRC (i.e., pro forma adjustment period) to update and correct them to reflect expectations about normal cost levels, however, our model uses unadjusted historic test year values for ratemaking purposes. The particular rate design of the utility consists of a combination of a volumetric energy charge ($/kWh), volumetric demand charge ($/kW), and fixed customer charge ($/customer) for a particular customer class. Model inputs specify the relative share of different types of utility costs that are collected from each of these three rate components.

The rates established in a GRC are then applied to the actual billing determinants in future years to calculate utility collected revenue in those years. The model accounts for a period of regulatory lag whereby rates that are established in a GRC do not go into effect until some specified number of years after the GRC. In between general rate cases, certain costs are passed directly to customers through rate-riders (e.g., fuel-adjustment clause or FAC). The model derives an average all-in retail rate metric that reflects the average revenue collected per unit of sales at the utility or customer-class level and accounts for periodic setting of new rates, rate-riders, and delays in implementing new rates.

The financial performance of the utility is measured by achieved after-tax earnings and achieved after-tax ROE. We calculated the prototypical utilities' achieved after-tax ROE in each year as the current year's earnings divided by current year's outstanding equity (i.e., the equity portion of the ratebase). The FINDER Model does not take into account changes in financing costs that may result from under- or over-recovery of costs, which may impact ROE. Achieved after-tax ROE may, and often does, differ from the utility's authorized ROE. The authorized ROE is typically established by regulators during a regulatory proceeding and used in a GRC to determine the amount of return that a utility may receive on its capital investments. Actual utility revenues and costs may—and nearly always do—differ from those in the test year, leading to achieved earnings, and hence achieved ROE, that deviates from the authorized level. In a GRC, utility rates are set such that the test-year revenue requirement produces earnings that are sufficient to reach the authorized after-tax ROE based on the test year costs and billing determinants. In general, achieved ROE will be less than authorized ROE if, between rate cases, utility costs grow faster than revenues. Conversely, achieved ROE will generally be greater than authorized ROE if utility costs grow slower than revenues between rate cases.

FINDER calculates the prototypical utilities' achieved after-tax earnings as collected revenues minus costs in each year. Achieved after-tax earnings can be different than the utility's authorized earnings, because the achieved earnings are based on actual profitability in a given year and the authorized earnings are set in the GRC revenue requirement, based on the authorized ROE. Technically, state regulators do not explicitly authorize earnings in a GRC; they authorize a ROE, which when applied to the undepreciated portion of a utility's share of equity financed ratebase produces a level of earnings. For our purposes in this report, we refer to that as authorized earnings.

Alternative regulatory mechanisms and rate structures can also be implemented in FINDER: decoupling mechanisms (i.e., sales based or revenue-per-customer), lost revenue adjustment mechanisms, and shareholder incentive mechanisms. Alternative rate structures (e.g., high fixed customer charge) are represented by changing the way utility revenues are collected among different billing determinants.

Appendix B Detailed Sensitivity Case Results

Table A1 shows the sensitivity case results expressed in percentage changes. The change in DPV value is relative to the DPV value in the Base case at the respective DPV deployment level (i.e., 4% or 8%). The change in achieved earnings and average all-in retail rates is relative to a case without DPV.

Table A1. Full sensitivity case results at 4% and 8% DPV.

	Sensitivity Case	Change in DPV Value (% Relative to Base Case)	Change in Earnings (% Change Relative to Respective no DPV)	Change in Average Retail Rates (% Change Relative to no DPV
4% DPV Deployment	Base	n/a	−1.5%	1.1%
	1.A. (Low Capacity Contribution)	−4.3%	−1.5%	1.4%
	1.B. (High Capacity Contribution)	6.1%	−1.5%	0.9%
	2.A. (Incremental D CapEx & OpEx)	−87.7%	−2.4%	1.6%
	2.B. (Incremental Utility-Owned Generation)	−9.5%	−1.4%	1.3%
	3.A. (Low CapEx Deferral)	−27.3%	−1.6%	1.1%
	3.B. (High CapEx Deferral)	10.5%	−1.5%	1.0%
8% DPV Deployment	Base	n/a	−3.1%	2.4%
	1.A. (Low Capacity Contribution)	−4.4%	−3.1%	3.1%
	1.B. (High Capacity Contribution)	6.3%	−3.1%	2.3%
	2.A. (Incremental D CapEx & OpEx)	−93.4%	−4.8%	3.4%
	2.B. (Incremental Utility-Owned Generation)	−9.9%	−4.1%	2.7%
	3.A. (Low CapEx Deferral)	−28.5%	−3.2%	2.5%
	3.B. (High CapEx Deferral)	3.7%	−3.0%	2.3%

References

1. SEIA/Wood Mackenzie. *Solar Market Insight 2018 Year in Review*; Solar Energy Industries Association and Wood Mackenzie Power & Renewables: Washington, DC, USA, 2016.
2. Barbose, G.; Darghouth, N.; LaCommare, K.H.; Millstein, D.; Rand, J. *Tracking the Sun: Installed Price Trends for Distributed Photovoltaic Systems in the United States-2018 ed.*; Berkeley Lab: Berkeley, CA, USA, 2018.
3. Sergici, S.; Yang, Y.; Castaner, M.; Faruqui, A. Quantifying net energy metering subsidies. *Electr. J.* **2019**, *32*, 106632. [CrossRef]
4. Gautier, A.; Jacqmin, J.; Poudou, J.-C. The prosumers and the grid. *J. Regul. Econ.* **2018**, *53*, 100–126. [CrossRef]
5. EEI. *A Policy Framework for Designing Distributed Generation Tariffs*; Edison Electric Institute: Washington, DC, USA, 2013.
6. *The 50 States of Solar: 2017 Policy Review and Q4 2017 Quarterly Report*; Clean Energy Technology Center: Raleigh, NC, USA, 2018.
7. Gagnon, P.; Sigrin, B.; Gleason, M. *The Impacts of Changes to Nevada's Net Metering Policy on the Financial Performance and Adoption of Distributed Photovoltaics*; National Renewable Energy Laboratory: Golden, CO, USA, 2017.
8. Barbose, G.; Miller, J.; Sigrin, B.; Reiter, E.; Cory, K.; McLaren, J.; Seel, J.; Mills, A.D.; Darghouth, N.R.; Satchwell, A. *Utility Regulatory and Business Model Reforms for Addressing the Financial Impacts of Distributed Solar on Utilities*; Berkeley Lab: Berkeley, CA, USA, 2016.
9. Martín-Martínez, F.; Sánchez-Miralles, A.; Rivier, M.; Calvillo, C.F. Centralized vs distributed generation. A model to assess the relevance of some thermal and electric factors. Application to the Spanish case study. *Energy* **2017**, *134*, 850–863. [CrossRef]
10. Averch, H.; Johnson, L. Behavior of the Firm Under Regulatory Constraint. *Am. Econ. Assoc.* **1962**, *51*, 1052–1069.
11. Barbose, G. *Putting the Potential Rate Impacts of Distributed Solar into Context*; Berkeley Lab: Berkeley, CA, USA, 2017.
12. Sandiford, M.; Forcey, T.; Pears, A.; McConnell, D. Five Years of Declining Annual Consumption of Grid-Supplied Electricity in Eastern Australia: Causes and Consequences. *Electr. J.* **2015**, *28*, 96–117. [CrossRef]

13. Schittekatte, T.; Momber, I.; Meeus, L. Future-proof tariff design: Recovering sunk grid costs in a world where consumers are pushing back. *Energy Econ.* **2018**, *70*, 484–498. [CrossRef]
14. Borenstein, S. The economics of fixed cost recovery by utilities. *Electr. J.* **2016**, *29*, 5–12. [CrossRef]
15. Satchwell, A.J.; Cappers, P.A. Recent developments in competition and innovation for regulated electric utilities. *Util. Policy* **2018**, *55*, 110–114. [CrossRef]
16. Gorman, W.; Mills, A.; Wiser, R. Improving estimates of transmission capital costs for utility-scale wind and solar projects to inform renewable energy policy. *Energy Policy* **2019**, *135*. [CrossRef]
17. Cohen, M.A.; Kauzmann, P.A.; Callaway, D.S. Effects of distributed PV generation on California's distribution system, part 2: Economic analysis. *Solar Energy* **2016**, *128*, 139–152. [CrossRef]
18. Cossent, R.; Gomex, T.; Olmos, L.; Mateo, C.; Frias, P. Assessing the Impact of Distributed Generation on Distribution Network Costs. In Proceedings of the 2009 6th International Conference on the European Energy Market, Leuven, Belgium, 27–29 May 2009.
19. E3. *Introduction to the Net Energy Metering Cost Effectiveness Evaluation*; CPUC: San Francisco, CA, USA, 2010.
20. Flinn, J.; Webber, C. *Residential Zero Net Energy Building Integration Cost Analysis*; DNV GL: Oslo, Norway, 2017.
21. Horowitz, K.A.W.; Palmintier, B.; Mather, B.; Denholm, P. Distribution system costs associated with the deployment of photovoltaic systems. *Renew. Sustain. Energy Rev.* **2018**, *90*, 420–433. [CrossRef]
22. Mills, A.D.; Barbose, G.; Seel, J.; Dong, C.; Mai, T.; Sigrin, B.; Zuboy, J. *Planning for a Distributed Disruption: Innovative Practices for Incorporating Distributed Solar into Utility Planning*; LBNL: Berkeley, CA, USA, 2016.
23. Sena, S.; Quiroz, J.; Broderick, R. Analysis of 100 Utility SGIP PV Interconnection Studies. In Proceedings of the 2014 IEEE 40th Photovoltaic Specialist Conference (PVSC), Denver, CO, USA, 8–13 June 2014.
24. Wolak, F. *The Evidence from California on the Economic Impact of Inefficient Distribution Network Pricing*; NBER: Cambridge, MA, USA, 2018.
25. Passey, R.; Watt, M.; Outhred, H.; Spooner, T.; Snow, M. *Study of Grid-connect Photovoltaic Systems-Benefits, Opportunities, Barriers, and Strategies*; University of New South Wales: Sydney, Australia, 2007.
26. Stanton, E.; Daniel, J.; Vitolo, T.; Knight, P.; While, D.; Keith, G. *Net Metering in Mississippi: Costs, Benefits, and Policy Considerations*; For PSC of Mississippi; Synapse: Cambridge, MA, USA, 2014.
27. Xcel. *Costs and Benefits of Distributed Solar Generation on the Public Service Compnay of Colorado System*; Colorado PUC: Denver, CO, USA, 2013.
28. Hoff, T.E.; Norris, B.; Wayne, G. *Potential Economic Benefits of Distributed Photovoltaics to the Nevada Power Company*; Clean Power Research: Napa, CA, USA, 2003.
29. Pudjianto, D.; Djapic, P.; Dragovic, J.; Strbac, G. *Direct Costs Analysis related to Grid Impacts of Photovoltaics*; Imperial College London: London, UK, 2013.
30. TVA. *Distributed Generation-Integrate Value: A Methodology to Value DG on the Grid*; Tennessee Valley Authority: Knoxville, TN, USA, 2015.
31. Beck, R.W.; Solutions, E.; Company, P.E.; Consulting, S.B. *Distributed Renewable Energy Operating Impacts and Valuation Study*; For Arizona Public Service; R.W. Beck: Seattle, WA, USA, 2009.
32. Perez, R.; Zweibel, K.; Hoff, T.E. Solar power generation in the US: Too expensive, or a bargain? *Energy Policy* **2011**, *39*, 7290–7297. [CrossRef]
33. Poullikkas, A. A comparative assessment of net metering and feed in tariff schemes for residential PV systems. *Sustain. Energy Technol. Assess.* **2013**, *3*, 1–8. [CrossRef]
34. Christoforidis, G.; Panapakidis, I.; Papadopoulos, T.; Papagiannis, G.; Koumparou, I.; Hadjipanayi, M.; Georghiou, G. A Model for the Assessment of Different Net-Metering Policies. *Energies* **2016**, *9*, 262. [CrossRef]
35. Eid, C.; Reneses Guillén, J.; Frías Marín, P.; Hakvoort, R. The economic effect of electricity net-metering with solar PV: Consequences for network cost recovery, cross subsidies and policy objectives. *Energy Policy* **2014**, *75*, 244–254. [CrossRef]
36. Picciariello, A.; Vergara, C.; Reneses, J.; Frías, P.; Söder, L. Electricity distribution tariffs and distributed generation: Quantifying cross-subsidies from consumers to prosumers. *Util. Policy* **2015**, *37*, 23–33. [CrossRef]
37. Johnson, E.; Beppler, R.; Blackburn, C.; Staver, B.; Brown, M.; Matisoff, D. Peak shifting and cross-class subsidization: The impacts of solar PV on changes in electricity costs. *Energy Policy* **2017**, *106*, 436–444. [CrossRef]
38. Satchwell, A.; Cappers, P.; Goldman, C. Customer bill impacts of energy efficiency and net-metered photovoltaic system investments. *Util. Policy* **2018**, *50*, 144–152. [CrossRef]

39. Satchwell, A.; Mills, A.; Barbose, G. Quantifying the financial impacts of net-metered PV on utilities and ratepayers. *Energy Policy* **2015**, *80*, 133–144. [CrossRef]
40. WECC. *TEPPC Study Report: 2026 PC1 Common Case*; Western Electricity Coordinating Council: Salt Lake City, UT, USA, 2017.
41. Darghouth, N.; Barbose, G.; Wiser, R. The impact of rate design and net metering on bill savings from distributed PV for residential customers in California. *Energy Policy* **2011**, *39*, 5243–5253. [CrossRef]
42. Darghouth, N.; Barbose, G.; Wiser, R. *Electricity bill Savings from Residential Photovoltaic Systems: Sensitivities to Changes in Future Electricity Market Conditions*; Berkeley Lab: Berkeley, CA, USA, 2013.
43. Darghouth, N.; Barbose, G.L.; Mills, A.D.; Wiser, R.H.; Gagnon, P.; Bird, L. *Exploring Demand Charge Savings from Residential Solar*; Berkeley Lab: Berkeley, CA, USA, 2017.
44. EIA. *Capital Cost Estimates for Utility Scale Electricity Generating Plants*; Energy Information Administration: Washington, DC, USA, 2016.
45. National Renewable Energy Laboratory Annual Technology Baseline. Data and Documentation. Available online: https://atb.nrel.gov/electricity/2018/index.html (accessed on 18 June 2018).
46. Satchwell, A.; Mills, A.; Barbose, G.; Wiser, R.; Cappers, P.; Darghouth, N. *Financial Impacts of Net-Metered PV on Utility and Ratepayers: A Scoping Study of Two Prototypical US Utilities*; Berkeley Lab: Berkeley, CA, USA, 2014.
47. National Renewable Energy Laboratory Distributed Generation Energy Technology Operations and Maintenance Costs. Data and Documentation. Available online: https://www.nrel.gov/analysis/tech-cost-om-dg.html (accessed on 8 August 2018).
48. CPUC. *Residential Zero Net Energy Building Integration Cost Analysis*; California Public Utilities Commission: San Francisco, CA, USA, 2017.
49. Blair, N.; DiOrio, N.; Freeman, J.; Gilman, P.; Janzou, S.; Neises, T.; Wagner, M. *System Advisor Model (SAM) General Description (Version 2017.9.5)*; National Renewable Energy Laboratory: Golden, CO, USA, 2018.
50. Rhodes, J.D.; Upshaw, C.R.; Cole, W.J.; Holcomb, C.L.; Webber, M.E. A Multi-objective Assessment of the Effect of Solar PV Array Orientation and Tilt on Energy Production and System Economics. *Sol. Energy* **2014**, *108*, 28–40. [CrossRef]
51. Satchwell, A.; Cappers, P.; Schwartz, L.; Fadronc, E.M. *A Framework for Organizing Current and Future Electric Utility Regulatory and Business Models*; Berkeley Lab: Berkeley, CA, USA, 2015.
52. Darghouth, N.; Wiser, R.H.; Barbose, G.L.; Mills, A. Net Metering and Market Feedback Loops: Exploring the Impact of Retail Rate Design on Distributed PV Deployment. *Appl. Energy* **2016**, *162*, 713–722. [CrossRef]

energies

MDPI

Article

Strategic-Agent Equilibria in the Operation of Natural Gas and Power Markets

Sheng Chen [1,*] and Antonio J. Conejo [2]

[1] College of Energy and Electrical Engineering, Hohai University, Nanjing 210098, China
[2] Department of Integrated Systems Engineering and Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210, USA; conejo.1@osu.edu
* Correspondence: chenshenghhu@163.com

Received: 20 January 2020; Accepted: 11 February 2020; Published: 17 February 2020

Abstract: We consider strategic gas/power producers and strategic gas/power consumers operating in both gas and power markets. We build a flexible multi-period complementarity model to characterize day-ahead equilibria in those markets. This model is an equilibrium program with equilibrium constraints that characterizes the market behavior of all market agents. Using a realistic case study, we analyze equilibria under perfect and oligopolistic competition. We also analyze equilibria under different levels of information disclosure regarding market outcomes. We study as well equilibria under different ownership schemes: no hybrid agent, some hybrid agents, and only hybrid agents. Finally, we derive policy recommendations for the regulators of both the gas and the power markets.

Keywords: natural-gas market; electricity market; equilibrium analysis

1. Introduction

Electric power systems and natural-gas systems are generally operated independently, with limited or no coordination [1]. This is the result of how these systems were created and have evolved over time. In fact, gas has not played a significant role as a primary fuel in electricity production until recently, and thus, gas and power system coordination has not been important until recently.

Due to the increasing availability of gas and its competitive price, during the last decade, an increasing number of combined cycle gas turbines (CCGTs) have been incorporated into the generation mix of many power systems. This has resulted in an increasingly strong interdependency between gas systems and power systems [2]. In fact, this interdependency can no longer be disregarded if the gas and power systems are to be operated efficiently [1].

However, tools to comprehend the effect of such interdependency are limited. Many of these tools adopt a centralized perspective, in which a single operator manages both the gas and power systems [3–14] , which is unrealistic. Representative references are briefly discussed below. Chen et al. [3] develop a unit commitment model that includes an enhanced second order conic gas flow model, where the interdependency between gas and power prices is investigated. Byeon and Van Hentenryck [4] introduce a unit commitment problem with gas network awareness, where bid-validity constraints are imposed on gas-fired units. He et al. [5] propose an integrated gas and power system operation model that considers demand response and uncertainty via distributionally robust optimization. He et al. [6] develop a decentralized operation model for multi-area gas and power systems. Chen et al. [7] develop a joint gas and power market model that addresses wind power uncertainty and gas system congestion. Ameli et al. [8] quantify the value of the flexibility of the gas system in accommodating intermittent renewable energy sources. Yang et al. [9]

propose a two-stage robust operation model that considers gas network dynamics and wind power uncertainty. Bai et al. [10] develop a robust scheduling model that considers N-1 contingencies of power transmission lines or gas pipelines. Zlotnik et al. [11] analyze the economic and security benefits of a coordinated scheduling of interdependent gas and power systems. Chen et al. [12] propose a two-stage robust day-ahead dispatch model for urban electric and gas systems. Antenucci and Sansavini [13] investigate the impacts of gas-system operational constraints on a stochastic unit commitment model with large renewable penetration. Ordoudis et al. [14] develop an integrated electricity and gas market-clearing model, in which the value of the gas system flexibility to accommodate high shares of renewables is discussed.

Complementarily, Ref. [15] proposes an equilibrium model of the type we propose in this paper, but for distribution systems and [16] describes an equilibrium model at the bulk level, but uses and heuristic solution approach.

We propose in this paper an equilibrium model that allows studying the interactions of both gas/power producers and gas/power consumers (referred generically to as agents) through both the gas and the power markets. This model expands the one reported in [17] as it considers a multi-period framework and carries out a comprehensive analysis. Each market agent (producer of consumer) is represented as a bi-level model (see Appendix A.3 of the Appendix) with an upper-level problem that pursues maximum profit (revenue minus cost or utility minus payment) for the agent (see Appendices A.3.1 and A.3.2, respectively of the Appendix), and two lower-level problems representing the clearing of the gas and the power markets (see Appendices A.1 and A.2 , respectively, of the Appendix). We then jointly consider the bi-levels problems of all the agents participating in the gas and power markets, and solve the resulting Equilibrium Problem with Equilibrium Constraints (EPEC) using a direct approach [18,19] that does not rely on heuristics.

We consider hybrid producers that own both gas and power production facilities as well as non-hybrid ones. Likewise we consider hybrid consumers that consume both gas and electricity and non-hybrid ones.

The study horizon that we consider for both the gas and the power markets is one day divided in a number of periods to capture inter-temporal effects, such as steep ramping requirements due to the variability of the production of renewable units.

The proposed model represents in detail the gas and power network, the latter using linear (dc) equations (see Appendix A.2 of the Appendix) and the former via second order conic equations (see Appendix A.1 of the Appendix).

We consider that gas/power producers and gas/power consumers are both strategic and seek to alter gas/power clearing prices to their respective benefits and analyze equilibria under three conditions, namely:

1. Perfect and imperfect competition.
2. Aggregated price information from the gas market, as in [20], which is common in practice.
3. Diverse ownership of the gas and power facilities, including no hybrid agent, some hybrid agents and only hybrid agents.

The equilibrium analysis reported in this paper is particularly relevant to the regulator, as it helps devising market adjustments and coordination rules to maximize social welfare in both the gas and power markets.

The contributions of this paper are twofold:

1. To formulate and solve a multi-period EPEC to characterize the outcomes of interrelated gas and power markets with strategic agents.
2. To analyze market outcomes under (i) different degrees of imperfect competition, (ii) market-clearing information granularity, and (iii) ownership structure.

The rest of this paper is organized a follows. Section 2 describes in a generic manner the bi-level model of a strategic agent (producer or consumer), Section 3 describes the considered EPEC, Section 4 shows how to solve it, Section 5 provides an illustrative example, Section 6 and 7 describe and discuss results from two realistic test systems, and Section 8 draw conclusions. The Appendix provides detailed descriptions of the models considered and metrics used.

2. Single-Agent Model

A generic bi-level model to represent the profit-seeking behavior of a single strategic agent (producer or consumer) is provided below:

$$\max_{\Xi^{(i)}} \quad \pi^{(i)}\left(x_g^{(i)}, x_e^{(i)}, \lambda_g^{(i)}, \lambda_e^{(i)}\right) \tag{1}$$

$$\text{s.t.} \quad o_g^{(i)} \in \mathcal{O}_g^{(i)}, o_e^{(i)} \in \mathcal{O}_e^{(i)} \tag{2}$$

$$\min_{x_g} \quad f_g(x_g, o_g) \tag{3}$$

$$\text{s.t.} \quad h_g(x_g) = 0 : \ \lambda_g \tag{4}$$

$$g_g(x_g, o_g) \leq 0 : \ \mu_g \tag{5}$$

$$\min_{x_e} \quad f_e(x_e, o_e) \tag{6}$$

$$\text{s.t.} \quad h_e(x_e) = 0 : \ \lambda_e \tag{7}$$

$$g_e(x_e, o_e) \leq 0 : \ \mu_e, \tag{8}$$

where $\Xi^{(i)} = \{o_g^{(i)}, o_e^{(i)}\} \cup \{x_g, x_e, \lambda_g, \mu_g, \lambda_e, \mu_e\}$.

The notation used is described below:

$\pi^{(i)}(\cdot)$ is the profit of agent (i),
x_g the vector of gas variables,
$x_g^{(i)}$ the sub-vector (of vector x_g) of gas variables that pertains to agent (i),
λ_g, μ_g vectors of dual gas variables,
$\lambda_g^{(i)}$ the sub-vector (of vector λ_g) of dual gas variables that pertains to agent (i),
x_e the vector of power variables,
$x_e^{(i)}$ the sub-vector (of vector x_e) of power variables that pertains to agent (i),
λ_e, μ_e vectors of dual power variables,
$\lambda_e^{(i)}$ the sub-vector (of vector λ_e) of dual power variables that pertains to agent (i),
o_g the gas offer/bid vector,
$o_g^{(i)}$ the gas offer/bid sub-vector (of vector o_g) pertaining to agent (i),
o_e the power offer/bid vector,
$o_e^{(i)}$ the power offer/bid sub-vector (of vector o_e) pertaining to agent (i),
$\mathcal{O}_g^{(i)}$ the feasible set of gas offers/bids of agent (i), and
$\mathcal{O}_e^{(i)}$ the feasible set of power offers/bids of agent (i).

Upper-level problem (1) and (2) represents the profit of the agent (revenue minus cost for a producer and utility minus payment for a consumer), while lower-level problems (3)–(5) and (6)–(8) represent the clearing of the gas and power markets, respectively.

The detailed models of a strategic gas/power consumer and a strategic gas/power producer are provided in Appendices A.3.1 and A.3.2, respectively, of the Appendix. Detailed descriptions of the gas clearing model (3)–(5) and the power clearing model (6)–(8) are provided in Appendices A.1 and A.2, respectively, of the Appendix.

Assuming that lower-level problems (3)–(5) and (6)–(8) are convex or have been convexified [21], we replace them with their corresponding Karush-Kuhn-Tucker (KKT) optimality conditions [18,19,22], rendering the Mathematical Program with Equilibrium Constraints (MPEC) below:

$$\max_{\Xi^{(i)}} \quad \pi^{(i)}\left(x_g^{(i)}, x_e^{(i)}, \lambda_g, \lambda_e\right) \tag{9}$$

$$\text{s.t.} \quad o_g^{(i)} \in \mathcal{O}_g^{(i)}, o_e^{(i)} \in \mathcal{O}_e^{(i)} \tag{10}$$

$$\nabla_{x_g} f_g(\cdot) + \lambda_g^\top \nabla_{x_g} h_g(\cdot) + \mu_g^\top \nabla_{x_g} g_g(\cdot), \quad h_g(\cdot) = 0, \quad 0 \le \mu_g \perp g_g(\cdot) \le 0 \tag{11}$$

$$\nabla_{x_e} f_e(\cdot) + \lambda_e^\top \nabla_{x_e} h_e(\cdot) + \mu_e^\top \nabla_{x_e} g_e(\cdot), \quad h_e(\cdot) = 0, \quad 0 \le \mu_e \perp g_e(\cdot) \le 0, \tag{12}$$

Since MPEC (9)–(12) might be complex to solve/transform and considering that the gas problem is formulated as a second order conic problem (SOCP) [21] and that the power problem is formulated as a linear programming problem, each of these problems can be replaced by its primal constraints, its dual constraints, and its strong duality equality. Thus, instead of considering (9)–(12), we consider:

$$\max_{\Xi^{(i)}} \quad \pi^{(i)}\left(x_g^{(i)}, x_e^{(i)}, \lambda_g, \lambda_e\right) \tag{13}$$

$$\text{s.t.} \quad o_g^{(i)} \in \mathcal{O}_g^{(i)}, o_e^{(i)} \in \mathcal{O}_e^{(i)} \tag{14}$$

$$\textit{primal-constraints}_g, \quad \textit{dual-constraints}_g, \quad \textit{strong-duality-equality}_g \tag{15}$$

$$\textit{primal-constraints}_e, \quad \textit{dual-constraints}_e, \quad \textit{strong-duality-equality}_e. \tag{16}$$

Problem (13)–(16) is generally better behaved than problem (9)–(12), and the KKT optimality conditions of (13)–(16) (single agent optimality conditions) are easily obtained [22] and represented as:

$$\text{KKT}^{(i)} \tag{17}$$

Deriving KKT conditions is a relatively simple exercise. For example, the solver EMP (Extended Mathematical Programming), (https://www.gams.com/latest/docs/UG_EMP.html) which is available in GAMS (General Algebraic Modeling System) (https://www.gams.com), derives KKT conditions automatically.

We note that since problem (13)–(16) is generally non-convex and its constraints might be non-regular, its optimality conditions as given by (17) identify points that might or might not be extrema.

3. Multiple-Agent Model: EPEC

To search for equilibria, we jointly solve (17) for all market agents, which constitutes an Equilibrium Problem with Equilibrium Constraints (EPEC) [19]. This is expressed as:

$$\left\{\text{KKT}^{(i)} \quad \forall i\right., \tag{18}$$

which is a system of nonlinear equalities and inequalities difficult to solve. How to solve EPEC (18) is addressed in Section 4 below.

We note that since the constraints of problem (13)–(16) might be non-regular, (18) identifies equilibria and other stationary points [23].

4. EPEC Solution

The auxiliary problem below can be used to solve (18), i.e., to search for equilibria:

$$\max \quad o(\cdot) \tag{19}$$

$$\text{s.t.} \quad \text{KKT}^{(i)} \quad \forall i, \tag{20}$$

where $o(\cdot)$ is a suitable objective function. We consider in the example and case study (Section 5 and 6, respectively) three objective functions (19), namely:

1. Total Producers' Profit (TPP).
2. Total Consumers' Profit (TCP).
3. Social Welfare of both markets (SW).

The corresponding EPECs (19)–(20) are referred to as:

1. Max TPP EPEC.
2. Max TCP EPEC.
3. Max SW EPEC.

Since (19)–(20) is generally nonlinear and non-convex, its solution can be attempted via linearization or using global solvers, such as BARON [24].

Once potential equilibrium points (solutions of (19)–(20)) are found, a diagonalization algorithm [22] can be used to verify if these points are indeed equilibria.

5. Illustrative Example

For the sake of illustration, we consider in this section a simple example. We analyze a two-bus power system (bus is used to refer to a power-system node) and a two-node gas system (node is used to refer to a gas node), the topology of which is shown in Figure 1. The gas-fired power unit at power bus 2 receiving gas from gas node 2 couples the two systems.

We consider two hybrid agents:

1. Agent 1 owns power unit 1 and gas source 1
2. Agent 2 owns power unit 2 and gas source 2.

For simplicity, we do not consider strategic bids by consumers in this example. In addition, we consider a perfect gas price information interchange between the gas market and the owner of gas-fired power unit 2 (Agent 2).

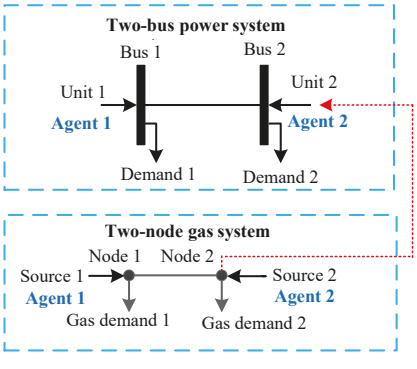


Figure 1. Example: two-bus power system and two-node gas system.

5.1. Data

The capacities of the two power units at buses 1 and 2 are 50 MW and 20 MW, respectively. The marginal production cost of the power unit at bus 1 is 18 \$/MWh. The non-fuel cost of the gas-fired unit at bus 2 is 1 \$/MWh, and its energy conversion coefficient associated with gas consumption is 0.0045 Mm^3/MWh.

Regarding the two gas sources at nodes 1 and 2, their capacities are 0.5 Mm^3/h and 0.7 Mm^3/h, respectively, and their marginal production cost are 3000 \$/Mm^3 and 3500 \$/Mm^3, respectively.

The transmission capacity of the power transsmission line connecting buses 1 and 2 is 18 MW. The lower and upper gas pressure limits at gas nodes are 25 bar and 40 bar, respectively. We note that these gas nodal pressure bounds do not restrict the gas flows through the pipeline connecting nodes 1 and 2.

The baseline utility of the power demands at buses 1 and 2 are 30 \$/MWh and 35 \$/MWh, respectively. The baseline utility of the gas demands at buses 1 and 2 are 4000 \$/Mm^3 and 4200 \$/Mm^3, respectively. The marginal utility factors of both gas and power demands during time periods 1–8, 9–16, and 17–24 are 0.8, 1.0, and 1.2 relative to their baseline utilities, respectively.

Finally, Figure 2 depicts the 24-h total non-generation-related gas demand and the total electricity demand.

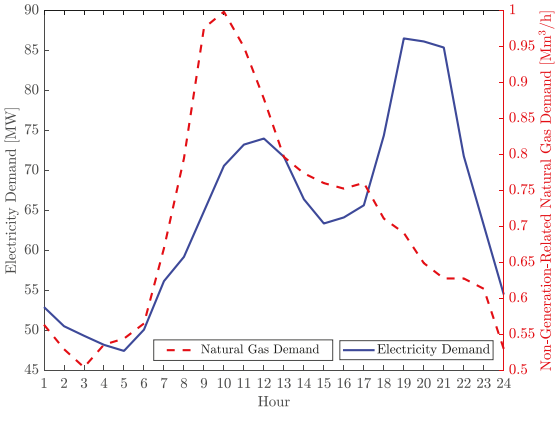


Figure 2. Example: non-generation-related gas demand and power demand.

5.2. Results

We considered two equilibrium models (19)–(20), whose objective functions were total producers' profit and social welfare of both markets, i.e., Max TPP EPEC, and Max SW EPEC, respectively. Table 1 summarizes the market equilibria obtained from the two models. We observed that the equilibrium model that maximized TPP yielded a lower SW but a higher TPP than the corresponding SW and TPP obtained from the equilibrium model that maximized SW. In addition, these two equilibrium models resulted in differences in the distribution of profits between the two production agents. Specifically, Agent 1 earned a higher profit from the model that maximized SW, while the model that maximized TPP was more beneficial for Agent 2.

Table 1. Example: profits and social welfare (\$ thousand) for two equilibrium models.

Equilibrium	**Profit**		**TPP**	**SW**
Model	**Agent 1**	**Agent 2**		
Max TPP	24.4	10.9	35.3	38.6
Max SW	24.6	10.4	35.0	39.2

Additionally, we considered a gas-shortage case, where the capacity of gas-fired unit 2 was reduced to 10 MW. Table 2 provides results for the base case and the gas-shortage case obtained from the Max TPP EPEC. This table shows that the gas-shortage case resulted in a higher profit for Agent 1, earned from the electricity market. This is because power unit 1 accounted for an increased share of electricity supply. Additionally, the gas shortage resulted in lower profits for the two production agents earned from the gas market due to reduced generation-related gas demands.

These results show how the operation of the gas system impacts production agents' profits earned from both gas and power markets. In practice, gas-fired power producers should be aware of potential gas-system bottlenecks, which determine the availability and reliability of their fuel supply.

Table 2. Example: profits and social welfare ($ thousand) for two cases (Max Total Producers' Profit (TPP) Equilibrium Problem with Equilibrium Constraints EPEC).

Case	Profit				TPP	SW
	Agent 1 (E) *	Agent 1 (G) *	Agent 2 (E) *	Agent 2 (G) *		
Base	11.3	13.1	7.5	3.4	35.3	38.6
Gas shortage	13.6	12.0	3.5	0.6	29.7	32.1

* Agent 1/2 (E) and Agent 1/2 (G) represent Agent 1's/2's profits earned in the power and gas markets, respectively.

The EPEC model (19)–(20) was solved using BARON [24] under GAMS on a computer with a 2.1-GHZ Intel Core-i7 processor with 8 GB of memory. The solution time of any instance analyzed was below 190 seconds.

6. Case Study

This section examines a case study comprising the IEEE-57 bus system [25] and a tree-like 134-node Greek gas system (http://gaslib.zib.de/).

We consider (i) strategic offers/bids from both producers and consumers, (ii) disaggregated and aggregated gas price information, and (iii) diverse ownership of gas and power production units.

Taking into account the computational machinery used and for the sake of simplicity and tractability, we consider a time horizon of 3 h.

6.1. Data

The gas system consists of three gas sources, 45 demand nodes, 132 pipelines, and one gas compressor. The power system includes seven power units, being the units at buses 1, 2 and 3 gas-fired and connected to gas nodes 2, 8, and 15, respectively. This system includes 22 demand nodes and 80 transmission lines.

We consider three strategic agents, agents 1 and 2 being hybrid producers, and agent 3 a hybrid consumer. Specifically:

1. Agent 1 owns the power units at buses 1–3 and 12, and gas sources at nodes 1 and 20.
2. Agent 2 owns the power units at buses 6, 8, and 9, and the gas source at node 80.
3. Agent 3 owns electricity demands at 10 buses and gas demands at 18 nodes.

All power units and gas sources are owned by either by Agent 1 or 2 and submit strategic offers. However, a number of electricity/gas demands are not owned by Agent 3, and hence bid competitively.

6.2. From Perfect to Oligopolistic Competition

Table 3 summarizes the market equilibria obtained from the competitive model and three oligopolistic models:

1. Max SW EPEC.
2. Max TPP EPEC.
3. Max TCP EPEC.

Table 3. Case study: profits and social welfare ($ thousand) under different equilibrium models.

Equilibrium Model	Profit			TCP (Agent 3)	SW
	Agent 1	Agent 2	Total		
Competitive *	56	17	73	58	222
Oligopoly–Max SW	86	56	142	25	222
Oligopoly–Max TPP	90	57	147	19	207
Oligopoly–Max TCP	83	54	137	28	222

* In the competitive model, all agents are non-strategic and offer/bid at their marginal costs/utilities.

Figures 3 and 4 provide the load-weighted electricity and gas locational marginal prices (LMPs), respectively, obtained from the four models.

The results obtained allow the following conclusions:

1. Since no market power was exercised, the competitive model yielded the highest SW and the lowest electricity and natural gas LMPs.
2. The oligopolistic model that maximized SW resulted in the same SW as the competitive one. However, the profits of the producers (Agents 1 and 2) obtained from the oligopolistic model were nearly twice those obtained from the competitive one.
3. The oligopolistic model that maximized TPP resulted in lower SW but higher TP than the oligopolistic model that maximized SW. This is because the model that maximized TPP allowed producers further exercising market power, which yielded higher gas and power LMPs.
4. The oligopolistic model that maximized TCP yielded the highest TCP, and the same SW than the oligopolistic model that maximized SW.
5. Among the three oligopolistic models, the one that maximized TPP resulted in the highest gas and power LMPs, while the model that maximized TCP resulted in the lowest gas and power LMPs. Hence, supply-side market power increases energy prices, while the demand-side market power decreases them.

The EPEC models that maximized SW, TPP, and TCP required approximately 1681 s, 4123 s, and 3124 s, respectively, of wall-clock time to solve.

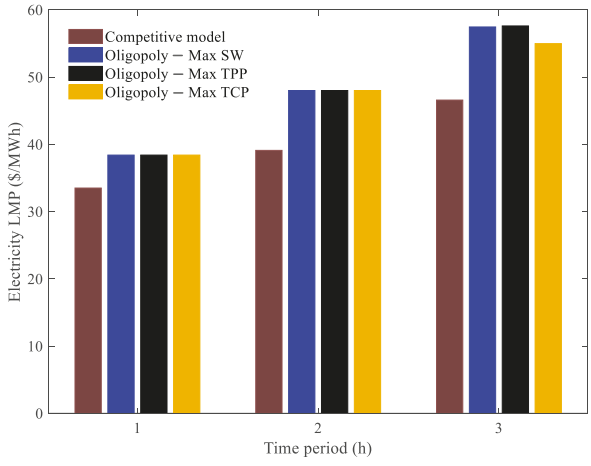


Figure 3. Case study: load-weighted electricity locational marginal prices (LMPs) obtained from four equilibrium models.

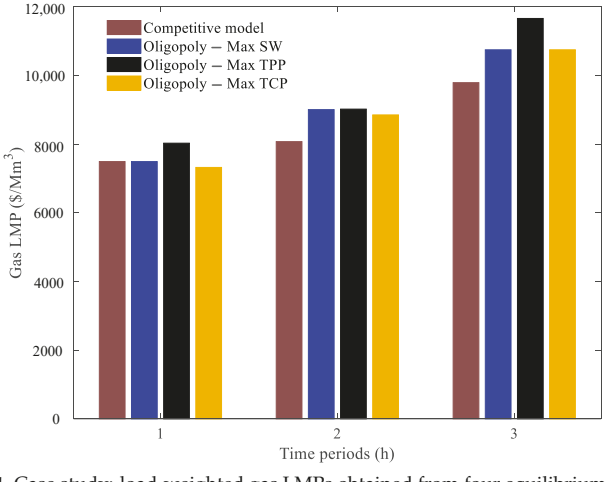


Figure 4. Case study: load-weighted gas LMPs obtained from four equilibrium models.

6.3. Aggregated Gas Prices

This subsection investigates the impact of temporal/spatial aggregation of gas prices on the market equilibria reached.

Considering the Max SW EPEC, Table 4 summarizes the market equilibria obtained from perfect pricing, spatial averaging pricing, temporal averaging pricing, and combined spatial and temporal averaging pricing.

The spatial averaging pricing derived a single price per hour by performing a load-weighting average across nodes of all gas LMP that hour (see (A32) in the Appendix). Similarly, the temporal averaging pricing derived a single price per node by performing a load-weighting average across hours of all gas LMP in that node (see (A33) in the Appendix). Finally, the combined spatial and temporal averaging pricing did both, deriving a single gas price per day (see (A34) in the Appendix).

We observe from Table 4 that the imperfect-pricing cases resulted in lower SW. Specifically, both spatial averaging pricing and temporal averaging pricing models yielded a lower TPP and a slightly higher TCP. However, the combined averaging pricing model resulted in a loss of both TPP and TCP.

Table 4. Case study: profits and social welfare ($ thousand) for a perfect pricing and three imperfect pricing cases (Max SW EPEC).

Equilibrium Model	Profit			TCP (Agent 3)	SW
	Agent 1	Agent 2	Total		
Perfect pricing	86.0	56.1	142.1	25.2	222.6
Spatial averaging	83.0	56.6	139.6	25.6	220.9
Temporal averaging	87.8	53.2	141.0	26.3	221.1
Combined averaging	80.8	56.8	137.5	25.0	220.3

These results show that highly granular pricing practices are desirable to co-ordinate gas and power markets. This is so because such practices prevent loss of SW and increased profits of gas/power producers.

6.4. Ownership Structure

We investigate in this section the impact of ownership structure on market equilibria. This was done by considering three cases involving all hybrid agents, some hybrid agents, and no hybrid agent. The Max TPP EPEC was considered. Table 5 describes the three cases considered.

Table 5. Case study: ownership structure. A: power units at buses 1–3 and 12. B: power units at buses 6, 8, and 9. C: gas sources at nodes 1 and 20. D: gas source at node 80.

Ownership	Production units Owned by			
	Agent 1	Agent 2	Agent 3	Agent 4
All hybrid agents	A and C	B and D	–	–
One hybrid agent	A	B and D	C	–
No hybrid agent	A	B	C	D

The resulting market equilibria are provided in Table 6. This table shows that the all hybrid agents' cases resulted in the highest TPP and the lowest SW. In comparison, the case of no hybrid agent resulted in the lowest TPP and the highest SW. These changes in TPP and SW are due to differences in the market power exercised by gas/power producers. In the all hybrid agents' cases, each agent accounted for a larger gas/power production capacity, and thus it could potentially exercise higher market power to its own profit, which, consequently, reduced the SW.

Table 6. Case study: profits and social welfare ($ thousand) for different market ownership cases (Max TPP EPEC).

Ownership Structure	Profit					TCP	SW
	Agent 1	Agent 2	Agent 3	Agent 4	Total		
All hybrid agents	90	57	n/a	n/a	147	19	207
One hybrid agent	30	60	55	n/a	145	21	212
No hybrid agent	26	33	58	25	142	23	214

7. Case Study 2

This section summarizes numerical results from a realistic Belgian 24-node power system and 20-node gas system [17], the topology of which is shown in Figure 5. The power units at buses 2, 3, 6, 8, 16, 15, and 22 are gas-fired and connected to nodes 4, 3, 4, 4, 6, 11, and 13, respectively. We considered three strategic producers: agents 1, 2 and 3. Agent 1 owned power units in area 1 (see upper left-hand-side of Figure 5); agent 3 owned gas sources in area A (see upper right-hand-side of Figure 5); agent 2 owned power units in area 2 (see lower left-hand-side of Figure 5) and gas sources in

area B (see lower right-hand-side of Figure 5). A fourth strategic agent owned electricity demands at buses 7, 9, 23, and 24 and gas demands at nodes 10, 12, 19, and 20. We considered a time horizon of 6 h.

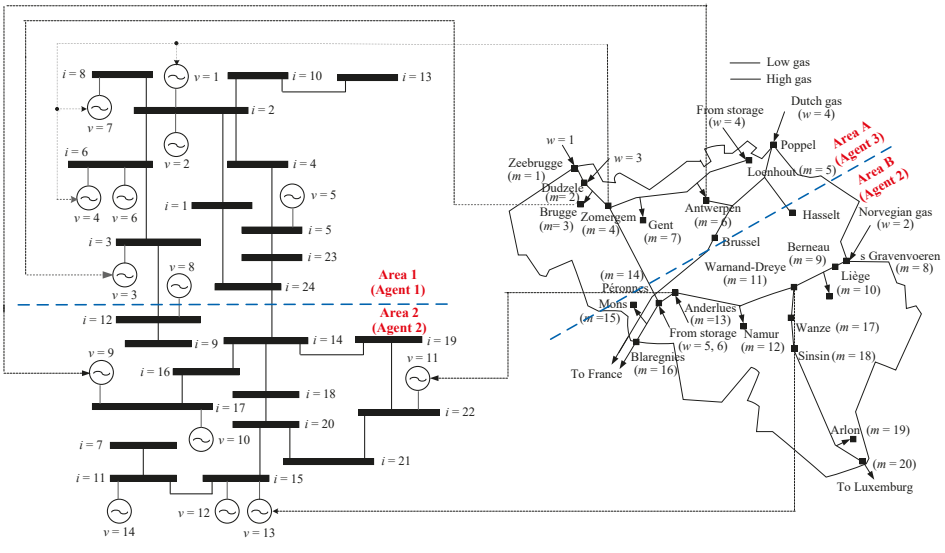


Figure 5. Case study 2: Belgian 24-node power system and 20-node gas system.

We investigated the impact of gas-pressure limits on the market equilibria reached. This was done by comparing the results obtained from two cases, in which the ranges of nodal gas pressures were between 30 bar and 70 bar and between 35 bar and 65 bar, respectively. Table 7 and Figure 6 summarize the equilibrium results obtained from the two cases. These results indicate that a strict gas-pressure limit resulted in 1) a lower TPP, TCP, and SW, 2) higher gas LMPs, and 3) lower profits of agents 1 and 2 obtained from the power market owing to increased fuel cost for gas-fired units.

Table 7. Case study 2: profits and social welfare ($ thousand) for two sets of gas pressure limits (Max TPP EPEC).

Gas Pressure	Profit				TPP	TCP	SW
Range (bar)	Agent 1	Agent 2 (E) *	Agent 2 (G) *	Agent 3		(Agent 4)	
30 – 70	612	816	255	434	2117	163	2412
35 – 65	601	797	306	401	2105	152	2385

* Agent 2 (E) and Agent 2 (G) represent Agent 2's profits earned in the power and gas markets, respectively.

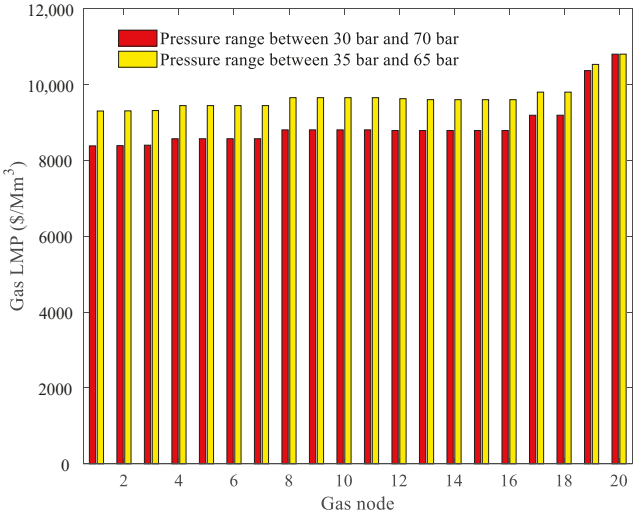


Figure 6. Case study 2: gas LMPs at the peak time period for two sets of gas pressure limits (Max TPP EPEC).

8. Conclusions

This paper proposes a multi-period EPEC model to analyze the interactions of both strategic power/gas producers and power/gas consumers that participate in power and gas markets. We investigate the impacts of (i) market power, (ii) aggregated gas prices and (iii) ownership structure of power/gas producers on the market equilibria reached. From the analysis carried out, the conclusions below are in order:

1. The proposed model is tractable and generally well-behaved, but complex. If larger instances and multi-period settings need to be considered, decomposition techniques and industry-grade computational resources are required.
2. We verify with our model that the exercise of market power results in reduced social welfare and arbitrary allocation of the extra profits among market agents. Moreover, exercising market power in either the gas or the power market impacts both the power and gas markets.
3. We find that bottlenecks in the gas system impact agents' profits earned from both gas and power markets.
4. Not transferring the true gas LMPs to the owners of gas-fired power units results in significant inefficiencies and potential intra-market and inter-market cross-subsidies.
5. We verify that the ownership structure determines the degree of market power that can be exercised by market agents: the lower the intra- and inter-market concentration, the higher the efficiency.
6. The model presented allows analyzing the impact of (i) a reduced disclosure of market outcomes (prices) and/or (ii) the impact of exercising market power by market agents. Such a model may help regulators to design market rules that encourages market-outcome disclosure, and discourages exercising market power.

Author Contributions: Conceptualization, A.J.C. and S.C.; methodology, A.J.C. and S.C.; software, S.C. and A.J.C.; validation, S.C. and A.J.C.; formal analysis, A.J.C. and S.C.; investigation, S.C. and A.J.C.; data curation: S.C. and A.J.C.; writing-original draft preparation, A.J.C. and S.C.; writing–review and editing, A.J.C. and S.C.; visualization, S.C. and A.J.C.; supervision, A.J.C. and S.C. All authors have read and agreed to the published version of the manuscript.

Funding: A. J. Conejo is partly supported by NSF Grant 1808169. S. Chen is partly supported by Fundamental Research Funds for the Central Universities under Grant 2019B05714.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Detailed Models

Appendix A.1. Gas Market Clearing

An SOC (Second Order Conic) formulation of the gas operation problem (3)–(5) is:

$$\max_{\Xi_G^P} \sum_{e\in\Lambda^{GO},t\in T} C_{et}^{GL}F_{et}^{L} + \sum_{l\in\mathbb{L},e\in\Lambda_l^{GL},t\in T} \varepsilon_{et}F_{et}^{L} + \sum_{m\in\mathbb{N},t\in T}\left(\sum_{v\in\Psi_m^G}\gamma_{vt}^{G}F_{vt}^{G} - \sum_{w\in\Psi_m^S}\beta_{wt}F_{wt}^{S}\right) \tag{A1}$$

$$\text{s.t.} \sum_{w\in\Psi_m^S}F_{wt}^{S} = \sum_{k\in\mathbb{C}(m)}(1+\vartheta_k)F_{kt}^{C} + \sum_{e\in\Psi_m^L}F_{et}^{L} + \sum_{v\in\Psi_m^G}F_{vt}^{G} + \sum_{n\in\mathbb{G}(m)}F_{mnt}; \forall m\in\mathbb{N}, t\in T \quad (u_{mt}) \tag{A2}$$

$$\bar{F}_{mnt} = (F_{mnt} - F_{nmt})/2; \forall m,n\in\mathbb{N}, t\in T \tag{A3}$$

$$F_{mnt} + F_{nmt} = L_{mnt} - L_{mn,t-1}; \forall m,n\in\mathbb{N}; t\in T \tag{A4}$$

$$L_{m,n,t} = K_{m,n}\cdot(\pi_{m,t}+\pi_{n,t})/2; \forall m,n\in\mathbb{N}; t\in T \tag{A5}$$

$$(\bar{F}_{mnt}/W_{mn})^2 \le \Pi_{mt}^2 - \Pi_{nt}^2; \forall m\in\mathbb{N}, n\in\mathbb{G}(m), t\in T \tag{A6}$$

$$\Pi_{kt}^{in}\rho_{C,k}^{\min} \le \Pi_{kt}^{out} \le \Pi_{kt}^{in}\rho_{C,k}^{\max}; \forall k\in\mathcal{K}, t\in T \tag{A7}$$

$$0 \le F_{kt}^{C} \le F_{k}^{C,\max}; \forall k\in\mathcal{K}, t\in T \tag{A8}$$

$$0 \le F_{wt}^{S} \le F_{w}^{S,\max}; \forall m\in\mathbb{N}, w\in\Psi_m^S, t\in T \tag{A9}$$

$$-F_{w}^{S,ramp} \le F_{wt}^{S} - F_{w,t-1}^{S} \le F_{w}^{S,ramp}; \forall m\in\mathbb{N}, w\in\Psi_m^S, t\in T \tag{A10}$$

$$0 \le F_{et}^{L} \le F_{et}^{L,\max}; \forall e\in\Lambda^G, t\in T \tag{A11}$$

$$\Pi_m^{\min} \le \Pi_{mt} \le \Pi_m^{\max}; \forall m\in\mathbb{N}, t\in T \tag{A12}$$

$$F_{mnt} \ge 0; \forall m\in\mathbb{N}, n\in\mathbb{G}(m), t\in T \tag{A13}$$

$$0 \le F_{vt}^{G} \le F_{v}^{G,\max}; \forall m\in\mathbb{N}, v\in\Psi_m^G, t\in T, \tag{A14}$$

where e is the index of gas demands in set Λ^{GO}, t the index of operating periods in set T, l the set of agents in set $\mathbb{L}$, m and n the indices of nodes in set $\mathbb{N}$, v the index of power units, Ψ_m^G the set of gas-fired units connected to node m, Ψ_m^S the set of gas sources connected to node m, $\mathbb{C}(m)$ the set of gas compressors connected to node m, Ψ_m^L the set of gas demands connected to node m, $\mathbb{G}(m)$ the set of nodes that are connected directly to node m, k the index of gas compressors in the set $\mathcal{K}$, and w the index of gas sources.

The parameters of the problem (A1)–(A14) are described below. C_{et}^{GL} is the marginal utility of demand e at time period t, $F_{k}^{C,max}$ the gas transportation capacity of compressor k, $F_{w}^{S,max}$ the production capacity of gas source w, $F_{et}^{L,\max}$ the quantity of gas demand e at time period t, $F_{v}^{G,\max}$ the maximum gas consumption of power unit v, K_{mn} the line-pack parameter of the pipeline connecting nodes m and n, W_{mn} the Weymouth constant of the pipeline connecting nodes m and n, $\rho_{C,k}^{\min}$ and $\rho_{C,k}^{\max}$ the minimum and maximum compression ratio of compressor k, ϑ_k the conversion efficiency of gas compressor k, and $\Pi_m^{\min}$ and $\Pi_m^{\max}$ the minimum and maximum gas pressures of node m, respectively.

The variables of the problem (A1)–(A14) are as follows. F_{et}^{L} is the consumption of demand e in time period t, F_{wt}^{S} the production of gas source w in time period t, F_{kt}^{C} the gas flow through compressor k in time period t, F_{mnt} the gas flow through the pipeline connecting nodes m and n in time period t, $\bar{F}_{mnt}$ the average gas flow through the pipeline connecting nodes m and n in time period t, L_{mnt} the line-pack in pipeline connecting nodes m and n in time period t, ε_{et} the bid of demand e in time period t, γ_{vt}^{G} the bid of gas-fired power unit v in time period t, and β_{wt} the offer of gas source w in time period t.

It should be noted that the variables ε_{et}, γ^G_{vt}, and β_{wt} are fixed by upper-level problems and thus are constant for this problem. u_{mt} denotes the dual variable of (A2), and represents the gas LMP of node m at time period t. The variables of problem (A1)–(A14) are those in set $\Xi^{\mathrm{P}}_{\mathrm{G}} = \left\{F^L_{et}, F^S_{wt}, F^C_{kt}, F_{mnt}, \bar{F}_{mnt}, L_{mnt}\right\}$.

The objective function (A1) is the gas SW that incorporates strategic offers from gas producers and strategic bids from power producers and gas consumers. Constraints (A2) represent the gas nodal balances. Constraints (A3) calculate average gas flows through pipelines. Constraints (A4) give the relationship between hourly changes in flows and line-pack in pipelines. Constraints (A5) determine the hourly line-pack in each pipeline, which is considered to be linear with the average gas pressure at the two ends of the pipeline. Constraints (A6) relate the average gas flow with the change in squared gas pressures between the upstream and downstream nodes for each pipeline. (A6) represent an SOC formulation of an exact gas flow model [3]. Constraints (A7) enforce minimum and maximum gas pressure ratios of gas compressors. Constraints (A8) impose transportation limits on gas compressors. Constraints (A9) and (A10) impose production capacities and ramping limits on gas sources, respectively. Constraints (A11) limit the amount of gas demands served. Constraints (A12) enforce minimum and maximum gas pressures of each node. Constraints (A13) assume that the direction of gas flows are known a priori, which is generally reasonable in short-term operations [3]. Constraints (A14) limit the amount of generation-related demands.

Appendix A.2. Power Market Clearing

An LP (Linear Programming) formulation of the power operation problem (6)–(8) is:

$$\max_{\Xi^{\mathrm{P}}_{\mathrm{E}}} \sum_{l\in\mathbb{L}, d\in\Lambda^{EL}_l, t\in T} \varsigma_{dt} P^L_{dt} + \sum_{d\in\Lambda^{EO}, t\in T} C^{EL}_{dt} P^L_{dt} - \sum_{v\in\Omega^E, t\in T} \alpha_{vt} P^G_{vt} \tag{A15}$$

$$\text{s.t.} \sum_{d\in\Theta^D_i} P^L_{dt} + \sum_{j\in\mathbb{E}(i)} b_{ij}\cdot(\delta_{it}-\delta_{jt}) = \sum_{v\in\Theta^G_i} P^G_{vt}; \forall i\in\mathbb{B}, t\in T \quad (\lambda_{it}) \tag{A16}$$

$$0 \le P^L_{dt} \le P^{L,\max}_{dt}; \forall d\in\Lambda^E, t\in T \tag{A17}$$

$$b_{ij}\cdot(\delta_{it}-\delta_{jt}) \le P^{\max}_{ij}; \forall i\in\mathbb{B}, j\in\mathbb{E}(i), t\in T \tag{A18}$$

$$0 \le P^G_{vt} \le P^{G,\max}_{v}; \forall v\in\Omega^E, t\in T \tag{A19}$$

$$P^{G,ramp}_{v} \le P^G_{vt} - P^G_{v,t-1} \le P^{G,ramp}_{v}; \forall v\in\Omega^E, t\in T \tag{A20}$$

$$\delta_{REF,t} = 0; \forall t\in T, \tag{A21}$$

where d is the index of electricity demands in set Λ^E, i and j the indices of electric buses in set $\mathbb{B}$, v the index of power units in set Ω^E, REF the index of the reference bus. Λ^{EL}_l the set of strategic consumers owned by agent l, Λ^{EO} the set of non-strategic consumers, Θ^D_i the set of electricity demands directly connected to bus i, Θ^G_i the set of power units directly connected to bus i, and $\mathbb{E}(i)$ the set of buses directly connected to bus i.

The parameters of the problem (A15)–(A21) are described below. C^{EL}_{dt} is the marginal utility of demand d in time period t, b_{ij} the susceptance of the line connecting buses i and j, $P^{L,\max}_{dt}$ the quantity of demand d in time period t, $P^{\max}_{ij}$ the transmission capacity of the line connecting buses i and j, $P^{G,\max}_{v}$ and $P^{G,ramp}_{v}$ the capacity and ramping limit of power unit v, respectively.

The variables of the problem (A15)–(A21) are as follows. P^L_{dt} is the quantity of demand d served in time period t, P^G_{vt} the power production of unit v in time period t, δ_{it} the phase angle of bus i in time period t, ς_{dt} the bid of demand d in time period t, and α_{vt} the offer of power unit v in time period t.

It should be noted that variables ς_{dt} and α_{vt} are determined by upper-level problems and thus are constants for this problem. λ_{it} denotes the dual variable of (A16), and represents the electricity LMP of bus i in time period t. The variables of problem (A15)–(A21) are those in the set $\Xi^{\mathrm{P}}_{\mathrm{E}} = \left\{P^L_{dt}, \delta_{it}, P^G_{vt}\right\}$.

The objective function (A15) is the power SW that considers the strategic offers of power producers and the strategic bids of power consumers. Constraints (A16) represent the active power balances, in which the DC power flow model is used. Constraints (A17) impose upper limits on power demands. Constraints (A18) enforce the transmission capacity of each line. Constraints (A19) and (A20) impose production capacities and ramping limits on power units, respectively. Constraints (A21) fix the phase angle of the reference bus to zero.

Appendix A.3. Agent Model

Appendix A.3.1. Strategic Consumer

The problem of a gas/power strategic consumer is:

$$\max_{\Xi_{\mathrm{UC}}} \sum_{d\in\Lambda_l^{EL},t\in T}(C_{dt}^{EL}-\lambda_{i(d),t})P_{dt}^{L}+\sum_{e\in\Lambda_l^{GL},t\in T}(C_{et}^{GL}-u_{m(e),t})F_{et}^{L} \tag{A22}$$

$$\text{s.t. } \varepsilon_{et}\geq 0;\forall e\in\Lambda_l^{GL},t\in T \tag{A23}$$

$$\varsigma_{dt}\geq 0;\forall d\in\Lambda_l^{EL},t\in T \tag{A24}$$

$$(A1)-(A21), \tag{A25}$$

where $i(d)$ is the bus at which electricity demand d is located, $m(e)$ the node at which gas demand e is located, and $\Xi_{\mathrm{UC}}=\left\{\Xi_{\mathrm{G}}^{\mathrm{P}},\Xi_{\mathrm{G}}^{\mathrm{E}},\varepsilon_{et},\varsigma_{dt}\right\}$.

The objective function (A22) is the profit of consumer l. Constraints (A23) and (A24) represent the non-negative bids of strategic gas consumer e and electricity consumer d, respectively. Constraints (A25) enforce market constraints.

Appendix A.3.2. Strategic Producer

The problem of a gas/power strategic producer is:

$$\max_{\Xi_{\mathrm{UP}}} \sum_{v\in\Omega_l^{C}\cup\Omega_l^{R},t\in T}\lambda_{i(v),t}P_{vt}^{G}-\sum_{v\in\Omega_l^{R},t\in T}C_{v}^{G}P_{vt}^{G}-\sum_{v\in\Omega_l^{G},t\in T}(C_{v}^{\mathrm{O}}+\eta_v u_{vt}^{GE})P_{vt}^{G}$$
$$+\sum_{w\in\Omega_l^{S},t\in T}(C_{w}^{S}-u_{m(w),t})F_{wt}^{S} \tag{A26}$$

$$\text{s.t. } \alpha_{vt}\geq 0;\forall v\in\Omega_l^{C}\cup\Omega_l^{G},t\in T \tag{A27}$$

$$\beta_{wt}\geq 0;\forall w\in\Omega_l^{S},t\in T \tag{A28}$$

$$\gamma_{vt}^{G}\geq 0;\forall v\in\Omega_l^{G},t\in T \tag{A29}$$

$$(A1)-(A21). \tag{A30}$$

where $i(v)$ is the bus at which power unit v is located, u_{vt}^{GE} the gas price information that the gas system sends to power unit v, and $m(w)$ the node at which gas source w is located, and $\Xi_{\mathrm{UP}}=\left\{\Xi_{\mathrm{G}}^{\mathrm{P}},\Xi_{\mathrm{G}}^{\mathrm{E}},\alpha_{vt},\beta_{wt},\gamma_{vt}^{G}\right\}$.

The objective function (A26) is the profit of each producer l. Constraints (A27) and (A28) represent non-negative offers of power producer v and gas producer w, respectively. Constraints (A29) represent non-negative bids of gas-fired power unit v in the gas market. Constraints (A30) enforce market constraints.

Appendix A.4. Perfect and Imperfect Gas Price Disclosure

We consider perfect and imperfect price coordination between gas and power markets. Specifically, we consider different levels of gas price granularity. In the perfect-pricing case, exact gas price information is exchanged as:

$$u_{vt}^{GE} = u_{m(v),t};\ \ \forall l \in \mathbb{L}, v \in \Omega_l^G, t \in T. \tag{A31}$$

where $m(v)$ denotes the node at which gas-fired power unit v is located.

In the imperfect-pricing cases, we consider spatial, temporal, and combined averaging of gas LMPs, in which the information interchange on gas prices are given by (A32), (A33), and (A34), respectively.

$$u_{vt}^{IN} = \sum_{e\in\Lambda^G} F_{et}^L u_{m(e),t} / \sum_{e\in\Lambda^G} F_{et}^L;\ \ \forall l \in \mathbb{L}, v \in \Omega_l^G, t \in T \tag{A32}$$

$$u_{vt}^{IN} = \frac{1}{|T|} \sum_{t\in T} u_{m(v),t};\ \ \forall l \in \mathbb{L}, v \in \Omega_l^G, t \in T \tag{A33}$$

$$u_{vt}^{IN} = \frac{1}{|T|} \sum_{t\in T} \left(\sum_{e\in\Lambda^G} F_{et}^L u_{m(e),t} / \sum_{e\in\Lambda^G} F_{et}^L \right);\ \ \forall l \in \mathbb{L}, v \in \Omega_l^G, t \in T. \tag{A34}$$

Additionally note that the true SW is given by:

$$\sum_{d\in\Lambda^E, t\in T} C_{dt}^{EL} P_{dt}^L - \sum_{v\in\Omega^E, t\in T} C_{vt}^G P_{vt}^G + \sum_{e\in\Lambda^{GO}, t\in T} C_{et}^{GL} F_{et}^L - \sum_{m\in\mathbb{N}, w\in\Psi_m^S, t\in T} C_{wt}^S F_{wt}^S. \tag{A35}$$

References

1. Gil, J.; Caballero, A.; Conejo, A.J. Power Cycling: CCGTs: The Critical Link Between the Electricity and Natural Gas Markets. *IEEE Power Energy Mag.* **2014**, *12*, 40–48. [CrossRef]
2. Chen, S.; Wei, Z.; Sun, G.; Cheung, K.W.; Sun, Y. Multi-linear probabilistic energy flow analysis of integrated electrical and natural-gas systems. *IEEE Trans. Power Syst.* **2017**, *32*, 1970–1979. [CrossRef]
3. Chen, S.; Conejo, A.J.; Sioshansi, R.; Wei, Z. Unit Commitment with an Enhanced Natural Gas-Flow Model. *IEEE Trans. Power Syst.* **2019**, *34*, 3729–3738. [CrossRef]
4. Byeon, G.; van Hentenryck, P. Unit Commitment with Gas Network Awareness. *IEEE Trans. Power Syst.* **2020**. [CrossRef]
5. He, C.; Zhang, X.; Liu, T.; Wu, L. Distributionally Robust Scheduling of Integrated Gas-Electricity Systems with Demand Response. *IEEE Trans. Power Syst.* **2019**, *34*, 3791–3803. [CrossRef]
6. He, Y.; Yan, M.; Shahidehpour, M.; Li, Z.; Guo, C.; Wu, L.; Ding, Y. Decentralized Optimization of Multi-Area Electricity-Natural Gas Flows Based on Cone Reformulation. *IEEE Trans. Power Syst.* **2018**, *33*, 4531–4542. [CrossRef]
7. Chen, R.; Wang, J.; Sun, H. Clearing and Pricing for Coordinated Gas and Electricity Day-Ahead Markets Considering Wind Power Uncertainty. *IEEE Trans. Power Syst.* **2018**, *33*, 2496–2508. [CrossRef]
8. Ameli, H.; Qadrdan, M.; Strbac, G. Value of gas network infrastructure flexibility in supporting cost effective operation of power systems. *Appl. Energy* **2017**, *202*, 571–580. [CrossRef]
9. Yang, J.; Zhang, N.; Kang, C.; Xia, Q. Effect of Natural Gas Flow Dynamics in Robust Generation Scheduling Under Wind Uncertainty. *IEEE Trans. Power Syst.* **2018**, *33*, 2087–2097. [CrossRef]
10. Bai, L.; Li, F.; Jiang, T.; Jia, H. Robust Scheduling for Wind Integrated Energy Systems Considering Gas Pipeline and Power Transmission N–1 Contingencies. *IEEE Trans. Power Syst.* **2017**, *32*, 1582–1584. [CrossRef]
11. Zlotnik, A.; Roald, L.; Backhaus, S.; Chertkov, M.; Andersson, G. Coordinated Scheduling for Interdependent Electric Power and Natural Gas Infrastructures. *IEEE Trans. Power Syst.* **2017**, *32*, 600–610. [CrossRef]
12. Chen, S.; Wei, Z.; Sun, G.; Cheung, K.W.; Wang, D.; Zang, H. Adaptive Robust Day-Ahead Dispatch for Urban Energy Systems. *IEEE Trans. Ind. Electron.* **2019**, *66*, 1379–1390. [CrossRef]

13. Antenucci, A.; Sansavini, G. Gas-constrained secure reserve allocation with large renewable penetration. *IEEE Trans. Sustain. Energy* **2018**, *9*, 685–694. [CrossRef]
14. Ordoudis, C.; Pinson, P.; Morales, J.M. An Integrated Market for Electricity and Natural Gas Systems with Stochastic Power Producers. *Eur. J. Oper. Res.* **2019**, *272*, 642–654. [CrossRef]
15. Wang, C.; Wei, W.; Wang, J.; Wu, L.; Liang, Y. Equilibrium of Interdependent Gas and Electricity Markets with Marginal Price Based Bilateral Energy Trading. *IEEE Trans. Power Syst.* **2018**, *33*, 4854–4867. [CrossRef]
16. Wang, C.; Wei, W.; Wang, J.; Liu, F.; Mei, S. Strategic Offering and Equilibrium in Coupled Gas and Electricity Markets. *IEEE Trans. Power Syst.* **2018**, *33*, 290–306. [CrossRef]
17. Chen, S.; Conejo, A.J.; Sioshansi, R.; Wei, Z. Equilibria in Electricity and Natural Gas Markets with Strategic Offers and Bids. *IEEE Trans. Power Syst.* **2020**. [CrossRef]
18. Ruiz, C.; Conejo, A.J. Pool Strategy of a Producer with Endogenous Formation of Locational Marginal Prices. *IEEE Trans. Power Syst.* **2009**, *24*, 1855–1866. [CrossRef]
19. Ruiz, C.; Conejo, A.J.; Smeers, Y. Equilibria in an Oligopolistic Electricity Pool with Stepwise Offer Curves. *IEEE Trans. Power Syst.* **2012**, *27*, 752–761. [CrossRef]
20. Chen, S.; Conejo, A.J.; Sioshansi, R.; Wei, Z. Operational Equilibria of Electric and Natural Gas Systems with Limited Information Interchange. *IEEE Trans. Power Syst.* **2020**, *35*, 662–671. [CrossRef]
21. Ben-Tal, A.; Nemirovski, A. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*; Society for Industrial and Applied Mathematics (SIAM): Philadelphia, PA, USA, 2001.
22. Gabriel, S.A.; Conejo, A.J.; Fuller, J.D.; Hobbs, B.F.; Ruiz, C. *Complementarity Modeling in Energy Markets*; Springer: New York, NY, USA, 2013.
23. Hu, X.; Ralph, D. Using EPECs to Model Bilevel Games in Restructured Electricity Markets with Locational Prices. *Oper. Res.* **2007**, *55*, 809–827. [CrossRef]
24. Tawarmalani, M.; Sahinidis, N.V. A polyhedral branch-and-cut approach to global optimization. *Math. Program.* **2005**, *103*, 225–249. [CrossRef]
25. Zimmerman, R.D.; Murillo-Sánchez, C.E.; Thomas, R.J. MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education. *IEEE Trans. Power Syst.* **2011**, *26*, 12–19. [CrossRef]

energies

MDPI

Article

Modeling the Price Stability and Predictability of Post Liberalized Gas Markets Using the Theory of Information

Anis Hoayek [1,†], Hassan Hamie [2,*,†] and Hans Auer [2]

[1] Faculty of Science, Institute Alexander Grothendieck, University of Montpellier, 34090 Montpellier, France; anisshoayek@hotmail.com

[2] Faculty of Electrical Engineering and Information Technology, Institute of Energy Systems and Electrical Drives, Energy Economics Group (EEG), Vienna University of Technology, 1040 Vienna, Austria; auer@eeg.tuwien.ac.at

* Correspondence: hassan.hamieh@student.tuwien.ac.at; Tel.: +43-9613824622

† These authors contributed equally to this work.

Received: 29 April 2020; Accepted: 3 June 2020; Published: 11 June 2020

Abstract: Energy markets in the United States and Europe are getting more liberalized. The question of whether the liberalization of the gas industry in both markets has led to stable prices and less concentrated markets has appealed great interest among the scientific community. This study aims to measure the power and efficiency of an information structure contained in the gas prices time series. This assessment is useful to the oversight duty of regulators in such markets in the post liberalized era. First, econometric and mathematical methods based on game theory, records theory, and Shannon entropy are used to measure the following indicators: level of competition, price stability, and price uncertainty respectively—for both markets. Second, the level of information generated by these indicators is quantified using the information theory. The results of this innovative two-step approach show that the functioning of the European market requires the regulator's intervention. This intervention is done by applying additional rules to enhance the competitive aspect of the market. This is not that case for the U.S. market. Also, the value of the information contained in both markets' wholesale gas prices, although in asymmetric terms, is significant, and therefore proves to be an important instrument for the regulators.

Keywords: gas markets; game theory-Cournot model; records theory; entropy; information theory

1. Introduction

Unlike the oil markets, the gas markets have witnessed regional divergence at several levels. However, the degree of competitiveness varies between the different gas markets.

Following extensive infrastructure development and regulation changes, the North American market developed transparent and competitive gas pricing hubs. Additional gas hubs emerged afterward in Europe, providing physical and virtual locations for trading gas. The abundance of gas and the presence of competition between different stakeholders at different levels of the value chain led to an increase in trade in both the spot and future markets. However, the price of gas does not reflect market fundamentals and forces in all markets.

The role of the regulators is to promote competitive conduct, domestic gas production, third-party access, price trade reporting, and to ensure the presence of futures trading. Once the liberalization measures are implemented and regulated, the status of the gas hub is confirmed as liquid and stable, which results in prices being indicative of market fundamentals.

In this study, we focus on the North American and European markets, in specific the United Kingdom (UK). The choice of these markets is explained by the fact that both attempted to liberalize the gas markets, and underwent intense regulations and policy changes over the past years [1].

Wholesale buyers used to follow long-term contracts indexed to the price of oil derivatives in both of the aforementioned markets [2]. Also, the gas industry was mostly dominated by state-owned monopolies.

However, the Federal Energy Regulatory Commission (FERC), encouraged the establishment of gas markets driven by free competition in the United States [3]. As a result, the Henry Hub (HH), known as the most successful hub, was created [4]. The success of the HH is marked by a large liquid portfolio of spot and future contracts, along with hub indexed prices which serve as a reference for the value of the gas commodity all over the North America.

Consequently, the UK and the European Union (EU) started their reforms. The Office of Gas and Electricity Markets (OFGEM) took lead and started the process of market liberalization since the 1990s. The reforms led to the establishment of the National Balancing Point (NBP), which serves as a physical platform for gas trading in the UK. Currently, the NBP is considered to be the most developed hub in Europe and has the longest standing European gas pricing point [5,6]. It is worth mentioning that the UK gas market and the European gas market are used interchangeably in the remainder of this study.

In 2016, the U.S. Natural gas consumption was roughly 750 billion cubic meters (BCM) [7]. The majority of the demand was satisfied through indigenous production, and the remaining was imported from Canada, via pipelines. Additional marginal gas is imported from Mexico (also via pipelines), and from around the world, via liquefied natural gas, LNG (see Figure 1).

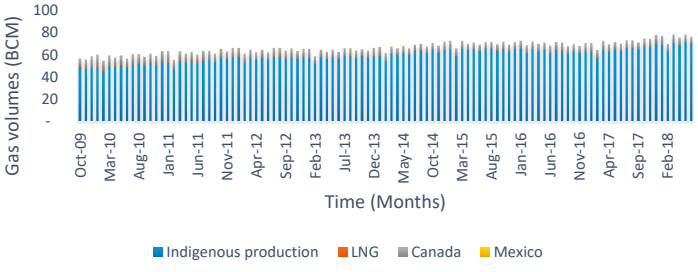


Figure 1. Indigenous production and monthly gas imports to U.S. Source: EIA (Available on EIA website, https://www.eia.gov/naturalgas/monthly/).

The UK natural gas consumption in 2016 is estimated at around 73 BCM [8], out of which 42 BCM are imported while the remaining volumes are produced locally (see Figure 2).

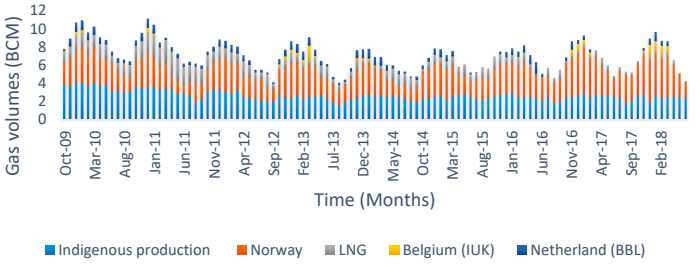


Figure 2. Monthly gas imports to the UK and indigenous production. Source: OFGEM –UK (Available on OFGEM website, https://www.ofgem.gov.uk/data-portal/gas-demand-and-supply-source-month-gb).

To reflect the gradual advances in supply-side competition, the functioning of a wholesale gas market should be measured quantitatively. This has attracted attention from the professional and scientific community, as in-depth analysis of gas markets have been conducted and published [2,9–13]. All studies confirm that parameters such as market participants, the monthly day ahead trades, and churn ratio give an indication and a feel of the market. The churn ratio, calculated as the ratio of traded gas volumes to the total gas demanded, is an indicative measure of the liquidity of a gas hub and market maturity. Additionally, it measures the confidence of traders and consumers in the market.

The numbers shown in Table 1, reveal a high churn ratio for both markets (above 15), which indicates that the gas prices registered at both hubs are liquid and reflect market conditions [11]. Therefore, clearing prices are accepted as a reference and indicator, which contain reliable information for all stakeholders involved in the gas value chain (traders, customers, regulators, etc.).

Table 1. The United States and UK traded volumes and churn. Source: OFGEM [14] and Cornerstone Research, IEA [15].

	2015–2016			
Hub	**Market Participants**	**Traded Volumes, Physical (BCM)**	**Traded Volumes, FUTURES (BMC)**	**Churn Ratio (Total)**
HH/U.S.	N/A	1776	53,968	67
NBP/UK	40	901	925	20

Analysis of recent trends in the European and North American gas markets shows that the prices of gas are fundamentally market-driven. However, rules and policies set by gas regulators are a must to guarantee that the market keeps on operating efficiently [16].

The information theory introduced by [17] is a probabilistic principle that helps to quantify the information generated by a random variable in an uncertain context. Information theory can explain observations without the need to rely on neither statistical assumptions regarding the distribution of random variables, nor the random noise [18]. Additional parametric assumptions such as estimates of demand and cost functions can also be avoided using the information theory. The mathematical tool that will be used in this work to measure the amount of information from the gas wholesale clearing prices is the statistical entropy.

Another tool to assess the information in a decision-making problem is the Blackwell approach [19]. However, this approach has several complexities that prevent a simple application of Blackwell's principle [20,21], especially at the level of cost and return function assumptions. Hence, to overcome all the mentioned difficulties, the entropy principle will be applied.

The first objective of this paper is to study whether the wholesale gas prices of two of the most liberalized gas markets carry valuable information that can serve as indicators for the relevant gas regulators. The value of these indicators will be quantified by using several econometric methods and mathematical theories. This analysis will guide and assist the decision-making process of regulators regarding the need for an intervention to stabilize the gas markets and improve the functioning of their internal markets. The second objective is to quantify and measure the accuracy and efficiency of the hidden information structure generated by these indicators.

All methods applied in this research are proven mathematical theories that have been used in previous studies [22–34]. All four theories (information, records, entropy, and game theory) have applications in the field of mathematics and statistics, in other words, econometrics. However, the novelty of our method lies in adopting a two-step approach that was not applied before in the literature. This approach is useful to assess the performance of a gas market in terms of information generated by several indicators and combined in one structure, called information structure. The indicators give an idea of the level of competition, level of price volatility, and price stability on one hand, and the level of the information structure, which measures the performance of the market. Among all gas stakeholders, this information is important for gas regulators. These will

be more confident and can trust the price indicators if the performance of the market is powerful and efficient.

The level of competition changes from one market to another, and if measured correctly defines the concentration of competing firms in the market. Limited number of firms imply a highly concentrated market, and that based on their strategies can dictate prices, otherwise known as price-setters. Besides, the fewer the number of firms the easier it is to abuse conduct and act collusively. Such firms adjust their strategies in conjunction with an agreed-upon understanding with the competing firms at the expense of the welfare of gas consumers and possibly smaller firms. A typical example of such a market and behavior is the presence of cartels in commodity markets.

The level of volatility indicates how fast gas prices change in the short term. The higher the volatility the harder it is to predict the future behavior of the changes, thus making the market uncertain.

On the other hand, price stability hints at the behavior of gas prices in the medium and long term. Commodity prices tend to have abrupt and rapid price shocks, this explains the sudden increase or decrease due to sudden changes in supply and demand characteristics. The longer it takes for a commodity price to witness a shock the more stable the market is.

The performance of the market is the measure of the power and efficiency of the information contained in the gas prices and the indicators. The more the information is efficient, the more reliable, and reflective the prices are in such a market.

In the first step of the approach, the authors identified three different indicators: market concentration, price stability, and price uncertainty. The authors then applied three appropriate mathematical and statistical theories to extract, from the gas prices time-series metric values that are most suitable to measure the relevant indicators. The formulation of each model is explained and justified in the next section.

In the second step, the three indicators are combined to create an information structure that will help the authors evaluate the performance of the gas market in question. These are assessed against the actions/states that could be executed by the regulator of such markets. Two actions are identified: to intervene or not to intervene. Intervention in the market is conducted by taking legal actions, such as issuing new directives to ensure a stable supply and demand equilibrium, and making sure that there is no abusive conduct by gas suppliers. Furthermore, the approach deals with a case where the information is neither completely absent nor perfectly known, which has rarely been dealt with in literature.

2. Methods/Data and Models Formulation

To avoid price abuse and manipulation, a gas regulator is expected to regulate firms' behavior by ensuring that customer welfare is maximized while maintaining the attractiveness and profitability for the producers and traders.

As stated in the introduction, price dynamics of a commodity in a liberalized market are indicative of the market structure. They contain consistent information that should, if adequately analyzed, help the regulators in assessing the performance of the market, namely the wholesale gas market in this case.

The authors have identified three main metrics that can signal information in the hidden structure of the price values of both hubs. These metrics are based on econometric and mathematical methods, and are used to inform the regulator in each market about the following:

- Indicator 1: Level of competition;
- Indicator 2: Market stability;
- Indicator 3: Volatility and uncertainty of prices.

The first indicator studies the degree of concentration in the two different gas markets by using game theory, specifically the non-parametric Nash-Cournot equilibrium test. In other words, if the test shows that traders are participating in the market by trying to maximize their profit as "the only

pure" strategy, then the market is considered efficient and the likelihood of anti-competitive behavior is negligible.

The second indicator employs the records theory, which relies on the analysis of the peak observations reached in a certain period. This indicator measures the degree of market stability, by calculating the probability of witnessing future peak prices. Therefore, the measure of probability is a measure of market stability and predictability. If the results point toward a tendency to score high probabilities of extreme gas prices, then the market can be characterized as unstable.

The third and final indicator studies the price predictability of both markets, by the use of Shannon entropy and the measure of volatility. This is done by analyzing the variation of prices and returns and assessing the degree of uncertainty and volatility which are present in gas prices. Simply, the higher the uncertainty in prices, the higher the volatility.

These indicators combined will inform the regulator about the functioning of the market. If the market shows signs of concentration, the likelihood of extreme prices, volatility, and uncertainty, then the regulator should intervene and use its policy enforcement power.

Since our indicators are based on econometric theory and models, it is important to assess the performance of such models. Therefore, a quantitative analysis that relies on the information theory is used to compare the power and efficiency of the information generated by all indicators in the two different selected markets. The market with the highest information power will give additional credibility to the indicators so that the gas prices time series speak for itself. Regulators in such a market have higher confidence and can trust the indicators, which will guide their decision of whether to intervene in the market or not.

The following part of Section 2 will list and define the three indicators metrics and will explain the econometric and mathematical methods that will be used in this manuscript. Section 2.5 will lay out an outline of the power of information. Results will be presented in Section 3 with an analysis of their significance and impact, with an overview of how they can be used by gas regulators in their assessment of wholesale gas markets. The final section will conclude the study and emphasize the importance of the dynamics of gas prices and the power of information that it provides.

2.1. Data

A description of the data used in the study is presented in Table 2. The variables set consist of monthly wholesale gas prices that were registered between October 2009 and June 2018.

Table 2. Data collection

Variable	Frequency	Number of Observations	Source
The National Balancing point, NBP gas prices (2009–2018)	Monthly	105	National Regulatory Authority—OFGEM Office of Gas and Electricity Markets—UK Government [a]
The Henry Hub, HH gas prices (2009–2018)	Monthly	105	U.S. Energy Information Administration (EIA) [b]

[a] Available on OFGEM website, https://www.ofgem.gov.uk/data-portal/gas-prices-day-ahead-contracts-monthly-average-gb; [b] Available on the U.S. Energy Information Administration (EIA) website, https://www.eia.gov/dnav/ng/hist_xls/RNGWHHDd.xls.

Figure 3 illustrates, in a time series plot, the evolution of the natural gas prices for the two different markets. There is a clear divergence that happened in the year 2009 and continues to date. This is mainly due to two main factors that took place in the United States. The first is the abundance and oversupply of new unconventional shale gas production in the local market. The second is that U.S. natural gas contracts in that period started to be decoupled from the crude oil prices.

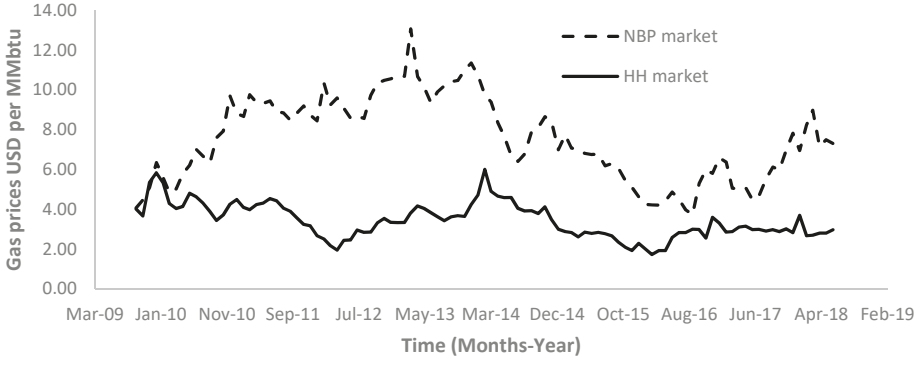


Figure 3. Monthly gas prices of both gas markets, USD/MMBtu.

The line plots of the two markets presented in Figure 3 show that there is no clear indication of a linear relationship between both variables.

2.2. Indicator 1—Level of Competition

The level of competition and market concentration method involves the classical Nash-Cournot equilibrium test. A Cournot equilibrium is reached when a given firm maximizes its profit by changing its output taking into account the other firm's output. One important feature of a Cournot model is that firms are not allowed to cooperate. Therefore, as long as the players are playing the latter strategy, the companies would be abiding by pure market profit-oriented strategies, trying to maximize their "utility function," and have no agreed-upon behavior (i.e., collusion). The market, where producers follow this trend is considered more liberalized.

The aim is to test whether the behavior of the gas producers in the respective markets follows a Cournot model. If a set of gas producers are not following the assumptions of a Cournot game, the test will identify them.

The optimal quantities of the suppliers in a given market are obtained by numerically solving the following set of equations:

$$\max_{q_{i,t} \in \mathbb{R}_+} \left(P_t(Q_t) \times q_{i,t} - C_{i,t}(q_{i,t}) \right), \tag{1}$$

At each observation t, the supplier i chooses quantity $q_{i,t} \in \mathbb{R}_+$ (where $\mathbb{R}_+$ represents the set of non-negative real numbers) to maximize its profit given the output of competent supplier j, at its optimal choice, $Q_{j,t}$. Q_t is the total quantity supplied to the market, and $P_t(Q_t)$ is the inverse demand function, from which the gas price is deduced. The latter function depends on the total quantity of gas supplied to the market. Finally, production and transmission costs are represented by the cost function $C_{i,t}(q_{i,t})$. The demand function is normally represented by a decreasing straight line and is estimated using regression methods. However, this methodology has its limitation due to endogeneity.

A non-parametric method, with no assumptions on cost and demand functions, has been developed by [24–27]. This analysis does not make any prior assumptions about the cost and demand functions. Therefore, instead of relying on an incomplete set of data related to demand and cost, the non-parametric method avoids such constraints. Nonetheless, many authors have contributed to the literature and exploited the parametric approach, by solving the equilibrium using some sensitivity analysis on cost and demand assumptions [22,23,35].

The marginal cost of supplier i at time t will be denoted by $MC_{i,t}$. The condition of the first order of the optimization problem defined in Equation (1) is given by:

$$Q_{i,t} \times P'_t(Q_t) + P_t(Q_t) - MC_{i,t} \; = \; 0, \tag{2}$$

where with $MC_{i,t} \geq 0$ and $Q_{i,t}$ the solution of (1). Now, we consider the observations given by

$$C = \left\{P_t(Q_t), (Q_{i,t})_{i \in \mathcal{I}}\right\}_{t \in \mathcal{T}}, \quad (3)$$

where $\mathcal{I}$ = *the set of suppliers* and $\mathcal{T}$ = *the set of period's indices* and say that C respects Cournot equilibrium if the following **conditions** are verified:

1. The matrix of data $\left\{MC_{i,t}\right\}_{(i,t) \in \mathcal{T} \times \mathcal{I}}$ must satisfy the following:

$$(P_t(Q_t) - MC_{1,t})/Q_{1,t} = (P_t(Q_t) - MC_{2,t})/Q_{2,t} = \cdots = (P_t(Q_t) - MC_{N,t})/Q_{,t} \geq 0 \ \forall t \in \mathcal{T} \quad (4)$$

where with N = Number of suppliers in the market.

Condition 1 compares the values of marginal costs for both firms. This means that the marginal cost of the firm with the higher marginal cost will produce a quantity that is lower than the firm with the lower marginal costs.

2. Optimal solutions $Q_{i,t}$ must verify the following:

$$(MC_{i,t'} - MC_{i,t})(Q_{i,t'} - Q_{i,t}) \geq 0 \ \forall t \neq t' \in \mathcal{T} \text{ and } \forall i \in \mathcal{I} \quad (5)$$

Condition 2 allows us to compare the costs of different firms at different times. For instance, if firm i changes the produced quantity from $Q_{1,t}$ to $Q_{1,t'}$ the marginal cost at time t' must be lower than the marginal cost at time t. The same analysis for firm j leads to an arrangement of marginal costs for each firm and at each time in increasing order.

To conduct the problem of detection of Cournot's equilibrium, a numerical algorithm was developed. The result of which indicates whether or not the data being tested respects the Cournot equilibrium.

The algorithm starts with an assumption on the initial marginal cost of firm i to produce quantity $Q_{i,t}$ that is equal to the price $P_t(Q_t)$ and tests for conditions 1 and 2. This is typical in a fully competitive market, where the price of any commodity (i.e., gas in our case) should be as close as possible to the delivery cost. This procedure is repeated in several iterations by changing the marginal cost each time until conditions 1 and 2 are fully met. If the algorithm does not converge, then the set of observations C does not respect a Cournot equilibrium. The ratio of the number of observations that respects the Cournot equilibrium (more specifically conditions 1 and 2 in our algorithm) to the total number of observations, is then calculated and is called the Cournot acceptance rate, defined as δ, and is expressed as a percentage (%).

The Cournot Theorem states that, for any market producing and selling a certain commodity, the price converges to its production cost, whenever the number of market participants N tends to be infinity.

$$\lim_{N \to \infty} P = \text{marginal cost} \quad (6)$$

This means that the more companies participate in the market, the more they are unable to affect the market price. The company will become a price-taker and must accept the equilibrium price as is. So it is normal and logical to start by the highest possible marginal cost, which is equal to the gas price at that specific time, and then reducing the cost values until we test the whole array $\left\{MC_{i,t}\right\}_{(i,t) \in \mathcal{T} \times \mathcal{I}}$. This is in line with our initial assumption.

For additional information about the algorithm, the readers are invited to check the four simple steps found in Appendix A.

In simple words, the algorithm can have the following outcomes:

1. Companies are competing based on a Cournot model, trying to maximize their profit by acting strategically. In this case, Cournot acceptance is high.

2. Companies are cooperating and not acting by the rules of Nash Cournot equilibrium. In this case, Cournot acceptance is low.
3. Other strategies and objectives can be set by competing firms, however, the Cournot acceptance rate is only used to differentiate if the latter companies are respecting conditions 1 and 2, which are linked to the Nash-Cournot model.

2.3. Indicator 2—Market Stability

Records theory studies observations that are concentrated in the tail of a given distribution [30] and will be used in this context to test the stability of two different gas markets.

Previous modelers of the record theory have obtained results in the case of independent and identically distributed (*i.i.d*) underlying observations, called the classical case (see [29]). In our application, we consider the absolute value of the difference between two consecutive gas prices as underlying observations. The most popular record model beyond the *i.i.d*. case was introduced by Yang and developed by Nevzorov [28,29,31], and it is currently called the Yang-model. In the latter model, the observations are considered to be independent but not identically distributed.

Considering a time series $\{X_t,\ 1 \le t \le T\}$, where T denotes the present time, the observation X_j is said to be an upper record if and only if $X_j > \max_{t<j} X_t$ and record indicators are a sequence of random variables δ_t defined by:

$$\delta_t = \begin{cases} 1 \; if \; X_t \; is \; a \; record \\ 0 \quad otherwise \end{cases} \tag{7}$$

The total number of records in the considered time series is given by:

$$N_T = \sum_{t=1}^{T} \delta_t \tag{8}$$

Additionally, the probability that the observation X_t represents a record in the *i.i.d* case is called record rate and is given by:

$$P_t = \frac{1}{t} \tag{9}$$

This has been justified by [36], and it can be deduced that the $\lim_{t\to\infty} P_t = 0$, this means that the chance of witnessing a record on a long term level is minimal and rare.

It is only reasonable to go beyond the classical case and test if the Yang model fits our set of data. However, before doing so, one should test if the data comes from a sequence of *i.i.d* random variables. The test is based on the statistic

$$\mathcal{N}_T = (N_T - logT)/\sqrt{logT} \tag{10}$$

which was shown by [29] to have an asymptotic normal behavior.

Moving to the Yang model, the time between two consecutive records converges asymptotically to a geometric distribution and the record rate verifies the following equation:

$$P_t = \frac{\gamma^t(\gamma-1)}{\gamma(\gamma^t-1)} \tag{11}$$

where γ is a parameter that needs to be estimated.

Unlike the classical case, the Yang record rate converges to a constant value in the long term and is given by:

$$\lim_{t\to\infty} P_t = (\gamma-1)/\gamma \tag{12}$$

This means that records are always expected in the long term, and not only observed among the first observations as in the classical case.

2.4. Indicator 3—Volatility and Uncertainty of Prices

The third quantitative method used in this study is represented by Shannon's probabilistic entropy and is used on a time series analysis, to test the predictability power hidden in the underlying probabilistic distribution of the considered time series. A time series with a high predictability power is considered to have a high level of stability with an anticipated pattern.

The classical definition of entropy is as follows: for a given source of information represented by a finite discrete random variable X with n possible outcomes, each possible outcome x_i having a probability p_i to appear, the Shannon entropy H of the random variable X is defined by:

$$H(X) = -\sum_{i=1}^{n} p_i log_2 p_i, \tag{13}$$

In general, a logarithm of base 2 ($\log_2$) is used because the entropy is generally expressed in bits [18]. Several researchers have previously attempted to predict the entropy of the commodity markets (oil, more specifically Brent and West Texas Intermediate, WTI, and other commodities) and tried to measure the information from statistical observations [32–34]. Brent and WTI are two different crude oil grades (quality) and are known to be the most important oil pricing benchmark around the globe. As previously explained, the gas markets in question have been liberalized, and the influence of oil prices on gas prices is shrinking. Gas prices are becoming more influenced by gas to gas competition. To the knowledge of the authors, no previous researchers have worked on predicting the entropy of the gas markets.

To compute the Shannon entropy of a time series, which is a continuous random variable, a particular discretization method is introduced:

First, the returns of the prices are computed, this is a requirement for normalizing the data set

$$r_t = 100 \times \frac{P_t - P_{t-1}}{P_{t-1}}, \tag{14}$$

where P_t and P_{t-1} are the prices at times t and $t-1$ respectively.

It is trivial that the series of observations of returns r_t has an underlying continuous distribution. Therefore, the second step is to introduce the discrete random variable s_t defined by:

$$s_t = \begin{cases} 1 \ if \ \ r_t \geq 0 \\ 0 \ if \ \ r_t < 0 \end{cases} \tag{15}$$

The random variable has a binary output. Unity is when the returns are positive, which means that the prices are increasing, and zero means that the prices tend to diminish.

Based on the observed values of s_t, and by denoting the total number of observations by n, the corresponding probability distribution is computed:

$$p_1 = \mathbb{P}[s_t = 1] = \frac{\sum_{t=1}^{n} s_t}{n}, \tag{16}$$

and,

$$p_0 = \mathbb{P}[s_t = 0] = 1 - p_1 \tag{17}$$

Hence, the entropy related to the random variable s_t is given by:

$$H(s_t) = -p_0 \times log_2(p_0) - p_1 \times log_2(p_1) \tag{18}$$

To compute the underlying entropy of each gas market by the above-explained method, we rely on daily values instead of monthly values. The price variable is divided into year windows, with 252 observations for each. The passage from one window to the other is done by removing the

first observation while adding another from the remaining ones, and so on. By applying the latter procedure, we obtain a series of entropies that should be represented by the mean parameter, as a representative of a series of entropy observations.

However, a major disadvantage of the mean is that it is sensitive to the distribution with a thick queue which can be caused by the presence of outliers and extreme observations. Besides, the mean may have a false interpretation in case of a highly skewed distribution for the considered data. To overcome these weaknesses, a second statistical parameter will be adopted: the median value. Moreover, it has been shown that the median is useful when comparing sets of data. Once the median entropy of each market is computed, the values for each market should be compared. Besides, a non-parametric mathematical test will be conducted to check the statistical significance of the difference between both markets [37]. The Kruskal test is used to compare two independent samples and checks if the observations originate from two different distributions or not.

Finally, the volatility for each market is computed and can serve to validate the results of the price unpredictability. It is the degree of variation in the price series of each market and is measured by the classical standard deviation of the returns.

Note that one can find in literature a version of the entropy adapted to the continuous random variable cases, called differential entropy [38], given by:

$$H(X) \;=\; -\int_{x} f(x) \times \log_2 f(x)\, \mathrm{d}x \tag{19}$$

where $f(\cdot)$ is the probability density function of the underlying distribution of the continuous random variable X. However, this method has many flaws:

- The density function $f(\cdot)$ is in general unknown. This is a weakness because the users will make assumptions about the distribution type. Nonetheless, users can utilize numerical methods to estimate the density function empirically, however, one could face several challenges concerning errors of estimation.
- The properties and the interpretation of discrete random variables entropies are not known to be conserved when passing to the continuous case. In other words, the differential entropy does not share all properties of discrete entropy.

2.5. Information Theory

After defining the indicators, which can be extracted from the gas prices of each market, the regulator has to make important decisions. If the market indicators indicate market concentration and price instability, then certain measures should be taken to bring back stability to the gas prices. Therefore, the two defined states in this study are either for the regulators to take action or keep the business as usual (BAU). This is defined by s_i, $i = 1, 2$ respectively. The indicators previously defined are denoted by y_k, $k = 1, 2, 3$ respectively. Also, we denote by p_{ij} the conditional probability that the market is in state s_i after receiving the indicator y_k i.e.,

$$p_{ik} \;=\; \mathbb{P}\left[\frac{S \;=\; s_i}{Y \;=\; y_k}\right] \tag{20}$$

The probabilities p_{ik} are categorized into three classes: Low, medium, and high. Each class has the following respective probabilities: $\frac{1}{10}, \frac{1}{2}, and \frac{9}{10}$. The information structure is illustrated in Table 3.

Table 3. Information structure conditional probabilities.

States	y_1—Level of Competition	y_2—Market Stability	y_3—Volatility and Uncertainty of Prices
s_1—Action needed	p_{11}	p_{12}	p_{13}
s_2—No action (BAU)	p_{21}	p_{22}	p_{23}

The methods used in this paper reduce the subjectivity of the probability distribution. The indicators are complemented by econometric models founded by economic parameters of the relevant gas markets, and by data analysis on its gas prices, from which the information is extracted.

The probability of being in a certain set (either s_1 or s_2) after receiving the indicator (either y_1, y_2, or y_3) can take either a value of 0.1, 0.5, or 0.9, which constitutes the possible events on the probability set. The sum of the probability of being in s_1 or s_2 after receiving the same indicator y_1, however, should be equal to 1. This is normal as there are only two sets considered in this study.

$$\sum_{i=1}^{2} p_{i1} = p_{11} = \mathbb{P}\left[\frac{S = s_1}{Y = y_1}\right] + p_{21} = \mathbb{P}\left[\frac{S = s_2}{Y = y_1}\right] = 1 \tag{21}$$

Note that the first step is to compute entropy, called "apriori" entropy, based on the distribution of the states S before the reception of any additional information. Then, we start by:

$$H(S) = -\sum_{i=1}^{2} \pi_i \log_2 \pi_i \tag{22}$$

where $\pi_i = \mathbb{P}[S = s_i]$ is the probability of being in the state s_i before receiving additional information called "*apriori probability*". As π_i is defined based on no previous information, it is reasonable to consider a distribution, which has the highest level of uncertainty. In other words, the regulator has no information that can lead him to make an action, and the probability of either of the two states is equally likely. This is a uniform distribution with the following "*apriori probabilities*" $\left(\pi_1 = \pi_2 = \frac{1}{2}\right)$.

Now, after receiving a specific indicator of information y_k, the conditional entropy of the random variable S relative to the indicator y_k is defined as:

$$H_k\left(\frac{S}{y_k}\right) = -\sum_{i=1}^{2} p_{ik} \log_2 p_{ik} \tag{23}$$

Hence, to assess the power of information generated by the whole information structure (composed by y_1, y_2, and y_3), the "*posterior entropy*" is defined, and is compared to the "*apriori probability*":

$$H\left(\frac{S}{Y}\right) = \sum_{k=1}^{3} q_k \times H_k\left(\frac{S}{y_k}\right) \tag{24}$$

where q_k is the weight given for each indicator. This number is equally distributed for the three indicators, as they are equally important and each contributes to the understanding of the gas market in different ways.

Finally, the amount of reduced uncertainty, due to the additional received information, is measured by a quantity called mutual information:

$$I(S, Y) = H(S) - H\left(\frac{S}{Y}\right) \tag{25}$$

Thus, to compare the power of information generated by two information structure related to two different gas market, one should consider the one with the highest $I(S, Y)$. This is equivalent to say

that the structure of information that reduces most of the uncertainty of the random variable S will be most efficient and powerful.

3. Results and Discussions

3.1. Results of the Non-Parametric Cournot Test

Considering Table 1, it is evident that both gas markets are competitive. Nonetheless, this is considerable and significant in the case of HH. Table 1 draws attention to two main numbers: the first being the big difference between the volumes of gas traded in the future and the volume traded on the physical, which indicates the excessive participation for traders and financial players in the virtual market. The second being the large numbers of churn ratios, which indicates high liquidity and healthy trading platform, an attractive characteristic for all stakeholders. Unlike the U.S. gas market, the Herfindahl index for the European gas market is relatively high [11]. This is a sign of healthy competition, and this simply means that out of the many gas suppliers in the U.S. market, none has market power on its own. However, this violates one of the main assumptions of a Cournot competition model, where firms have market power, and each firm's output decision affects the gas prices. In a nutshell, there is no risk of market manipulation in such a market, therefore the market concentration is minimal and close to zero. All U.S. gas suppliers should be price takers in this case, and the Cournot acceptance rate is no longer valid.

As explained in section two, the data that is used for the Cournot test consist of the gas prices and the gas supplies to the relevant market. Gas supplies are shown in Figures 1 and 2. The suppliers are represented by countries of origin. Results can be more indicative if the data related to gas supplies are composed of volumes of the suppliers (shippers, trader, and companies) directly rather than the country (market) where the gas was purchased. The authors acknowledge the need for the traders' suppliers' data and the need to perform the Cournot test on the American market, however with no publically available information on the supply market shares of companies, this is not possible. Therefore, we encourage the publishing agencies to list such data on their website (or upon request). The FERC Form 552 provides a database of trading activity and lists the data related to the largest companies (Top 20) with the largest total transaction volume from year to year. The list found in [15] is incomplete and contains yearly data only. Thus additional data related to suppliers' portfolios is needed to have valid test results. The suppliers in North Western Europe are oligopolistic [39]; therefore, the usage of the data will lead to conclusive and significant results, when using the algorithm.

The non-parametric test results for the European market gives a Cournot acceptance rate of 51%. The results can be analyzed as follows: the behavior of the large gas suppliers in the European can be explained by a Cournot model, where suppliers are trying to maximize their payoffs by competing over quantities. However, the other half of the acceptance rate means that there are companies that are not behaving as such. This could implicate that some of the suppliers have other strategies such as collusion, or strategies that are not "pure" profit maximizers.

An example of a possible collusive behavior has been witnessed in the oil markets under the Organization of the Petroleum Exporting Countries, OPEC back in the 1970s. These countries used to control a major share of the world oil supplies, and together they form a cartel that cooperates, to increase prices and limit external competition.

Other examples that can be used to illustrate possible reasons why these suppliers are not seeking a "profit only" strategy under the Nash-Cournot umbrella are listed below:

Authors such as [40], suggest that Gazprom, a major gas suppliers, is maximizing its "utility function" not only by limiting itself on one strategy that is focused on making a profit, but also by contemplating other strategies such as seeking to eliminate competition, even if this leads to some losses in profits initially.

Other authors, such as [41], enumerate other reasons that are preventing some of the European gas suppliers from exerting their oligopolistic power, and these are due to old legacy contracts that

are still effective, and perhaps new regulations. In short gas prices mechanism in old legacy contracts are mainly indexed to oil prices, and this type of contract does not offer the needed flexibility to gas suppliers. These valid assumptions are among many, possible reasons why the Cournot acceptance rate is not that elevated in Europe.

The first indicator is informative and the analysis of the prices of the NBP wholesale gas prices is indicative for the regulator in this market.

3.2. Results of the Records Theory

The second indicator is assessed using the records theory. To anticipate if the data belong to an *i.i.d* sequence of variables, the goodness of fit test is used.

The results shown in Table 4 indicate that the European market rejects the null hypothesis. The test results were computed at a confidence level of 5%. Accordingly, and from an empirical perspective, the gas prices recorded in this market are characterized by price variations and sudden price shifts.

Table 4. Goodness of fit test results.

Markets	**HH**	**NBP**
p-value	0.89	0.05
Result	Accept H_0	Reject H_0

Based on the analysis of Table 5, the result is not surprising, as it confirms that the European market has a high number of records relative to the small number of observations. This indicates that the European gas price records are not grouped in one section of the time series, and are instead more spread, while the U.S. market is rather more stable and that price shifts are rarely observed all along the time series.

Table 5. Number of records and record index.

Markets	**HH**	**NBP**
Number of Records	*2*	*8*
Record index	November-09 December-09	November-09 December-09 January-10 October-10 December-10 February-12 March-13 April-13

Looking further, in an attempt to measure the probability of witnessing a record in each of the gas markets, the Yang model will be used for the European market and the classical model for the U.S. market. The computed probabilities were computed for each market, and Table 6 shows the result of the probability that matches the date of June 2018.

Table 6. Probability of records results.

Markets	**Probability of Records**	**Probability of Having a Record on t = 105 (June 2018)**
HH	$\mathbb{P}[\delta_t = 1] = \mathrm{P_t} = \frac{1}{t}$	$\mathbb{P}[\delta_t = 1] = 0.01$
NBP, where γ is equal to 1.016	$\mathbb{P}[\delta_t = 1] = \mathrm{P_t}(\gamma) = \left(\gamma^t(\gamma-1)\right)/(\gamma(\gamma^t-1))$.	$\mathbb{P}[\delta_t = 1] = 0.02$

The probability of witnessing a new record is higher in the European market. The results from the above analysis can be summarized as follows:

1. The record rate in the European model converges to a constant value in the medium and long term (as explained in Section 2). This shows that the market could prove to be unstable over time.
2. In contrast, the record rate and the time index in the U.S. model tend to have a negative correlation. This means that the probability of record diminishes over time, i.e., when t increases, the probability of record diminishes. This shows that the market is rather more stable in the medium and long term.

3.3. Results of the Shannon Entropy

By applying the procedure described in Section 2.4 dealing with Shannon entropy, the representative median entropy of each considered market in addition to the p-value of the Kruskal non-parametric test is calculated. The results of the entropy approach are presented in Table 7.

Table 7. Entropy results.

Markets	HH	NBP
Median Entropy	0.9939	0.9991
Kruskal p-value	2.2×10^{-16}	

If a random variable X follows a discrete uniform distribution with n possible outcomes, the corresponding entropy is $H(X) = log_2 n$ [42]. Hence, in our context, the values of the entropies are both close to the case of uniform distribution $log_2 2$, which is equal to unity, and this means that both markets are far from being predictable.

As the values of the median entropies of the considered two markets are close to each other, it is substantial to test if the two considered median entropies are issued from two different distributions. If it is the case, then this indicates a significant difference between the two medians. The non-parametric Kruskal test is applied to verify the latter hypothesis.

Based on Table 7, the p-value of the Kruskal test is close to zero, and therefore less than 5%. Accordingly, the difference between the markets in terms of entropy is significant, i.e., the market with the higher median value, European market in our case, has an entropy significantly higher than the U.S. market.

Also, the volatility in the U.S. market is very low (0.7), whereas it is significantly high in the European market (2.5).

Thus, for indicator number three, the U.S. market is significantly more predictable and has lower uncertainty than the European market.

3.4. Synthesis of Indicator Results

In an attempt to better illustrate the results of the three mathematical models used in the previous sections to measure the market indicators, Table 8 summarizes the results and lists the main findings for each market.

Table 8. Synthesis of indicator results.

Market	y_1—Level of Competition	y_2—Market Stability	y_3—Volatility and Uncertainty of Prices
EU Market	The Cournot acceptance rate is calculated: • $\delta = 51\%$ This number gives a good indication that the oligopolistic gas suppliers in Europe are playing a Cournot game, where the supplier/company is trying to maximize its payoff by competing over quantities. The Cournot acceptance rate also indicates not all the companies are acting as profit maximizers and playing a Cournot game. In fact, less than half of the observations do not respect conditions 1 and 2, which means, companies are not respecting the rules of a Nash-Cournot model. This indicates that some companies are either adopting strategies that are not "pure" profit maximizers on one hand or that they are colluding to increase profit on the other hand. This is unacceptable and thus will reduce consumer welfare.	As per the results of the Goodness of fit test (Table 4), the EU gas prices behave in a non-*i.i.d* case. The record rate calculated for the Yang model converges to a constant value in the long term.The results of the Yang model are as follows: • $\mathbb{P}[\delta_t = 1] = 0.02$, for t = 105, equivalent to the month of June 2018 • $\mathbb{P}[\delta_t = 1] = 0.02$, for t = 1100, equivalent to the month of May 2100 The probability of records converges to constant in the long term, which means that there is always a possibility for sudden and drastic change (spike or drop) in gas prices.	The entropy for the EU market is close to a uniform distribution. • The median entropy for the NBP market is 0.9991. • The vatility of the NBP market is relatively higher and is at 2.5. As for the entropy of the U.S. market, its entropy is also high. Thus, both markets are close to the uniform distribution, which means that both are far from being predictable. The Kruskal non-parametric test result is: • p-value is 2.2×10^{-16} This implies that, relative to each other, the entropy of the NBP gas market is significantly higher than the HH gas market. Thus high volatility and higher median entropy imply that the European gas prices have bigger uncertainty in the medium term and thus not easily predictable.
U.S. Market	As mentioned in Section 3.1, the U.S. market is characterized by: 1. The number of traders participating in the market is high. 2. A large number of churn ratio. 3. Herfindahl index compared to the European gas hubs is low. This simply means, that unlike the European market, no single supplier trading in the U.S. has sufficient market power. In this case, the Cournot acceptance rate is not valid, since this violates one of the main assumptions of a Cournot competition model, where firms have market power, and each firm's output decision affects the gas prices. • $\delta = \mathrm{N/A}$	As per the results of the Goodness of fit test (Table 4), the U.S. market is modeled as the case of a classical record. The results of the classical model are as follows: • $\mathbb{P}[\delta_t = 1] = 0.01$, for t = 105, equivalent to the month of June 2018 • $\mathbb{P}[\delta_t = 1] = 0.00$, for t = 1100, equivalent to the month of May 2100 The probability of records converges to zero in the long term. This simply means that the market is stable, and there is no risk of witnessing sudden price changes.	The entropy for the U.S. market is also close to a uniform distribution. • The median entropy for the HH market is 0.9939. • The volatility of the HH market is very low and is at 0.7. Thus low volatility and lower median entropy imply that U.S. gas prices have smaller uncertainty in the medium term and easier to predict.

3.5. Results of the Information Theory

Concentrated markets raise regulatory and antitrust concerns, as this is a clear sign of market power in the hands of suppliers. Appropriate actions need to be taken by the regulator to make sure that neither collusion, nor cooperation between companies, nor any kind of strategic decisions that do not end up in favor of consumer welfare, are permitted.

The regulator in such a case should ensure that under no circumstances, the companies communicate and have the agreed-upon understanding to raise prices and profit margins at the expense of consumer welfare. The barrier to entry for new companies should also be considered and reduced by regulators in such markets, to increase competition and diversify supplies. These are some examples of actions that the regulator can impose on the suppliers.

Markets that witness price volatility and uncertainty in the medium term, as well as price instability in the long term, are also raising concerns for regulators. In such a case the key to determining the movement of gas prices are supply and demand fundamentals [43]. A slowdown in global demand is a key downside risk for suppliers, as they will eventually earn less while trying to sell their gas. On another hand, a sudden slowdown in supply is a key downside risk for another player in the gas value chain, which is the consumer. The latter will have to pay more to purchase the commodity.

In both cases, regulators should anticipate such results by acting in favor of a continuous supply and demand equilibrium, by trying to diversify supply (indigenous production, imports, and storage), while also ensuring that the consumers have the appropriate infrastructure and financial means to

buy such a commodity. However in a market characterized by gas prices that are predictable in the medium term and prices that are stable over the long term, then there is no need for further actions by the regulator.

Moving forward, we start by assigning the relevant conditional probabilities p_{ik} which indicate to the regulator the state of nature of the gas market. As previously mentioned, the probabilities are categorized into three classes: low, medium, and high. Also important to remember that the sum of $\sum_{i=1}^{2} p_{i1}$ is equal to 1, as we only have two possible sets. The same thing applies to indicators 2 and 3.

After presenting the results of the three indicators in Sections 3.1–3.4, Table 9 explains the process of probability category selection and lists the results for both the U.S. and European markets.

Table 9. Information structure conditional probabilities for both markets.

U.S. Market—HH			
State	y_1**—Level of Competition**	y_2**—Market Stability**	y_3**—Volatility and Uncertainty of Prices**
s_1**—Action Needed**	A liberalized and non-concentrated market, therefore the probability assigned for any action by the relevant regulator is minimal. **LOW**, $p_{11,\text{U.S.}} = 0.1$	Low record probability that converges to zero on a long term, implies that the market is stable and there is no need for additional market oversight measures **LOW**, $p_{12,\text{U.S.}} = 0.1$	Low volatility and lower entropy values (relative to the European market) means that the gas prices are far from being unpredictable, although the entropy values are not that low when analyzed with no benchmark. **MEDIUM**, $p_{13,\text{U.S.}} = 0.5$
s_2**—No Action (BAU)**	The regulator does not need to intervene, and the legal framework along with the supply and demand fundamentals are ensuring a smooth functioning market **HIGH**, $p_{21,\text{U.S.}} = 1 - p_{11,\text{U.S.}} = 0.9$	No intervention needed **HIGH**, $p_{22,\text{U.S.}} = 1 - p_{12,\text{U.S.}} = 0.9$	No intervention needed **MEDIUM**, $p_{23,\text{U.S.}} = 1 - p_{13,\text{U.S.}} = 0.5$
EU Market—NBP			
State	y_1**—Level of Competition**	y_2**—Market Stability**	y_3**—Volatility and Uncertainty of Prices**
s_1**—Action Needed**	High Cournot acceptance rate, implies that the large gas suppliers are trying to maximize their payoffs by competing over quantities. However, the other half of the acceptance rate means that there are companies that are not behaving as such. A possible indication of market abuse, therefore the regulator in invited to intervene. **MEDIUM**, $p_{11,\text{EU}} = 0.5$	High record probability that does not converge to zero on a long term, implies that the market is not stable and there is a need for additional market oversight measures **MEDIUM**, $p_{12,\text{EU}} = 0.5$	High volatility and higher entropy values (relative to the U.S. market) means that the gas prices are not predictable. **HIGH**, $p_{13,\text{EU}} = 0.9$
s_2**—No Action (BAU)**	The regulator needs to intervene, to adjust the legal framework along with the supply and demand fundamentals. This is essential as it ensures a smooth functioning gas market **MEDIUM**, $p_{21,\text{U.S.}} = 1 - p_{11,\text{EU}} = 0.5$	The regulator has to take action. **MEDIUM**, $p_{22,\text{U.S.}} = 1 - p_{12,\text{EU}} = 0.5$	The regulator must take action. **LOW**, $p_{23,\text{U.S.}} = 1 - p_{13,\text{EU}} = 0.1$

The results listed in Table 9, give a clear indication that the market in the United States is functioning smoothly and that the regulator does not need to add other measures. In other words, the BAU case is favored.

This is not the case however, for the European market. The regulator is more inclined to intervene. UK's regulator OFGEM has to intervene and investigate the reason behind some instability and signs of non-competitive behavior, where some firms are not only focused on profit maximization.

To compute the global power of information generated by the considered information structure, we start by assessing the level of uncertainty of each receiving indicator by computing the conditional entropy of the latter $H_k\left(\frac{S}{y_k}\right)$; then we get:

The "*posterior entropy*," previously defined by $H\left(\frac{S}{Y}\right)$ is then computed, and compared with the "*apriori entropy*," which is defined in Section 2.5 as the entropy of a uniform distribution (one that has the highest level of uncertainty), and given a value of 1.

Tables 10 and 11 illustrate the results of the entropies, conditional to the relevant indicators, which is then used to compute the outcome of these indicators in aggregated dimension and for each market.

Table 10. Conditional entropy of each indicator.

Market	y_1—Level of Competition	y_2—Market Stability	y_3—Volatility and Uncertainty of Prices
U.S.—HH	0.47	0.47	1
Europe—NBP	1	1	0.47

Table 11. "Posterior" and "apriori" entropies.

Market	*Posterior*	*Apriori*
U.S.—HH	0.64	1
Europe—NBP	0.82	1

The difference between the "*posterior entropy*" and the "*apriori entropy*" will help assess the level and amount of information, previously defined as $I(S, Y)$ gained by analyzing the gas prices data in each market. In other words, the indicator analyses that is measured by the various econometric methods used in this study constitute additional information that the regulators can use to assess the status of the market. The more the additional information increases (i.e., the difference between the "*posterior entropy*" and the "*apriori entropy*"), the more confident the regulator is about the power of information generated.

The amount of reduced uncertainty, due to the additional information received from the indicators, is estimated at 0.38 for the U.S. market and 0.18 for the European market, which means that the level of uncertainty has been reduced in the European and U.S. market respectively by 18% and 38%. The value of the information contained in both markets, although in asymmetric terms, is significant, powerful, and can serve as a reliable and efficient source of information.

4. Conclusions

Overall, this work presents four econometric and mathematical methods that are used collectively to estimate the level of information contained in gas prices in two separate wholesale gas markets, i.e., the European and the U.S. gas markets. The theories employed are Cournot theory, records theory, Shannon entropy, and information theory.

By analyzing the efficiency of the gas market and assessing the need for additional measures and intervention, the work of gas regulators with regard to market oversight is likely to be improved. The value of the information is based on three market indicators: the possibility of non-competitive behavior by gas firms, market stability, and uncertainty in prices.

Our findings suggest that the U.S. gas market is stable. The information value contained in the wholesale gas prices gives a clear indication that there is no need for additional market oversight. However, this is not the case in the UK (the most developed European gas market), where results show signs of market instability and non-competitive behavior. In other words, some firms are not only focused on profit maximization; therefore, the wholesale prices are not solely the product of classical law of supply and demand.

Interestingly, the value of the additional information brought about by the indicator analysis and included in both markets has contributed to reducing uncertainty. This makes the information carried in the gas prices of both markets, although asymmetric, powerful and efficient. The regulators in both markets can, therefore, act accordingly by using the two-step approach to assess the level of competition, price stability, and price predictability.

The originality of the two-step approach applied in this document can be summarized as follows: it is the first time that several multidisciplinary econometric methods have been combined to create a probabilistic structure assessing the underlying information of a gas market. Furthermore, the approach deals with a case where the information is neither completely absent nor perfectly known, which has rarely been dealt with in literature.

Worthy to mention, the authors have chosen three market indicators and four different econometric methods in this study. It is believed that additional mathematical/statistical analysis can be used for this topic. For further research, one can work on estimating the entropy generated (the third indicator) using another discretization procedure. This work is a growing research track and needs a large number of observations. Besides, one can also work on creating estimators for the underlying probability distribution of each indicator. A starting point is to apply goodness of fit techniques or a more empirical to perform a bootstrap process.

Author Contributions: Formal analysis, H.H.; methodology, A.H. and H.H.; project administration, H.A.; software, A.H. and H.H.; writing—original draft, H.H.; writing—review and editing, H.A. and H.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors acknowledge TU Wien Bibliothek for financial support through its Open Access Funding Program.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Algorithm for Non-parametric Cournot test

- ***Step 1:*** Consider the prices at time t (P_t) as a starting point of the upper bound of the marginal cost ($MC^{ub}_{i,t}$) of the supplier i at time t. i.e., $MC^{ub}_{i,t} = P_t$.
- ***Step 2:*** Define for each supplier i and at time t the variable $\gamma_{i,t}$:

$$\gamma_{i,t} = \min\left\{\min_{\{t' \neq t:\ Q_{i,t'} > Q_{i,t}\}}\left\{MC^{ub}_{i,t'}\right\}, MC^{ub}_{i,t}\right\}$$

  Note that if $\left\{t' \neq t:\ Q_{i,t'} > Q_{i,t}\right\} = \phi$ set $\gamma^{ub}_{i,t} = MC^{ub}_{i,t}$. This step ensures that Condition 1 is respected.
- ***Step 3:*** Define for each supplier i and at each time t the variables λ_t and $\gamma^{ub}_{i,t}$:

$$\lambda_t = \max_j\left\{\frac{P_t - \gamma_{j,t}}{Q_{j,t}}\right\}$$

  and

$$\gamma^{ub}_{i,t} = P_t - \lambda_t Q_{i,t}$$

  This step ensures that Condition 2 is respected.
- Step 4:

  i. If $\exists(i,t)$ such that $\gamma^{ub}_{i,t} < 0$, then the algorithm is stopped, and it is concluded that Cournot equilibrium conditions are not satisfied.
  ii. If $\forall(i,t)$, $MC^{ub}_{i,t} = \gamma^{ub}_{i,t}$, then the algorithm is stopped, and it is concluded that Cournot equilibrium conditions are satisfied.
  iii. Otherwise, return to Step 1 for a new iteration by letting $MC^{ub}_{i,t} = \gamma^{ub}_{i,t}$ for all (i,t).

- ***Finally,*** if the algorithm does not stop at or before iteration T (number of periods), it is stopped by force, and it is concluded that Cournot equilibrium conditions are not satisfied.

References

1. EIA, Perspectives on the Development of LNG Market Hubs in the Asia Pacific Region, Washington, 2016. Available online: https://www.eia.gov/analysis/studies/lng/asia/pdf/lngasia.pdf (accessed on 11 June 2019).
2. Franza, L. Long-Term Gas Import Contracts in Europe. The Evolution in Pricing Mechanisms, Netherland, 2014. Available online: http://www.clingendaelenergy.com/inc/upload/files/Ciep_paper_2014-08_web_1.pdf (accessed on 11 June 2019).
3. FERC. Market Oversight—Federal Energy Regulatory Commission (FERC), FERC Website. 2016. Available online: https://www.ferc.gov/market-oversight/guide/guide.asp?csrt=11553422067829331379 (accessed on 16 June 2019).
4. James, R.P.S. Gas Regulation—United States, Getting the Deal Through. 2018. Available online: https://gettingthedealthrough.com/area/15/jurisdiction/23/gas-regulation-united-states/ (accessed on 16 June 2019).
5. Brown, Gas Regulation—United Kingdom, Getting the Deal through M. 2018. Available online: https://gettingthedealthrough.com/area/15/jurisdiction/22/gas-regulation-united-kingdom/ (accessed on 16 June 2019).
6. Shi, X. Development of Europe's gas hubs: Implications for East Asia. *Nat. Gas Ind. B* **2016**, *3*, 357–366. [CrossRef]
7. Huntington, H.G. Industrial natural gas consumption in the United States: An empirical model for evaluating future trends. *Energy Econ.* **2007**, *29*, 743–759. [CrossRef]
8. OFGEM, Gas Demand and Supply Source by Month (GB). 2020. Available online: https://www.ofgem.gov.uk/data-portal/gas-demand-and-supply-source-month-gb (accessed on 5 June 2019).
9. Fulwood Mike, Asian LNG Trading Hubs: Myth or Reality, 2018. Available online: https://energypolicy.columbia.edu/research/report/asian-lng-trading-hubs-myth-or-reality (accessed on 7 June 2019).
10. Petrovich, B. *The Cost of Price de-Linkages between European Gas Hubs*; The Oxford Institute for Energy Studies: Oxford, UK, 2015.
11. Heather, P. *The Evolution of European Traded Gas Hubs*; The Oxford Institute for Energy Studies: Oxford, UK, 2015.
12. Heather, P.; Petrovich, B. European Traded Gas Hubs. In *European Traded Gas Hubs*; The Oxford Institute for Energy Studies: Oxford, UK, 2017.
13. Petrovich, B. *Do We Have Aligned and Reliable Gas Exchange Prices in Europe?* The Oxford Institute for Energy Studies: Oxford, UK, 2016.
14. OFGEM, Gas Trading Volumes and Monthly Churn Ratio by Platform (GB). 2018. Available online: https://www.ofgem.gov.uk/data-portal/gas-trading-volumes-and-monthly-churn-ratio-platform-gb (accessed on 5 June 2019).
15. Cornerstone Research, Characteristics of US Natural Gas Transactions 2016, (**2016**) 8. Available online: https://www.cornerstone.com/Publications/Reports/Characteristics-of-US-Natural-Gas-Transactions-2016 (accessed on 12 June 2019).
16. Haris, M.; Tao, J. Role of governance in creating a commodity hub: A comparative analysis. *Nat. Gas Ind. B* **2016**, *3*, 367–376. [CrossRef]
17. Shannon, C.E. Communication Theory of Secrecy Systems. *Bell Syst. Tech. J.* **1949**, *28*, 656–715. [CrossRef]
18. Yang, J. Information Theoretic Approaches in Economics. *J. Econ. Surv.* **2017**, *32*, 940–960. [CrossRef]
19. Crémer, J. A simple proof of Blackwell's "comparison of experiments" theorem. *J. Econ. Theory* **1982**, *27*, 439–443. [CrossRef]
20. Bohnenblust, Reconnaissance in Game Theory, RAND Memorandum. 1949. Available online: https://www.rand.org/pubs/research_memoranda/RM208.html (accessed on 21 June 2019).
21. Jones, R.A.; Ostroy, J.M. Flexibility and Uncertainty. *Rev. Econ. Stud.* **1984**, *51*, 13. [CrossRef]
22. Golombek, R.; Gjelsvik, E.; Rosendahl, K.E. Effects of Liberalizing the Natural Gas Markets in Western Europe. *Energy J.* **1995**, *16*, 85–111. [CrossRef]
23. Golombek, R.; Gjelsvik, E.; Rosendahl, K.E. Increased Competition on the Supply Side of the Western European Natural Gas Market. *Energy J.* **1998**, *19*, 1–18. [CrossRef]
24. Carvajal, A.; Deb, R.; Fenske, J.; Quah, J.K.-H. Revealed Preference Tests of the Cournot Model. *Econometrica* **2013**, *81*, 2351–2379. [CrossRef]
25. Forges, F.; Minelli, E. Afriat's theorem for general budget sets. *J. Econ. Theory* **2009**, *144*, 135–145. [CrossRef]

26. Ray, I.; Zhou, L. Game Theory via Revealed Preferences. *Games Econ. Behav.* **2001**, *37*, 415–424. [CrossRef]
27. Sprumont, Y. On the Testable Implications of Collective Choice Theories. *J. Econ. Theory* **2000**, *93*, 205–232. [CrossRef]
28. Yang, M.C.K. On the distribution of the inter-record times in an increasing population. *J. Appl. Probab.* **1975**, *12*, 148–154. [CrossRef]
29. Arnold, B.C.; Balakrishnan, N.; Nagaraja, H.N. *Records*; Wiley: Hoboken, NJ, USA, 1998; pp. 7–50.
30. Hamie, H.; Hoayek, A.; Auer, H. Modeling the price dynamics of three different gas markets-records theory. *Energy Strat. Rev.* **2018**, *21*, 121–129. [CrossRef]
31. Nevzorov, V.B.; Balakrishnan, N. 19 A record of records. *Handb. Stat.* **1998**, *16*, 515–570. [CrossRef]
32. Mensi, W.; Beljid, M.; Managi, S. Structural breaks and the time-varying levels of weak-form efficiency in crude oil markets: Evidence from the Hurst exponent and Shannon entropy methods. *Int. Econ.* **2014**, *140*, 89–106. [CrossRef]
33. Bariviera, A.F.; Zunino, L.; Rosso, O.A. Crude oil market and Geopolitical Events: An Analysis Based on Information-Theory-Based Quantifiers. *Fuzzy Econ. Rev.* **2016**, *21*, 41–51. [CrossRef]
34. Benedetto, F.; Giunta, G.; Mastroeni, L. On the predictability of energy commodity markets by an entropy-based computational method. *Energy Econ.* **2016**, *54*, 302–312. [CrossRef]
35. Blitzer, C. *Western European Natural Gas Trade Model*; MIT International Gas Trade Project; Center for Energy Policy Research, MIT Energy Laboratory: Cambridge, MA, USA, 1986.
36. Nevzorov Valery, B. *Records-Mathematical Theory*; American Mathematical Society: Providence, RI, USA, 2001.
37. Kruskal, W.H.; Wallis, W.A. Use of Ranks in One-Criterion Variance Analysis. *J. Am. Stat. Assoc.* **1952**, *47*, 583–621. [CrossRef]
38. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; Wiley: Hoboken, NJ, USA, 2005.
39. Yang, Z.; Zhang, R.; Zhang, Z. An exploration of a strategic competition model for the European Union natural gas market. *Energy Econ.* **2016**, *57*, 236–242. [CrossRef]
40. Jansen, T.; Van Lier, A.; Van Witteloostuijn, A.; Von Ochssée, T.B. A modified Cournot model of the natural gas market in the European Union: Mixed-motives delegation in a politicized environment. *Energy Policy* **2012**, *41*, 280–285. [CrossRef]
41. Egging, R.; Gabriel, S.A.; Holz, F.; Zhuang, J.; Egging, R. A complementarity model for the European natural gas market. *Energy Policy* **2008**, *36*, 2385–2414. [CrossRef]
42. Gray, R.M. *Entropy and Information Theory*; Springer Publishing Company: Berlin/Heidelberg, Germany, 2011. [CrossRef]
43. Hamie, H.; Hoayek, A.; Auer, H.; Kamel, M. Northwestern European wholesale natural gas prices Comparison of several parametric and non-parametric forecasting methods. *Int. J. Glob. Energy Issues* **2020**. [CrossRef]

energies

MDPI

Article

A New Methodology to Obtain a Feasible Thermal Operation in Power Systems in a Medium-Term Horizon

Luis Montero *, Antonio Bello and Javier Reneses

Institute for Research in Technology (IIT), ICAI School of Engineering, Comillas Pontifical University, 28015 Madrid, Spain; antonio.bello@iit.comillas.edu (A.B.); javier.reneses@iit.comillas.edu (J.R.)
* Correspondence: luis.montero@iit.comillas.edu

Received: 27 March 2020; Accepted: 10 June 2020; Published: 12 June 2020

Abstract: Nowadays, electricity market paradigms are constantly changing. On the one hand, the deployment of non-dispatchable renewable energy sources is bringing out the necessity of representing hourly dynamics in medium-term fundamental models. On the other, the promotion of new interconnection capacity and the integration of markets (as is the case of the European market) makes necessary the simultaneous modeling of multiple electricity systems. Thus, the large size of power markets, together with the consideration of uncertainty in some inputs, make it computationally intractable to work rigorously on an hourly detailed time span. Temporal aggregation, integer programming relaxation or less accurate generation modeling are usually employed to obtain reasonable computation times. However, the application of these techniques often leads to infeasible or suboptimal operational outputs. This paper proposes a new soft-linking methodology to meet reliable results from medium-term models, such as hourly prices or aggregated productions, with a feasible and detailed representation of the thermal generation, considering technical constraints and risk aversion. The results of a fundamental model that represents the competitive behavior between market players in a multi-area power system are used as the starting point for the methodology. Then, a post-processing method is applied to optimize and make feasible the thermal portfolio of a market agent. The final output is a feasible hourly scheduling and an ample space for optimization, where the introduction of a strategic term represents the rational behavior of a player who tries to maximize its profit.

Keywords: electricity markets; feasible operation; medium-term representation; optimization models; power systems; thermal generation; unit commitment

1. Introduction

The liberalization of electricity markets over the last decades has highlighted the importance for generation companies to optimize the production of their thermal units in order to maximize profits and be competitive. Power market models have traditionally been used as a supporting tool for this purpose and have proven their effectiveness. However, the accurate representation of current market trends in real size systems has become a great challenge. In particular, there are four main sources of the increasing complexity in electricity market modeling.

First, the maturity status of the onshore and offshore wind generation, together with the recent drop in the price of photovoltaic facilities, are leading to major changes in traditional power mixes. The variability of electricity supply will be accentuated in the near future with an increasing penetration of these renewable energy sources. This trend brings out the need to consider a more detailed time granularity in the horizons of energy models, as well as detailing the flexibility of thermal units [1].

This is especially crucial when trying to represent and assess the behavior of storage (pumping, batteries, etc.) facilities, which are highly influenced by time chronology.

Second, the current market paradigms do not only demand great modeling detail. The energy integration policies adopted in many regions have promoted a notable increase in the interconnection capacities between countries. Additionally, the diversification strategies in markets embraced by some players to cover risks have also contributed to enlarging the simulation dimensions, making it necessary to represent multi-area power systems in order to analyze an electricity market [2]. In particular, the European electricity market is becoming more and more integrated and is now cleared on a multi-country basis.

Third, the mentioned increase in intermittent renewable generation makes it necessary to properly address the technical constraints, costs, and flexibility of thermal units. The fast-ramping ability of gas-fired power plants makes them the best medium-term option to assure a reliable electricity service, covering demand peaks, plant falls and drastic variations in the renewable generation. This fact points out the importance of an accurate representation of the operation of flexible thermal units like combined cycle gas turbines (CCGTs) [3].

Finally, the competitive behavior of generation companies should also be included in energy models. This is relevant both regarding profit maximization and risk aversion. The continuous evolving of energy regulation, the ongoing generation switching, as well as the increasing deployment of intermittent energy resources, bring out the increasing importance of the uncertainty consideration [4]. This means that models have to be run using a number of scenarios of the different risk variables.

These changing market conditions can be represented accurately through the application of short-term methodologies, which allow a rigorous representation of large power systems using a short and detailed time horizon. However, generators, retailers and large consumers also need medium-term tools to optimize the operation of their assets and support the decision-making process, like fuel purchases, hydro-thermal management, emissions allowances trading or financial contracting [5]. In particular, the role of natural gas as a vector in the energy transition towards a renewable mix makes it essential for generation companies to know how to obtain the maximum return on these assets [6].

Despite the continuous computational improvements, the extension of the accurate short-term techniques to longer horizons is computationally intractable. For this reason, modeling simplification is necessary, and some assumptions are carried out in the literature in order to reduce the size of the considered problems. Nevertheless, the combination of high-detail modeling, multiple areas, hourly granularity, and uncertainty consideration at once, is increasingly desired. These aspects are analyzed in depth in Section 2.

This paper proposes a new soft-linking methodology to fill this gap, meeting reliable results from medium-term models, such as hourly prices or aggregated productions, with a feasible and detailed representation of the thermal generation, considering technical constraints and risk aversion.

2. Literature Review

The operation of power systems is a subject widely studied in the literature. The representation and optimization of the hydro-thermal generation has been deeply addressed in order to increase its profitability. For this purpose, several modeling techniques have been developed to reach an accurate performance of the simulation tools.

A suitable example for the rigorous representation of large power systems is the unit commitment (UC) problem. It provides the optimal dispatch of thermal units according to price and demand forecasts. Technical constraints, operating costs, and profit maximization can be easily modelled with this methodology and useful results are obtained in reasonable running times. In most cases of the literature, e.g., [7–13], UC considers one day to one week time spans in an hourly basis, performing a precise simulation of thermal generation that cannot be extended to longer time horizons.

The representation of the strategic behavior between players in competitive electricity markets is also a problem widely discussed in the literature. A great variety of methodologies have been proposed

to address the Nash's game theory [14] applied to electricity markets [15]. Diverse models, based on the mixed complementary problem [16,17], heuristic techniques [18] or dynamic programming [19] have successfully described competitive behavior in the medium term.

However, market equilibrium tends to be represented as an optimization problem. With this aim, behavior assumptions are made, such as perfect competition [20] or Cournot competition [21]. Both competitions are encompassed by conjectural-variation (CV) models, as well as a wide range of intermediate situations between these two extreme behaviors. Quite accurate simulations of oligopolistic markets can be performed through these models. In [22], a medium-term CV-based model is proposed, where the equivalence of this optimization problem to a market equilibrium problem is also mathematically demonstrated.

Medium-term equilibrium models experience a tight trade-off between modeling detail and run time. Despite the improvements in computational techniques, a completely hourly detail in real size cases with uncertainty consideration is still intractable. Nevertheless, renewable penetration and market integration developments bring out the necessity of a multi-area medium-term modeling in an hourly basis, in order to achieve an adequate representation of the importance of the greater supply variability and the interconnection capacity increase.

Generic formulations are frequently employed in many academic and commercial models for the medium- and long-term representation of energy systems [23–37]. These open presentations increase their flexibility and brings custom options to the users. Nonetheless, these formulations are not always followed by a case study to show the real scope of the model. Temporal horizons, thermal unit details or multi-area limitations are not usually fenced when a new model is presented.

Many of these methodologies offer a wide modeling catalogue to perform a meticulous simulation. High granularity in temporal representation, long-term horizons, multi-area representation, integer programming, even the inclusion of uncertainty, are often available in the same model. Nevertheless, it is computationally intractable to consider every technique at a once in a medium- or long-term representation of a real power system.

As an example, Table 1 shows the particularities for each model formulated in [23–37], where the difficulties of considering every single accurate modeling technique at once are exposed. It is demonstrated that simplifications are necessary when the time span exceeds short-term horizons.

The uncertainty representation in a detailed medium-term horizon frequently implies to renounce to integer programming and its accurate modeling of the power system operation [38,39]. In fact, integer variables are also relaxed if a high granularity along the whole time span is desired [40–42]. Otherwise, if a MIP performance is necessary, time aggregation techniques are usually implemented [30,35], either temporal decoupling [43] or a drastic reduction in the problem size [31]. Other representative examples of the trade-off between modeling detail and computational resolution of medium-term models are illustrated in [29,44,45].

As expected, energy market representations in the long term require the same modeling simplifications. The combination of integer programming with a complete hourly resolution can be hardly afforded even if a tiny power system is considered [46]. Integer variable relaxation and temporal aggregation techniques [37] or a low time granularity [32], are needed if the representation of real size electricity markets is desired. In conclusion, medium- and long-term models cannot afford the whole modeling details at once in an hourly basis.

These cases highlight the need to use temporal aggregation techniques for the representation of large electricity markets in the medium term if a high temporal granularity is desired. Load blocks [47] have been traditionally employed as clustering technique to reduce the consumption of computational resources in hourly time spans. These clusters represents the system through demand levels.

The variability introduced by the penetration of non-dispatchable renewable technologies can be modelled with a load duration curve that characterizes the net demand [48]. This formulation was overcome by the system states [49,50], which ushered the inclusion of multiple system features and introduced cluster-transition concepts to increase the accuracy of the model.

Table 1. Modeling simplifications of each case study. IP: Integer Programming.

Ref	Model	Case Study	IP	Resolution	Detail
[23]	Backbone	Multi-area - [38]	No	Clusters	Cluster moving window Aggregation per technology Stochastic programming
[24]	Balmorel	Multi-area - [40]	No	Hourly	8760 time steps per year Only RES generation
[25]	Calliope	Multi-area - [44]	No	Clusters	550 time steps per year Aggregation per technology
[26]	COMPETES	Multi-area - [45]	No	Clusters	12 time steps per year Detailed thermal units + RES
[27]	DIETER	Single-area - [41]	No	Hourly	8760 time steps per year Aggregation per technology
[28]	EMMA	Multi-area - [39]	No	Hourly	8760 time steps per year Aggregation per technology Monte Carlo simulation
[29]	EnergyScopeTD	Single-area - [29]	No	Typical days	288 hourly steps per year Aggregation per technology
[30]	ESO-XEL	Single-area - [30]	Yes *	Clusters	12 time steps per year 1722 thermal units + RES
[31]	Ficus	Single-area - [31]	Yes	15 min	35040 time steps per year 1 single factory
[32]	MultiMod	Multi-area - [32]	No	10 years	1 time step per 10 years Aggregation per technology
[33]	OSeMOSYS	Single-area - [42]	No	Clusters	12 time steps per year Aggregation per technology
[34]	PLEXOS	Multi-area - [43]	Yes	30 min	Daily moving window 760 thermal units + RES
[35]	Switch	Multi-area - [35]	Yes	Typical hours	144 hourly steps per year 578 thermal units + RES
[36]	TIMES	Single-area - [46]	Yes	Typical days	288 hourly steps per year 6 generation units (RES incl.)
[37]	URBS	Multi-area - [37]	No	Typical weeks	1008 hourly steps per year Aggregation per technology

* Integer variables are relaxed when the time span exceeds one year.

Nonetheless, the chronology between clusters was not taken into account until [51,52], where new constraints were formulated to keep the technical information in the transitions between clusters. In this way, chronological relationships are maintained using system states or enhanced representative periods. Regardless, as every single temporal aggregation method, they sacrifice some detail to gain in resolution time.

The application of these techniques to optimization models often leads to suitable results in the medium-term forecasting of power systems. The methodology proposed in [53] for a medium-term fundamental model based on conjectural variations represents uncertainty in risk variables such as demand, water incomes, wind generation, fuel costs, CO_2 prices and unavailability of thermal units, reaching a great characterization of the hourly prices obtained for the Spanish electricity market as a case study. Nevertheless, it does not offer a feasible hourly scheduling.

The application of integer programming to cases of such a large size and detail as the previously mentioned is computationally intractable in hourly horizon representations. The inclusion of clustering techniques in the unit commitment problem [54–56], open the door to extend the time spans of models

that work with integer variables, allowing to represent in detail the operation of large energy markets in the medium term.

Moreover, the use of temporary aggregation techniques complicates the representation of the start-ups, since its cost depends on the number of hours that the thermal unit has been offline. The representation of this cost as a single step [56] does not provide a great detail. However, the formulation of [54] seems to overcome this problem, but the representation by centroids will always result in a loss of the real variability existing between the elements that integrate the cluster.

In [57], a new methodology is proposed in order to preserve the variability between hours in the performance of a medium-term model that considers system states. The use of statistical techniques results in a notable improvement in the representation of the hourly prices of the electricity market. Nevertheless, their use cannot be extrapolated to the production outputs. In turn, it would not solve the problem of the hourly infeasibilities either, which appears when the problem size or the uncertainty consideration oblige to use relaxed variables if reasonable run times are sought.

Integer programming relaxation is widespread in medium-term models. The use of continuous variables provide differentiability, exploited by powerful current solvers to solve large LP and QC problems without the consuming too much time or computing resources. However, its utilization means a loss of fidelity of results because of its acceptance of fractional values as levels of decision variables like the commitment status or the start-ups and shut-downs of thermal units. For this reason, some outputs are far from the real behavior of the power system.

Continuous variables allow partial commitment status for thermal units, as well as non-integer internalization of the costs of start-ups and shut-down processes. This means that production results and hourly costs cannot always be extrapolated to reality. However, since a minor detail perspective such as a monthly vision, some outputs of these models like productions, incomes and costs are quite close to the real values for each player, being a useful information for generation companies.

According to the high usefulness of these results for companies and their approximation to real values, a novel post-processing method is proposed in this paper in order to obtain a feasible hourly scheduling for the thermal generation portfolio of a market player. The scope of this methodology is to overcome the mentioned gaps of the existing models in the literature, being flexible to be applied to several medium- or long-term simulation tools. The main contributions of this paper are:

- A new soft-linking methodology to meet the advantages of the accurate market representation of medium-term models, with the detailed and feasible schedules of the short-term modeling, is proposed. The market equilibrium of real-size multi-area power systems, where players have a competitive behavior, is represented under the consideration of uncertainty. In turn, the infeasible outputs from simplification techniques are rectified and hourly dynamics are properly captured.
- This methodology is opened to probabilistic considerations and risk management. The inclusion of these assumptions in the method returns valuable results for a market player, identifying the most profitable hours to place its production according to the technical constraints of its thermal units and the margins in which its sales are framed.
- The proposed methodology is flexible and allows the combination of the detailing considerations that the medium-term models cannot assume at once. In agreement with economic and technical constraints, this method can group low productions into thermal units with higher operation levels, beyond obtaining a feasible scheduling. This flexible processing scheme can reach the integrated optimization of a entire thermal portfolio if desired.

3. Methodology

3.1. Overview

As previously discussed, modeling simplifications are usually assumed in medium- and long-term models. Temporal aggregation, integer programming relaxation or less accurate modeling are used to achieve computational feasibility. Furthermore, risk representation notably increases the complexity of

the problem. Considering uncertainty is more difficult the longer the horizon is, either by stochastic programming or Monte Carlo simulations.

This section proposes a methodology to harmonize the performance of a medium-term model, subject to modeling simplifications, with short-term techniques which bring a detailed representation of power markets operation.

3.1.1. Medium-Term Fundamental Model

The methodology takes as inputs the high-reliable outputs from a medium-term market model. To this end, a medium-term fundamental market model will be solved. This model looks for a detailed representation of the operation of an electricity market, considering its regulation framework and technical constraints. The main features of this kind of models are described below:

- As mentioned above, this paper considers a market based on a multi-area system. The model should include every single thermal unit, hydro reservoirs and non-dispatchable generation technologies, as well as energy storages.
- It is also desirable a properly representation of the interconnection facilities, both between the considered areas, and with the external regions adjacent to the studied systems.
- Furthermore, market agents must be considered, since it is necessary to simulate the competitive behavior of the players to reach an accurate performance of the operation in the power system.
- This rigorous modeling should be complemented with uncertainty representation to capture the risks associated with some generation technologies or supply contract compliance.
- Given the changing current market trends, where non-intermittent generation is rapidly increasing, a time representation as closely to hourly modeling would be desirable.

Generally, some simplifications are needed in order to make these models tractable from a computational point of view. First, integer programming relaxation is typically needed, and this relaxation collaborate in getting good price signals, since the use of integer programming only reflects the variable costs of the units that are committed, when the dual variables which represent the prices are obtained. Secondly, the application of time aggregation or selection techniques is also required, either by aggregating similar hours in time blocks (and consequently losing the chronology), or with the selection of prototypical days or weeks. Either simplification is able to obtain accurate aggregated results (such as weekly or monthly productions) but will fail obtaining feasible and realistic hourly operations. Finally, some simplifications are needed regarding thermal units, as detailed start-up costs depending on the time that the unit has been offline.

The proposed methodology will be tested with a particular medium-term model, but the formulation is open to the consideration of any medium- or long-term model, whose operation results are desired to be made feasible and optimal on an hourly basis.

3.1.2. Post-Processing Methodology

Once the results of the medium-term model are obtained, they are used as an input data by the post-processing method, which provides a detailed, accurate, and feasible thermal schedule. The methodology will be stated from the point of view of a thermal agent trying to obtain a feasible operation of its thermal portfolio. This feasibility process is formulated as an optimization problem, which responds to the rational behavior of a player that wishes to maximize its profit considering risk aversion. The methodology uses as input data the expected hourly prices and the expected productions of the thermal portfolio through a considered time span. It is important to note that these results respond to a rational infrastructure management, like hydro reservoirs, fuel storage, maximum number of start-ups, minimum annual operation hours etc. These expected productions are the aggregated values of each thermal unit included in the portfolio. The compliance of the production goals has a certain flexibility degree, according to a strategic term that could be adjusted by the market agent. Hence, the optimization process can only consider some clearance to the production targets to avoid

infeasibilities, or it can have a more flexible character, in which a redistribution of productions between units is allowed. It will depend on the strategic term, that can be easily determined by the market player according to extra operational cost, logistics, opportunity costs and any other desired consideration.

The next section describes the mathematical formulation of the methodology. Its nomenclature is included in a glossary at the end of the document. Upper-case letters are used for denoting parameters and sets, while lower-case letters denote variables and indexes. Hourly intervals are considered for unit consistency.

3.2. Mathematical Formulation

The post-processing method is presented as an optimization problem in which a player wishes to maximize its profits, adjusting a given production of its thermal units. In addition to this input, an hourly price market forecast throughout the considered time span is also taken as an input. This time horizon is flexible, being able to cover days, weeks, or even months.

The objective function of the maximize optimization problem is shown below. It is subject to the restrictions formulated along this section:

$$max \sum_{g \in G} \sum_{t \in T} (p_{g,t} L_t - c_{g,t}^{PROD} - c_{g,t}^{SD} - c_{g,t}^{SU}) - \sum_{g \in G} c_g^{DIV} \tag{1}$$

3.2.1. Production Costs

The production cost of the thermal units are modelled as a quadratic function of the power output:

$$c_{g,t}^{PROD} = u_{g,t} C_g^{NL} + p_{g,t} C_g^{LV} + p_{g,t}^2 C_g^{QC} \qquad \forall g, t \tag{2}$$

This formulation is more accurate than the simplified linear production costs which are taken in [9,12,56]. However, it is less detailed than the piecewise approximation of [7], where the use of binary variables allows the non-convex and non-differentiable variables costs of the thermal units to be segmented and adjusted with high accuracy.

The choice of a quadratic function is based in the power of the current solvers to work with MIQCPs and prevent the MILP problem from slowing down with a big number of binary variables, whose resolution requires a long run time. Besides, it constitutes a high-accurate approximation to the actual cost function.

3.2.2. Start-Up and Shut-Down Costs

The exponential nature of the start-up cost function, where cost increases with the amount of hours that the unit has been offline, is usually represented by a stairwise approximation, as Figure 1 illustrate, or simplified to a single step cost [9,56].

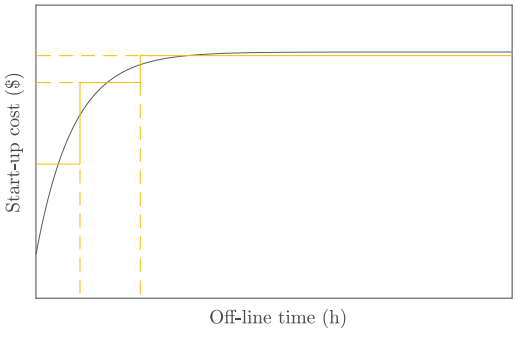


Figure 1. Stairwise approximation of the non-linear start-up cost function.

According to the advantages of the formulation of [12] over those of [10], the equations of [12] are chosen to represent the behavior of the start-ups, consuming less computational resources:

$$c^{SU}_{g,t} = \sum_{s \in S} \delta_{g,s,t} C^{SU}_{g,s} \qquad \forall g, t \tag{3}$$

$$\delta_{g,s,t} \leq \sum_{i=T^{SU}_{g,s}}^{T^{SU}_{g,s+1}-1} w_{g,t-i} \qquad \forall g, s \in [1, S_g), t \in [T^{SU}_{g,s+1}, T] \tag{4}$$

$$v_{g,t} = \sum_{s \in S_g} \delta_{g,s,t} \qquad \forall g, t \tag{5}$$

$$\delta_{g,s,t} = 0 \qquad \forall g, s \in [1, S_g), t \in (T^{SU}_{g,s+1} - T^{0}_{g}, T^{SU}_{g,s+1}) \tag{6}$$

Regarding shut-downs costs, its modeling is widely extended as a fixed cost:

$$c^{SD}_{g,t} = w_{g,t} C^{SD}_{g} \qquad \forall g, t \tag{7}$$

3.2.3. Diverting Target Production Costs

Diverting costs are not a real cost, but the way to model the possible transfer of production targets between the thermal groups, as a result of making the problem feasible (moving production targets that are below the minimum power outputs) or optimizing profits (moving productions targets from a group that would start for a few hours to one that is committed).

$$c^{DIV}_{g} = (f^{A}_{g} + f^{B}_{g}) \frac{C^{F}}{2} \qquad \forall g \tag{8}$$

Note that the definition of the diverting strategic term C^F will affect the optimization behavior, from a high C^F where the production targets of each unit are respected, to a C^F equal to zero where only the total production target A^T is respected. These events are detailed in Section 3.2.4 and analyzed in the case study proposed in Section 4.2.

3.2.4. Production Adjustment Equations

The post-processing method implements two balance equations in order to represent the strategic management of the expected production targets:

$$A^{T} = \sum_{g \in G} \sum_{t \in T} p_{g,t} \tag{9}$$

$$A_{g} = \sum_{t \in T} p_{g,t} + f^{A}_{g} - f^{B}_{g} \qquad \forall g \tag{10}$$

$$0 \leq f^{A}_{g} \qquad \forall g \tag{11}$$

$$0 \leq f^{B}_{g} \qquad \forall g \tag{12}$$

Equation (9) always respects the total aggregated production target, A^T, of the set of thermal units considered in the post-processing problem.

Additionally, Equation (10) make it possible to overcome unfeasible data for the production targets of each thermal unit, A_g, like production targets which are below the minimum power output as the result of using relaxed variables in the medium-term fundamental model. In turn, this equation also gives versatility to the profit optimization. The considered time span will depend on the reliability of the productions obtained with the medium-term model, frequently achieving solid values when the aggregation exceeds one week.

Variables f^A_g and f^B_g distribute production targets among the set of thermal units, to a greater or lesser extent depending on the value of C^F, which penalizes transfers in the objective function. If high

values are assigned to C^F, the model will always try to respect the objective production of each unit A_g, relocating productions with the sole purpose of avoiding infeasibilities. On the other hand, if C^F is set to zero, transfers are free and the operation of the portfolio increases its flexibility. In this case, A_g is ignored and the greatest possible profit, considering A^T, is obtained after the optimization.

This parameter opens the door to the analysis of different situations and strategic behaviors in the management of a thermal portfolio belonging to a generation company. The allocation of moderate values would allow from transfer few MWh to avoid that a thermal unit stretches its production at some hours that are not so profitable, to the possibility of preventing a unit from starting to be working for only one hour or similar.

The introduction of this strategic term refers to the internalization of some operation, logistic and opportunity costs that can not be considered otherwise. In that way, the post-processing method will naturally avoid the inefficiencies mentioned above, as well as return a feasible thermal scheduling. All of these events are analyzed in the case studies proposed in Section 4.

3.2.5. Basic Operating Constraints

These equations determine the chronological relationship between the hourly periods, defining the logic between commitments and startups/shutdowns throughout the time span:

$$v_{g,t} - w_{g,t} = u_{g,t} - u_{g,t-1} \qquad \forall g, t \in [2, T] \tag{13}$$

$$v_{g,t} - w_{g,t} = u_{g,t} - U_g^0 \qquad \forall g, t \in [1, 2) \tag{14}$$

$$u_{g,t} \in \{0, 1\} \tag{15}$$

$$v_{g,t} \in \{0, 1\} \tag{16}$$

$$w_{g,t} \in \{0, 1\} \tag{17}$$

Note that it is not necessary to formulate $w_{g,t}$ as binary variable. Its behavior is defined by differences between binaries and the only values that can be taken are 0 or 1.

Finally, the operating constraints of the thermal units are included, limiting their hourly production between its minimum and maximum power outputs when they are committed.

$$u_{g,t} P_g^{MIN} \leq p_{g,t} \qquad \forall g, t \tag{18}$$

$$p_{g,t} \leq u_{g,t} P_g^{MAX} \qquad \forall g, t \tag{19}$$

4. Case Study and Results

In order to show the usefulness of the methodology presented in this paper, a real size case study of an agent that wishes to make a feasible scheduling and optimize the production of four CCGTs in a medium-term horizon is analyzed. Section 4.1 describes the case study, as well as the origin of the inputs needed in the post-processing method. Section 4.2 shows the results of the application of this methodology. This section also presents three cases where the value of the diverting penalty is analyzed, representing a combined profit optimization and feasible scheduling post-processing.

4.1. Presentation of the Case Study and Its Medium-Term Fundamental Model

In this case study, the production of four CCGTs belonging to a generation company operating in the Iberian electricity market (MIBEL) will be made feasible and optimized. In the first phase of the methodology, a medium-term model is run. This model follows the mathematical formulation proposed in [22], representing the equilibrium between markets players through conjectural variations. Its validity to determine the market equilibrium as an optimization problem is also proved in [22], where it is demonstrated that if the cost function is convex and there are non-negative conjectures, the optimization problem is equivalent to an equilibrium problem.

This formulation was summarized in [53] as follows. The competitive behavior in the market is represented as an oligopoly, where the conjectured-price, $\theta_{i,a,p}$, of each market player i is considered as known. The function that relate the production cost for each player and area a, $C_{i,a}$, with its electricity generation, $q_{i,a,p}$, during the period p, is linear or quadratic. The electricity price, $\lambda_{a,p}$, is determined as the dual variable of the power balance constraint, which matches the total energy output with the demand, $D_{a,p}$. Finally, technical constraints, $\mathcal{H}$, are shorten in a generic formulation:

$$\min_{q_{i,a,p}} \sum_{i,a,p} \left(C_{i,a} q_{i,a,p} + \theta_{i,a,p} \frac{q_{i,a,p}^2}{2} \right) \tag{20}$$

subject to:

$$\sum_i q_{i,a,p} = D_{a,p} : \lambda_{a,p} \qquad \forall a, p \tag{21}$$

$$\mathcal{H}(q_{i,a,p}) \geq 0 \qquad \forall i, a, p \tag{22}$$

The main technical constraints applied to the thermal units are those related to the commitment status, maximum and minimum power outputs, operational costs, start-up and shut-down costs, maximum number of start-ups within a period, unplanned unavailability and maintenance schedules. Regarding river basins, the turbine and pumping capacities are modelled, as well as efficiency, storage capacity, inflows, topology and the upper and lower water bounds to guarantee a safety operation.

This model is used in an accurate representation of the Spanish, Portuguese and French electricity markets and its interconnections. Every single thermal unit is considered, as well as hydro reservoirs and the non-dispatchable renewable energy sources. The horizon comprises three years on an hourly basis. After the market clearing determination, the model also checks technical issues (such as network constraints), affairs as the Transmission System Operator does, committing some thermal units to guarantee the stability of the grid, if required.

In order to obtain reasonable run times, integer programming relaxation is applied and the time aggregation technique of system states [49] is used. This clustering process take into account different conditions of the power system and aggregates hours according to its corresponding thermal gap. The transition between clusters is considered with this method, but the equations to keep chronology are not included. In this way, 940 time steps represent the whole time span of three years.

The combination of these modeling techniques leads to very acceptable run times. On the other hand, it presents the drawback of being possible to obtain technically infeasible results. Besides, the representation of detailed thermal costs (as it is the case of start-up costs) is simplified with respect to the proposed post-processing methodology. Table 2 shows a comparison between a three-year case of the described fundamental model, with the post-processing method proposed in Section 3, applied to a four thermal-unit case through a 31-day time span on an hourly basis. The cases analyzed in this paper have been run in a computer Intel Xeon CPU E5-2660 v3 @2.60 GHz with 40 logical processors and 144 GB of installed RAM memory running 64-bit Windows Server 2012 R2, solved with the commercial solver CPLEX 12.10 [58] under GAMS [59].

Regarding the representation of uncertainty, a Monte Carlo simulation has been carried out. A total of 300 cases have been considered to represent different scenarios for the following risk factors: Power demand, hydro conditions, wind generation, solar generation, coal prices, natural gas prices, CO_2 emissions allowance prices and unplanned unavailability of thermal units.

Given the great variety of risk variables considered in the simulation, the spatial interpolation technique proposed in [60] has been applied, making it possible to obtain a high accuracy in the results with only 300 cases evaluated. Furthermore, the determination of correlations between variables and the scenario creation has been carried out in collaboration with a major utility present in the Spanish electricity market.

This Monte Carlo simulation, carried out with the medium-term model described above, has been used to obtain the necessary input data for the proposed methodology. The corresponding results needed in the post-processing method are shown in Table 3 and Figure 2, which gather the expected thermal productions and the electricity prices, respectively, for the four CCGTs considered in this case study along the 31 day-hourly time span. However, the real scope of this Monte Carlo simulation is longer.

Table 2. Problem sizes after the performance of CPLEX presolve.

Problem Size	Medium-Term Model	Post-Processing Case Study
# of constraints	718,949	17,195
# of cont. variables	1,123,672	11,246
# of binary variables	-	8928
# of non-zero elements	3,110,273	327,674
Run time (s)	∼2000	∼20

It is also important to mention that these outputs correspond to the mean value of the distribution function. Nevertheless, either mean values or those results that belong to any of the contemplated centiles (P10, P50, P90, etc.) can be easily handled with the post-processing methodology proposed in this paper.

Table 3. Expected production of the thermal units considered in the post-processing case study.

Thermal Unit	Productions (MWh)
Unit A	95,358
Unit B	130,635
Unit C	414,360
Unit D	190

Technical data of the thermal units considered in the post-processing are shown in Tables 4 and 5. Start-up costs of the thermal units are modelled with three steps, as mentioned in [1,3]. This representation improves the unique start-up cost of the medium-term model. In turn, the formulation described in Section 3.2.2, easily allows an increment of steps if a more accurate simulation is desired.

Table 4. Technical data of the thermal units and status in the first hour of the considered time span.

Thermal Unit	C_g^{NL} ($/h)	C_g^{LV} ($/MWh)	C_g^{QC} ($/MWh^2)	C_g^{SD} ($)	P_g^{MAX} (MW)	P_g^{MIN} (MW)	T_g^0 (h)	U_g^0
Unit A	1500	33	0.00050	5500	412	157	1	0
Unit B	2300	31	0.00056	5500	390	135	1	0
Unit C	4100	27	0.00027	9500	856	285	1	0
Unit D	1600	32	0.00053	5500	402	112	1	0

Table 5. Start-up parameters of the thermal units.

Thermal Unit	$C_{g,hot}^{SU}$ ($)	$C_{g,warm}^{SU}$ ($)	$C_{g,cold}^{SU}$ ($)	$T_{g,hot}^{SD}$ (h)	$T_{g,warm}^{SU}$ (h)	$T_{g,cold}^{SU}$ (h)
Unit A	15,000	23,000	24,500	1	6	32
Unit B	11,500	25,000	28,000	1	53	245
Unit C	28,000	37,000	43,500	1	21	75
Unit D	12,000	16,000	18,000	1	23	120

4.2. Analysis of Feasible Schedules Obtained with the Post-Processing Methodology

The application of the post-processing method after the performance of a medium-term model, like the one described in Section 4.1, offers many advantages. It achieves a feasible thermal scheduling, keeping reliable and quite valuable medium-term information, such as the hydraulic generation.

In turn, it also improves the representation of the technical operational constraints, since the medium-term model only uses a single-step start-up cost. This phase is more computationally flexible, being possible to approach the start-up cost to a multi-step function, which provides a much more accurate modeling.

Additionally, it is possible to introduce a strategic term to consider hidden operation preferences, allowing a more realistic management of these assets by a market player. These operational priorities can be easily quantified in economic terms and give the chance of transferring some production targets between the thermal units of the portfolio. In this section, three case studies will be considered to analyze the impact of the strategic term, C^F, on the thermal scheduling:

- The first case assigns a value of 500 \$/MWh to C^F. In this case, this number is high enough to avoid divertion of production targets between thermal units. Table 3 is respected without any flexibility. This situation would only allow non-zero values of f_g^A and f_g^B when there are infeasible production targets, like operations below the minimum power output.
- The second case applies a value of 100 \$/MWh to C^F. In this case, the strategic term avoids non-expected operational behaviors, such as the commitment of a thermal unit for operating during a single hour.
- The third case shows a global optimization of the total production target of the portfolio. The assignment of 0 \$/MWh to C^F allows the optimal distribution of A^T in order to maximize the profit.

It is important to keep in mind that the strategic divertion term behavior will depend on the hourly price forecast, because its higher or lower levels would promote or damp the transfers in the objective function. This strategic term can be easily assigned by generation companies, which know in depth its operation, logistic, and opportunity costs. In turn, the company can also play with this value to analyze different situations and risk scenarios.

Table 6 shows the outputs of the performance of the three cases. The comparison is carried through the obtained profits. As expected, the post-processing method reaches a greater profit when there is a higher availability to transfer productions between units, moving them to the most profitable hours and avoiding useless start-ups and the imposition of quantities that are far from the optimum values for this case. The gap established for the three cases is 1%. This value is accurate enough, but it is important to note that there is a trade-off between the desired accuracy and the run time and computational resources. Thus, the higher the number of thermal units involved in the post-processing methodology is, the higher the run time to reach an integer solution will be.

Table 6. Results of the evaluation of the three cases.

Output	Case 1	Case 2	Case 3
Unit A (MWh)	95,358	95,548	101,228
Unit B (MWh)	130,635	130,635	51,440
Unit C (MWh)	414,360	414,360	370,102
Unit D (MWh)	190	0	117,733
Run time (s)	17.8	19.7	9.4
Profits (\$)	8,896,632	8,913,115	9,163,300

In Case 1, every single production target is respected according to the high value assigned to C^F. On the contrary, the reduced production target of Unit D is quickly transferred to other units when C^F is relaxed in Case 2. The optimum value is reached moving 190 MWh from Unit D to Unit A. Finally, the total optimization of Case 3 shows that the optimal solution of the problem is to use each thermal unit along its most profitable hours. In this case, 79,155 MWh and 44,258 MWh are yielded by Unit B and Unit C, being 5870 MWh assumed by Unit A and 117,543 MWh by Unit D.

All of these cases provide a real picture of the detailed operation of the thermal portfolio. As it was expected, its production responds to the electricity price peaks dynamics, considering an optimal

management of start-ups and shut-downs to maximize profit. Figure 2 shows an example of feasible scheduling, where Unit A maximizes its benefit according to the expected production gathered in Table 3. The thermal schedule represented in this figure corresponds to the results of Case 1, where the strategic term C^F is high enough to avoid transfers of production targets.

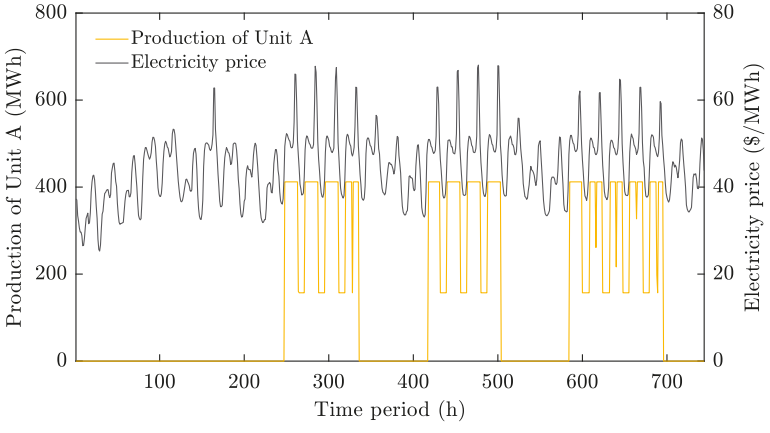


Figure 2. Hourly matching prices of the Spanish market obtained with the fundamental model of Section 4.1 and response of Unit A to its production target with a strategic term of 500 $/MWh as used in Case 1.

For the sake of simplicity, only one thermal unit has been included in Figure 2, allowing an easier interpretation of the operational behavior. Unit A, as well as the other thermal units, sets its production at its maximum power output during the most profitable hours. In addition, it usually reduces its production to the minimum output when prices drop, incurring in shut-downs if electricity peaks are too separated in time.

5. Conclusions

The changing reality of current electricity markets highlights the importance of a proper representation of power systems. The increasing variability of generation, due to the deployment of non-dispatchable RES, and the interconnection promotions between areas as a result of integration policies, points out the necessity of simulating multi-area power system on an hourly basis. Nowadays, electricity market models require a high level of detail and time granularity not only in the short term, but also in the medium term, especially in order to represent a real management of energy storage facilities. This fact, together with the uncertainty consideration in some input data, makes it imperative to simplify medium-term models to increase their computational tractability. However, these simplifications lead to infeasible and/or suboptimal operational outputs for thermal units.

A new soft-linking methodology to overcome these problems has been proposed in this paper. This method combines the advantages of medium-term models, selecting high reliable results from these tools, and using them in a post-processing phase. This step provides a feasible thermal scheduling for a well-detailed generation portfolio. In this way, the infeasible outputs obtained with the medium-term model as a consequence of the implemented simplifications, such as time aggregation, integer programming relaxation or less accurate modeling, are corrected.

In addition, the methodology allows the use of a strategic term, providing an alternative to jointly optimize the thermal generation portfolio of a market agent. This term improves the single representation of technical constraints, allowing the assignment of logistic and opportunity costs, as well as the inclusion of hidden flexibility possibilities in the operation of thermal units.

The formulation of the post-processing phase as an optimization problem also contributes to recreating the competition in power markets, where each player tries to maximize its profit. The whole methodology has been tested with a realistic case study, showing its validity.

Author Contributions: Conceptualization, L.M., A.B. and J.R.; methodology, L.M., A.B. and J.R.; software, L.M.; validation, L.M., A.B. and J.R.; formal analysis, L.M.; investigation, L.M.; resources, L.M.; data curation, L.M.; writing—original draft preparation, L.M.; writing—review and editing, L.M., A.B. and J.R.; visualization, L.M.; supervision, A.B. and J.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CCGT	Combined Cycle Gas Turbine
CV	Conjectural Variation
IP	Integer Programming
MIBEL	Iberian Electricity Market
MILP	Mixed Integer Linear Programming
MIQCP	Mixed Integer Quadratically Constrained Programming
RES	Renewable Energy Sources
UC	Unit Commitment

Nomenclature

Indexes & Sets

$g \in G$	Set of indexes of generating units
$s \in S$	Set of indexes of start-up segments
$t \in T$	Set of indexes of hourly periods of the time span

Parameters

A_g	Target production of an individual unit g throughout the time span T (MWh)
A^T	Total target production of the portfolio G throughout the time span T (MWh)
C^F	Strategic term for diverting target production between thermal units (\$/MWh)
C_g^{LV}	Linear variable cost of unit g (\$/MWh)
C_g^{NL}	Fixed cost of unit g (\$/h)
C_g^{QC}	Quadratic variable cost of unit g (\$/MWh^2)
C_g^{SD}	Shut-down cost of unit g (\$)
$C_{g,s}^{SU}$	Start-up cost for the start-up type s of unit g (\$)
L_t	Price of energy in period t (\$/MWh)
P_g^{MAX}	Maximum power output of unit g (MW)
P_g^{MIN}	Minimum power output of unit g (MW)
$T_{g,s}^{SU}$	Minimum time period that unit g must be offline for the start-up type s (h)
T_g^0	Hourly periods that unit g has been offline in the first period t of the time span T (h)
U_g^0	Commitment status of unit g in the first period t of the time span T

Variables

c_g^{DIV}	Diverting target production cost of unit g along the time span T (\$)
$c_{g,t}^{PROD}$	Production cost of unit g in period t (\$)
$c_{g,t}^{SD}$	Shut-down cost of unit g in period t (\$)
$c_{g,t}^{SU}$	Start-up cost of unit g in period t (\$)
$\delta_{g,s,t}$	Start-up type s of unit g in period t
f_g^A	Increase in target production of unit g due to divertions along the time span T (MWh)
f_g^B	Decrease in target production of unit g due to divertions along the time span T (MWh)
$p_{g,t}$	Power output above the minimum output of unit g in period t (MW)

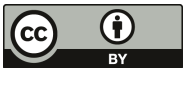
$u_{g,t}$	Commitment decision of unit g in period t
$v_{g,t}$	Start-up decision of unit g in period t
$w_{g,t}$	Shut-down decision of unit g in period t

References

1. Gonzalez-Salazar, M.A.; Kirsten, T.; Prchlik, L. Review of the operational flexibility and emissions of gas-and coal-fired power plants in a future with growing renewables. *Renew. Sustain. Energy Rev.* **2018**, *82*, 1497–1513. [CrossRef]
2. Ahmadi-Khatir, A.; Conejo, A.J.; Cherkaoui, R. Multi-area unit scheduling and reserve allocation under wind power uncertainty. *IEEE Trans. Power Syst.* **2013**, *29*, 1701–1710. [CrossRef]
3. Prina, M.G.; Fanali, L.; Manzolini, G.; Moser, D.; Sparber, W. Incorporating combined cycle gas turbine flexibility constraints and additional costs into the EPLANopt model: The Italian case study. *Energy* **2018**, *160*, 33–43. [CrossRef]
4. Peng, C.; Lei, S.; Hou, Y.; Wu, F. Uncertainty management in power system operation. *CSEE J. Power Energy Syst.* **2015**, *1*, 28–35. [CrossRef]
5. Reneses, J. Analisis de la Operacion de los Mercados de Generacion de Energia Electrica a Medio Plazo. Ph.D. Thesis, Comillas Pontifical University, Madrid, Spain, 2004.
6. Martinez, F.; Villar, J. Profitability Analysis of Spanish CCGTs under Future Scenarios of high RES and EV Penetration. In Proceedings of the 2019 16th International Conference on the European Energy Market (EEM), Ljubljana, Slovenia, 18–20 September, 2019, pp. 1–5.
7. Arroyo, J.M.; Conejo, A.J. Optimal response of a thermal unit to an electricity spot market. *IEEE Trans. Power Syst.* **2000**, *15*, 1098–1104. [CrossRef]
8. Nowak, M.P.; Römisch, W. Stochastic Lagrangian relaxation applied to power scheduling in a hydro-thermal system under uncertainty. *Ann. Oper. Res.* **2000**, *100*, 251–272. [CrossRef]
9. Arroyo, J.M.; Conejo, A.J. Modeling of start-up and shut-down power trajectories of thermal units. *IEEE Trans. Power Syst.* **2004**, *19*, 1562–1568. [CrossRef]
10. Carrión, M.; Arroyo, J.M. A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem. *IEEE Trans. Power Syst.* **2006**, *21*, 1371–1378. [CrossRef]
11. Ostrowski, J.; Anjos, M.F.; Vannelli, A. Tight mixed integer linear programming formulations for the unit commitment problem. *IEEE Trans. Power Syst.* **2011**, *27*, 39–46. [CrossRef]
12. Morales-España, G.; Latorre, J.M.; Ramos, A. Tight and compact MILP formulation for the thermal unit commitment problem. *IEEE Trans. Power Syst.* **2013**, *28*, 4897–4908. [CrossRef]
13. Knueven, B.; Ostrowski, J.; Watson, J.P. A novel matching formulation for startup costs in unit commitment. *Optim. Online* **2017**. [CrossRef]
14. Nash, J.F. Equilibrium points in n-person games. *Proc. Nat. Acad. Sci. USA* **1950**, *36*, 48–49. [CrossRef]
15. Ferrero, R.; Shahidehpour, S.; Ramesh, V. Transaction analysis in deregulated power systems using game theory. *IEEE Trans. Power Syst.* **1997**, *12*, 1340–1347. [CrossRef]
16. Hobbs, B.F.; Helman, U. Complementarity-based equilibrium modeling for electric power markets. In *Modeling Prices in Competitive Electricity Markets*; Bunn, D.W., Ed.; Wiley: New York, NY, USA, 2004.
17. Gabriel, S.A.; Siddiqui, S.A.; Conejo, A.J.; Ruiz, C. Solving discretely constrained Nash–Cournot games with an application to power markets. *Netw. Spat. Econ.* **2013**, *13*, 307–326. [CrossRef]
18. Tesser, M.; Pages, A.; Nabona, N. An oligopoly model for medium-term power planning in a liberalized electricity market. *IEEE Trans. Power Syst.* **2008**, *24*, 67–77. [CrossRef]
19. Scott, T.J.; Read, E.G. Modelling hydro reservoir operation in a deregulated electricity market. *Int. Trans. Oper. Res.* **1996**, *3*, 243–253. [CrossRef]
20. Shrestha, G.; Pokharel, B.K.; Lie, T.T.; Fleten, S.E. Medium term power planning with bilateral contracts. *IEEE Trans. Power Syst.* **2005**, *20*, 627–633. [CrossRef]
21. Barforoushi, T.; Moghaddam, M.P.; Javidi, M.; Sheik-El-Eslami, M. A new model considering uncertainties for power market. *Iranian J. Elect. Electron. Eng.* **2006**, *2*, 71–81.
22. Barquin, J.; Centeno, E.; Reneses, J. Medium-term generation programming in competitive environments: A new optimisation approach for market equilibrium computing. *IEE Proc. Gener. Transm. Distrib.* **2004**, *151*, 119–126. [CrossRef]

23. Helistö, N.; Kiviluoma, J.; Ikäheimo, J.; Rasku, T.; Rinne, E.; O'Dwyer, C.; Li, R.; Flynn, D. Backbone—An Adaptable Energy Systems Modelling Framework. *Energies* **2019**, *12*, 3388. [CrossRef]
24. Ravn, H.F.; Munksgaard, J.; Ramskov, J.; Grohnheit, P.; Larsen, H. *Balmorel: A Model for Analyses of the Electricity and CHP Markets in the Baltic Sea Region*; Technical Report; Balmorel Project: Denmark, 2001. Available online: http://www.balmorel.com/index.php/balmorel-documentation (accessed on 17 March 2012).
25. Calliope: A Multi-Scale Energy Systems (MUSES) Modeling Framework. Available online: https://calliope.readthedocs.io/en/v0.5.3/ (accessed on 17 March 2020).
26. Hobbs, B.F.; Rijkers, F.A. Strategic generation with conjectured transmission price responses in a mixed transmission pricing system-Part I: formulation. *IEEE Trans. Power Syst.* **2004**, *19*, 707–717. [CrossRef]
27. Zerrahn, A.; Schill, W.P. Long-run power storage requirements for high shares of renewables: Review and a new model. *Renew. Sustain. Energy Rev.* **2017**, *79*, 1518–1534. [CrossRef]
28. Hirth, L. The European Electricity Market Model EMMA Model Documentation. Available online: https://neon-energie.de/emma-documentation.pdf (accessed on 17 March 2020).
29. Limpens, G.; Moret, S.; Jeanmart, H.; Maréchal, F. EnergyScope TD: A novel open-source model for regional energy systems. *Appl. Energy* **2019**, *255*, 113729. [CrossRef]
30. Heuberger, C.F.; Rubin, E.S.; Staffell, I.; Shah, N.; Mac Dowell, N. Power capacity expansion planning considering endogenous technology cost learning. *Appl. Energy* **2017**, *204*, 831–845. [CrossRef]
31. Atabay, D. An open-source model for optimal design and operation of industrial energy systems. *Energy* **2017**, *121*, 803–821. [CrossRef]
32. Huppmann, D.; Egging, R. Market power, fuel substitution and infrastructure–A large-scale equilibrium model of global energy markets. *Energy* **2014**, *75*, 483–500. [CrossRef]
33. Howells, M.; Rogner, H.; Strachan, N.; Heaps, C.; Huntington, H.; Kypreos, S.; Hughes, A.; Silveira, S.; DeCarolis, J.; Bazillian, M.; et al. OSeMOSYS: The open source energy modeling system: An introduction to its ethos, structure and development. *Energy Policy* **2011**, *39*, 5850–5870. [CrossRef]
34. PLEXOS Market Simulation Software. Available online: https://energyexemplar.com/solutions/plexos/ (accessed on 17 March 2020).
35. Nelson, J.; Johnston, J.; Mileva, A.; Fripp, M.; Hoffman, I.; Petros-Good, A.; Blanco, C.; Kammen, D.M. High-resolution modeling of the western North American power system demonstrates low-cost and low-carbon futures. *Energy Policy* **2012**, *43*, 436–447. [CrossRef]
36. IEA-ETSAP Optimization Modeling Documentation. Available online: https://iea-etsap.org/index.php/documentation (accessed on 17 March 2020).
37. Schaber, K.; Steinke, F.; Hamacher, T. Transmission grid extensions for the integration of variable renewable energies in Europe: Who benefits where? *Energy Policy* **2012**, *43*, 123–135. [CrossRef]
38. Rasku, T.; Kiviluoma, J. A Comparison of Widespread Flexible Residential Electric Heating and Energy Efficiency in a Future Nordic Power System. *Energies* **2018**, *12*, 1–27. [CrossRef]
39. Hirth, L. The benefits of flexibility: The value of wind energy with hydropower. *Appl. Energy* **2016**, *181*, 210–223. [CrossRef]
40. Göransson, L.; Johnsson, F. Cost-optimized allocation of wind power investments: A Nordic–German perspective. *Wind Energy* **2013**, *16*, 587–604. [CrossRef]
41. Schill, W.P.; Zerrahn, A. Long-run power storage requirements for high shares of renewables: Results and sensitivities. *Renew. Sustain. Energy Rev.* **2018**, *83*, 156–171. [CrossRef]
42. Welsch, M.; Deane, P.; Howells, M.; Gallachóir, B.Ó.; Rogan, F.; Bazilian, M.; Rogner, H.H. Incorporating flexibility requirements into long-term energy system models–A case study on high levels of renewable electricity penetration in Ireland. *Appl. Energy* **2014**, *135*, 600–615. [CrossRef]
43. Deane, J.; Driscoll, A.; Gallachóir, B.Ó. Quantifying the impacts of national renewable electricity ambitions using a North–West European electricity market model. *Renew. Energy* **2015**, *80*, 604–609. [CrossRef]
44. Pfenninger, S.; Keirstead, J. Renewables, nuclear, or fossil fuels? Scenarios for Great Britain's power system considering costs, emissions and energy security. *Appl. Energy* **2015**, *152*, 83–93. [CrossRef]
45. Hobbs, B.F.; Rijkers, F.A.; Wals, A.F. Strategic generation with conjectured transmission price responses in a mixed transmission pricing system-part II: Application. *IEEE Trans. Power Syst.* **2004**, *19*, 872–879. [CrossRef]
46. Pina, A.; Silva, C.; Ferrão, P. The impact of demand side management strategies in the penetration of renewable electricity. *Energy* **2012**, *41*, 128–137. [CrossRef]

47. Boiteux, M. La tarification des demandes en pointe: Application de la théorie de la vente au coût marginal. *Rev. Gen. Electr.* **1949**, *58*, 321–340.
48. Caramanis, M.C.; Tabors, R.D.; Nochur, K.S.; Schweppe, F.C. The introduction of nondiispatchable technologies a decision variables in long-term generation expansion models. *IEEE Trans. Power App. Syst.* **1982**, *PAS-101*, 2658–2667. [CrossRef]
49. Wogrin, S.; Dueñas, P.; Delgadillo, A.; Reneses, J. A new approach to model load levels in electric power systems with high renewable penetration. *IEEE Trans. Power Syst.* **2014**, *29*, 2210–2218. [CrossRef]
50. Wogrin, S.; Galbally, D.; Reneses, J. Optimizing storage operations in medium-and long-term power system models. *IEEE Trans. Power Syst.* **2015**, *31*, 3129–3138. [CrossRef]
51. Tejada-Arango, D.A.; Wogrin, S.; Centeno, E. Representation of storage operations in network-constrained optimization models for medium-and long-term operation. *IEEE Trans. Power Syst.* **2017**, *33*, 386–396. [CrossRef]
52. Tejada-Arango, D.A.; Domeshek, M.; Wogrin, S.; Centeno, E. Enhanced representative days and system states modeling for energy storage investment analysis. *IEEE Trans. Power Syst.* **2018**, *33*, 6534–6544. [CrossRef]
53. Bello, A.; Reneses, J.; Muñoz, A.; Delgadillo, A. Probabilistic forecasting of hourly electricity prices in the medium-term using spatial interpolation techniques. *Int. J. Forecast.* **2016**, *32*, 966–980. [CrossRef]
54. Atakan, S.; Lulli, G.; Sen, S. A state transition MIP formulation for the unit commitment problem. *IEEE Trans. Power Syst.* **2017**, *33*, 736–748. [CrossRef]
55. Morales-España, G.; Tejada-Arango, D.A. Modeling the hidden flexibility of clustered unit commitment. *IEEE Trans. Power Syst.* **2019**, *34*, 3294–3296. [CrossRef]
56. Tejada Arango, D.A.; Lumbreras Sancho, S.; Sánchez Martín, P.; Ramos Galán, A. Which unit commitment formulation is best? A comparison framework. *IEEE Trans. Power Syst.* **2019**, accepted for publication. [CrossRef]
57. Bello, A.; Bunn, D.W.; Reneses, J.; Muñoz, A. Medium-term probabilistic forecasting of electricity prices: A hybrid approach. *IEEE Trans. Power Syst.* **2016**, *32*, 334–343. [CrossRef]
58. IBM ILOG CPLEX Optimization Studio V12.10.0 Documentation. Available online: https://www.ibm.com/support/knowledgecenter/SSSA5P_12.10.0/COS_KC_home.html (accessed on 17 March 2020).
59. GAMS User's Guide. Available online: https://www.gams.com/latest/docs/UG_MAIN.html (accessed on 17 March 2020).
60. Duenas, P.; Reneses, J.; Barquin, J. Dealing with multi-factor uncertainty in electricity markets by combining Monte Carlo simulation with spatial interpolation techniques. *IET Gener. Transm. Distrib.* **2011**, *5*, 323–331. [CrossRef]



energies

MDPI

Article

Modeling and Managing Joint Price and Volumetric Risk for Volatile Electricity Portfolios

Johannes Kaufmann [1,†], Philipp Artur Kienscherf [2,3,*,†] and Wolfgang Ketter [2,3]

1 Next Kraftwerke GmbH, 50825 Cologne, Germany; kaufmann@next-kraftwerke.de
2 Faculty of Management, Economics and Social Sciences, University of Cologne, 50923 Cologne, Germany; ketter@wiso.uni-koeln.de
3 Institute of Energy Economics, University of Cologne, 50827 Cologne, Germany
* Correspondence: philipp.kienscherf@wiso.uni-koeln.de
† These authors contributed equally to this work.

Received: 28 January 2020; Accepted: 3 July 2020; Published: 11 July 2020

Abstract: With an increasing share of renewable energy resources participating in electricity markets, there is a growing dependence between renewable power production and clearing prices of spot markets. Modeling this dependence using bivariate analysis can result in underestimation of market risks and adverse effects for stakeholders' risk management. To enable an undistorted risk assessment, we are using a copula approach to precisely and jointly model electricity prices and infeed volumes of wind power. We simulate the case of wind farm operators using power purchase agreements (PPAs) to shift the price risk to an energy trader, who integrates the infeed into its portfolio. The trader's portfolio can either be geographically dispersed, or highly localized. Based on its portfolio, the energy trader can decide to use derivatives such as futures to manage its risk exposure. The trader decides on future volumes subject to its portfolio's inherent volatility. With a given risk averse strategy, a sufficiently diverse portfolio can help reduce the necessity to trade futures and subsequently the disadvantage of having to pay potential risk premiums.

Keywords: portfolio; portfolio management; risk; risk assessment; energy trading; power purchase agreements; PPA; copula

1. Introduction

In Germany, as in many other countries, market penetration of volatile renewable electricity producers has reached a significant level. In accordance to federal government and European Union goals, the German power sector is set to increase its share of electricity produced by renewable energy sources (RES) to at least 35% by the end of 2020, at least 65% in 2030, and at least 80% in 2050 [1]. RES in this context are wind, solar, biomass, hydro and niche producers (e.g., geothermal). The share of RES in the gross electricity consumption reached 31.6% in 2016, double the share compared to 2008 [2]. This puts it on track to reach the stated goal. Because of the volatile nature of renewable production, the doubling of produced electricity was accompanied with a bigger increase in production capacity of 287%, corresponding to a share of the total production capacity of 52% [3].

The increase in renewable energy generation is primarily driven by expansion of wind and solar power. This expansion of volatile electricity production has measurable effects on price volatility and dependencies between renewable infeed and prices [4,5]. A principle component analysis (PCA) of price variation shows that seasonal factors, which affect renewable generation, are a major component [6]. A similar approach has been used to assess the role of prices spikes in electricity markets [7]. The volatility caused by RES expansion poses numerous challenges for actors in the energy system. Potential investors in new power plants need their assets to generate enough revenue

to cover fixed costs; policy makers have to ensure that energy demand can (almost) always be satisfied. These challenges can be tackled in numerous ways. New market designs can help to ensure matching of supply and demand [8,9], and advanced algorithmic techniques can be used to automate trading in energy markets [10,11]. Our work falls in the realm of statistical modelling that allows for advanced forecasting in the highly stochastic energy system.

There is a large body of work in statistical modelling of energy systems and markets, respectively. Using a GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model, [4] shows that wind power decreases the average price level, but increases volatility. The same relationship is shown by [12]. This effect is not only present in the German energy system, but has also been demonstrated for the Texan electricity market [13]. Electricity prices further inhibit statistically significant calendar effects [14]. While most renewable energy producers are currently shielded from these market effects by guaranteed infeed tariffs, this system is being phased out gradually in Germany. New plants do not get guaranteed remuneration for their infeed and old plants are dropping out of the compensation scheme. As a consequence of the shift from guaranteed infeed tariffs to market-based remuneration, there is a trend to market-based financing mechanisms for new installations and plants without fixed compensation. One of these are power purchase agreements (PPAs). Here, the production of a specific electricity producer is sold to an energy trading company or directly to a consumer. While there is a large potential for increasing use of PPAs, risk averse energy trading companies have to manage the acquired risk exposure. The underlying drivers to motivate risk aversion are diverse among different actors. Asset owners are typically risk averse because they carry the capital costs for new installations. To securely refinance their investment, they need to hedge against risks from regulation and technical failure [15]. In liberalized energy markets they also have to hedge against market risks. A common aspect of this risk for different actors along the energy value chain is the aforementioned problem of joint price and quantity risk. The seminal paper of [16] describes this problem for farmers wishing to protect themselves from output uncertainty and unknown market prices. Not only is the future production of a volatile (e.g., wind) portfolio unknown, the revenue from this production is also unclear. The adverse relationship of production and prices, i.e., lower prices in situations with high production and vice versa, exposes market actors to a higher risk than the two individual risk factors [17,18]. This also makes it risky to perform a simple volume hedge, where the hedged quantity is the expected production volume. Due to the dependence structure, this strategy would leave the market actor exposed to disregarded risk aspects.

Owners of RES regularly conclude agreements with market access providers, who offer them so PPAs in the form of "fixed-for-fluctuating-agreements", where the owner receives a fixed price for the future production and thus remains solely with the volumetric risk [19]. As the production volume is driven by weather phenomena, it can be assessed by project developers without in-depth knowledge of energy markets. Companies offering "fixed-for-fluctuating-agreements" or power purchase agreements (PPAs) are paying the producer fixed rates, while facing both unknown production volumes and market prices in the future. They are therefore motivated to hedge against both price and volumetric risks using different instruments. As prices and generation are both stochastic and cross-correlated, this is a complex task. A hedging decision which does not take the stochastic relationship of quantities and price into account risks undervaluing the situations with the highest negative impact on revenue.

Financial risks (not only in energy markets) are often quantified by the Value-at-Risk (VaR) metric. It describes the highest possible loss of a return distribution with a $(1-p)$ confidence, where p is the exogenously defined risk level [20]. Typical VaR levels are 5% (e.g., [21]) or 1% (e.g., [22]). An extension is Conditional Value-at-Risk (CVaR), which conditions VaR on information before a specific point in time [23]. A second measure common to risk management is Expected Shortfall (ES). It is the expected value of the Value-at-Risk at the $(1-p)$ confidence level. Expected Shortfall is better suited to conceptualize the risk for fat-tailed return distributions, because it reflects the resulting higher likelihood of extreme values in its value [24].

To optimize its market position, an energy trader has to model the dependence structure of production and prices accounting for its complexity, especially with regard to the joint distribution's tails. This work focuses on modeling this aspect with regard to the wholesale electricity market, as participation of volatile renewable energy sources in other markets (e.g., regulation) is uncommon in Germany. A mathematical tool to do so are copulas (see, e.g., [24]). Copulas disentangle the dependence structure of multiple variables from their marginal distributions. They are regularly used in applications of mathematical finance and economics, but gained interest in energy research in the past years [25,26]. With the combination of fitted marginal models and copulas, market prices and infeed volumes under a joint distribution can be simulated. The resulting values can be used to optimize the hedging decision using different hedging instruments and to reduce the risk an energy trading company faces. The process of the market position optimization is similar to the work of [19], who used the approach in an analysis of the Danish energy market. Using German data, we develop a better understanding of the relationship of wind power infeed and prices in this market. The final risk model assumes futures prices that equal the expected wholesale price, i.e., it assumes efficient (financial) markets. Doing so is common practice (see, e.g., [18,19,21]). This approach enables a focus on the variance effects of different hedging instruments, rather than expected revenue, for dealing with the price risk.

The remainder of this paper is structured as follows. First, data and the distribution function estimation process are presented. Then, the estimated distribution functions are used to bootstrap a simulation of joint infeed and price realizations. Using this simulation, different portfolios are optimized with regard to the remaining variance in revenues. The estimation and simulation procedure can be summarized by the following steps:

1. Apply outlier model to price data;
2. Apply logit-transform to infeed data;
3. Estimate seasonal models;
4. Estimate autoregressive and moving average components and variance terms;
5. Estimate suitable distributions for the standardized residuals;
6. Estimate suitable copula.

With the fitted model, Monte Carlo simulations can be performed:

1. Draw random samples from copula;
2. Re-transform these to price and infeed values for a chosen time-period;
3. Estimate values of hedging instruments for different portfolios;
4. Minimize variance of revenue distribution over different quantities of hedging contracts for different portfolios.

Using this approach, we show that a copula based variance minimizing hedge can reduce Conditional Value-at-Risk (CVaR) of a wind power portfolio significantly and improve with regard to expected shortfall (ES) compared to a simpler volume hedge (based on the expected production). Further, we build a variance reduced portfolio and show that needed hedging volumes are lower for both volume hedge und variance minimizing hedge. Our contribution is thus two-fold. First, we provide a statistical analysis of the complex joint relationship of wind power infeed and electricity prices in the German market. Second, we develop an initial set of tools for risk management of volatile portfolios in electricity markets with high RES penetration.

2. Data

The historical price data and wind infeed (min-max normalized) can be seen in Figure 1. Price data is based on German day-ahead clearing prices of EPEX SPOT (https://www.epexspot.com/en). While there are other marketplaces for electricity, and a large volume of over-the-counter (OTC) trading outside energy exchanges, the day-ahead auction is the exchange-based marketplace with highest liquidity in Germany. The original resolution of our data is EUR/MWh. While seemingly counter-intuitive for the analysis of volatility, it is common practice to aggregate the data to daily values. The main reason is that the day-ahead auction clears for all hourly slots of the next day simultaneously. Because of this, the prices of the day-ahead auction do not constitute a sequential series [27–29].

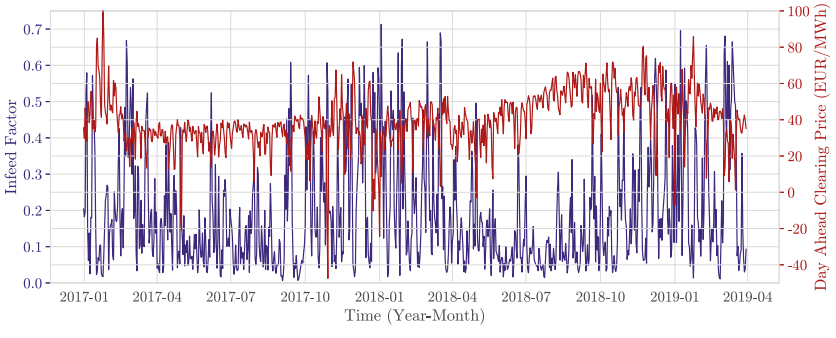
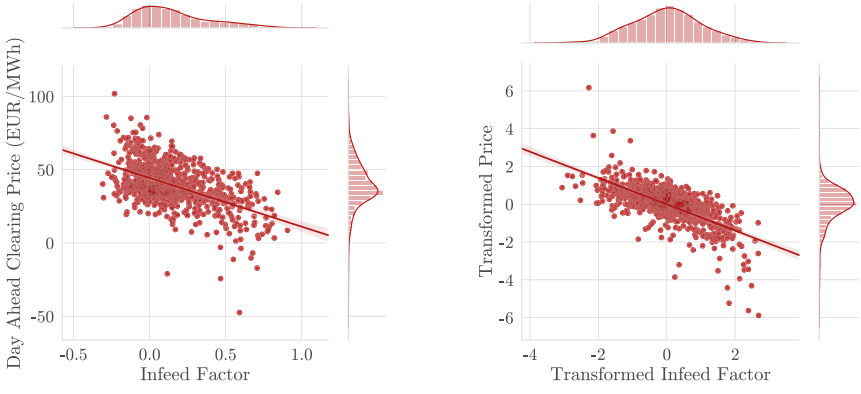


Figure 1. German wind infeed and Day Ahead spot prices from 2017 until spring 2019. Infeed is scaled as a factor of total available capacity. Periods of common high volatility are recognizable.

Part of the wind portfolio of Next Kraftwerke (https://www.next-kraftwerke.com/) constitutes the data source for wind power infeed. The portfolio is preprocessed such that data of 46 wind power plants for a time span of 820 days between January 2017 and March 2019 is available. It has a linear correlation coefficient of 0.9532 with the total German wind power infeed, meaning it is highly representative for the German geographic properties. The infeed is processed using two transformations. As it has an upward trend due to increasing installed capacity, it is standardized as a factor of the total installed capacity. Then, a logistic transformation is applied to the standardized time series. This is due to the fact that boundaries are problematic when modeling the mean and variance models [19]. Before fitting the models, the mean value is subtracted from the time series of wind and prices, i.e., they are centered.

Figure 2 shows the joint distribution of the infeed factor (as average across the portfolio) and the spot prices as well as the joint distribution of the transformed time series. A clear dependency is visible. The descriptive statistics in Table 1 confirm the visual analysis.



(a) **(b)**

Figure 2. Empirical distribution of wind infeed and electricity wholesale day-ahead prices. (**a**) Original infeed and price data; (**b**) Infeed and prices after transformation with the marginal models (see also Section 3).

Table 1. Descriptive statistics for the empirical infeed and price time-series from 2017–2019.

Measure	Infeed Factor	Prices (EUR/MWh)
Mean	0.1545	39.5089
Std. Error	0.1109	15.0836
Spearman's $\hat{\rho}$		-0.4525
Kendall's $\hat{\tau}$		-0.3121
Linear Corr.		-0.5132

Outliers and Seasonality

There are extreme spikes in electricity prices with a variety of methods to correct or replace these values [7,27,30]. Although extreme prices are correct data points, reflecting actual (if rare) economic regimes (cf. [31]), it is reasonable and common practice to treat the data when estimating stochastic models on it. This is because they disproportionately skew the time series [27,30]. This is especially true for hourly data that is more volatile than daily data, and also holds for the aggregated time series.

A simple method to treat extreme values is the fixed price threshold, where the time series is truncated subject to an upper and a lower bound. This method, amongst other similar ones, risks capturing either too few or too many outliers when dealing with data over a long time span. This due to the fact that electricity prices show strong mean variations of several years. An approach to tackle this issue is to remove the trend from the price data using a moving average before applying a filter to the residuals. Because the data not only varies in its mean but also in its variance, a variable price threshold of at least three standard deviations can be used. With some model extensions, this filter can be run iteratively until no more outliers are detected [30]. As the infeed is standardized on a $[0, 1]$-interval, no extreme values are present in the data. The preprocessing and choice of infeed data ensures that sufficiently long time spans with complete infeed data are available for all regarded power plants and no methods for interpolation of results are necessary.

An idiosyncratic aspect of electricity prices is their strong seasonal variation. To account for seasonality, we decompose the electricity price into three distinct components, a short-term and a long-term seasonal component (STSC and LTSC), and a stochastic component X_t. Thus, the random variable P_t^{DA} representing the day-ahead price can be described as follows.

$$P_t^{\mathrm{DA}} = f_t^s + f_t^l + X_t \tag{1}$$

The LTSC f_t^l is defined as a sinusoidal with a yearly period.

$$f_t^l = a_0 + a_1 \cdot \sin\left[2\pi\left(\frac{t}{365.25}\right)\right] + a_2 \cdot \cos\left[2\pi\left(\frac{t}{365.25}\right)\right] \tag{2}$$

Parameters a_1 and a_2 determine phase and amplitude of the sinusoidal, whereas a_0 determines mean. The denominator 365.25 is due to the fact that daily data is used.

The STSC f_t^s, representing daily patterns, is not modeled as a sinusoidal with higher frequency, but using a least squares dummy variable approach. This is due to the fact that daily electricity consumption (and hence, price) patterns do not follow a smooth trend, but are subject to distinct difference between days, e.g., Sunday and Monday. Hence, daily dummy variables are assigned to each day of the week.

$$f_t^s = \sum_{i=1}^{T_W} d_i^d \cdot D_i^d \tag{3}$$

The days of the week are defined by D_i^d, with the D_1^d being Monday. T_W denotes the length of one week in days, i.e., $T_W = 7$. d_i^d denotes the parameter for day i. For instance, $d_1^d = 1$ means that the short-term seasonal component for Mondays equals 1. Note that the random infeed quantity Q_t at time t can be modeled similarly to P_t^{DA}, however, there is no STSC as wind power does not follow a weekly pattern.

3. Estimation Procedure

The preprocessed data is used to estimate both marginal and joint distribution models. The widely used choice for the estimation of the marginal distribution are ARMA-GARCH models, which model the conditional mean and variance of the variables [4,6,19,24]. ARMA models describe stationary stochastic processes through autoregressive and moving-average terms. The autoregressive term uses p lags of the dependent variable and the moving-average term q lags of the error term. The errors ε_t are assumed to be independent and identically distributed (iid) and $\varepsilon_t \sim \mathcal{F}(0, \sigma^2)$, where $\mathcal{F}$ is usually the Normal distribution. The ARMA process for a random variable X_t is

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t, \tag{4}$$

where φ_i and θ_j are the coefficients of the respective lag orders i and j. The GARCH extension replaces the error term with another autoregressive function to account for heteroskedasticity in the errors. It is also defined with a lag order of p and q and can be written as

$$\varepsilon_t = v_t \cdot \sigma_t, \tag{5}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \tag{6}$$

where α_i and β_j are the coefficients of the respective lag orders i and j. The parameters are restricted to ensure stationarity, with $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, and $\sum \alpha_i + \sum \beta_j < 1$ [32]. Usually, $v_t \sim \mathcal{N}(0,1)$ for all t. The normality condition can be relaxed, so that $v_t | \mathcal{F}_{t-1} \sim \mathcal{F}(0,1)$ [21]. This not only permits more general parametric distributions for the error term, its distribution is also conditioned on its past. The conditioning on $\mathcal{F}_{t-1}$ includes past information not only from the variable in question but from all variables. In many cases, when there is no cross-dependency, this can be restricted to the respective variable while still ensuring that all models are conditioned using the same information [24]. While the mean and variance models are coupled through the error term, they can be estimated separately, with

the residuals of the ARMA model serving as input for the GARCH model [27]. This adds modeling flexibility and eases convergence. The lag order of the ARMA and GARCH models can be identified by comparing the Bayesian Information Criteria (BIC) or Akaike Information Criteria (AIC). The more widely used (e.g., [21,24]) BIC is defined as

$$\text{BIC} = -2 \log \mathcal{L} + d \log n,$$

where $\mathcal{L}$ denotes the likelihood function. BIC penalizes model complextiy depending on model size (number of lag parameters) d and sample size n to avoid overfitting [27]. The model with the lowest BIC is considered best. After successfully estimating a model for both mean and variance, standardized residuals can be obtained. These are then used to estimate the copula. Also, a suitable distribution is fitted on the residuals to re-transform the samples obtained from the copula [24].

3.1. Goodness of Fit

The goodness of fit (GoF) of ARMA and GARCH models can be evaluated by plotting the (partial) autocorrelation functions ((P)ACF) of the model's residuals. If the model is fitted well, no significant autocorrelation should remain. This can be tested using the Ljung-Box Q-test of serial independence. The test statistic is given by

$$Q = n(n+2) \sum_{k=1}^{h} \frac{\hat{\rho}_k^2}{n-k}. \tag{7}$$

Here, n is the sample size, $\hat{\rho}_k^2$ the sample autocorrelation at lag k, and h the maximum length for which the test is being performed [33]. Under H_0, the data is independently distributed. Thus, the test should not reject for the mean and variance model. Two widely used tests exist to evaluate the goodness of fit for the distributions which were estimated from the residuals. These are the Kolmogorov-Smirnov (KS) and Cramer-von-Mises (CvM) tests. Both are performed on the values of the residuals' empirical CDF and test the similarity with a known (specified) distribution, where both are the same under H_0:

$$\text{KS}_i = \max_t \left| U_t^f - \hat{U}_t \right| \tag{8}$$

$$\text{CvM}_i = \sum_{t=1}^{T} \left(U_t^f - \hat{U}_t \right)^2 \tag{9}$$

$\hat{U}_t$ is obtained using the empirical CDF and U_t^f using the fitted parametric distribution. As KS and CvM tests are also available to evaluate the GoF of the copula model, subscript i denotes the applicability to the individual models.

3.2. Copula Model

Copulas are used to model the dependence structure of random variables [34]. Whereas, e.g., multivariate normal distributions require all variables and their dependency to have a normal distribution, copulas allow modeling separate marginal distributions of multiple random variables and their dependence. This allows for high flexibility in choosing a suitable distribution and simplifies the estimation procedure, as it can be done in stages [24]. Copulas are common in risk management and econometric applications [35–38]. A d-dimensional copula is a cumulative distribution function on d uniform marginals [39].

$$C : [0,1]^d :\to [0,1].$$

Then, with $C(\mathbf{u}) = C(u_1, \ldots, u_d)$, three properties define a copula: 1) $C(\mathbf{u})$ is non-decreasing in every component u_i. 2) The marginal in component i can be obtained with $u_j = 1$ for all $j \neq i$ because of its uniform distribution $C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$. 3) When $a_i \leq b_i$, $P(U_1 \in [a_1, b_1], \ldots, U_d \in [a_d, b_d])$

is always non-negative. Assuming differentiability of the marginal distributions, the copula can be written as (see, e.g., [24])

$$C(\mathbf{u}) = F(F_1^{-1}(u_1), \ldots, F_d^{-1})). \tag{10}$$

Extensions of copula theory with regard to conditional distributions exist [40] and have been applied to energy modelling [21]. Consider for the bivariate case two random variables $\mathbf{X} \equiv (X_{1,t}, X_{2,t})'$ with a joint conditional distribution function $F(\cdot|\mathcal{F}_{t-1})$ and respective conditional marginal distribution functions $F_i(\cdot|\mathcal{F}_{t-1})$, $i = 1, 2$. Then, a conditional copula $C(\cdot|\mathcal{F}_{t-1})$ with two dimensions exists, such that

$$F((x_1, x_2)|\mathcal{F}_{t-1}) = C(F_1(x_1|\mathcal{F}_{t-1}), F_2(x_2|\mathcal{F}_{t-1})|\mathcal{F}_{t-1}). \tag{11}$$

Note that both the marginal models and the copula are conditioned on the past. If the marginals are continuous, the copula C is unique.

$$\mathbf{U}_t|\mathcal{F}_{t-1} \sim C(\cdot|\mathcal{F}_{t-1}) \tag{12}$$

with $\mathbf{U}_t \equiv (U_{1,t}, U_{2,t})'$. Each $U_{i,t} \sim U[0, 1]$ has the probability integral transform variable

$$U_{i,t} \equiv F_i(X_{i,t}|\mathcal{F}_{t-1}),\ i = 1, 2.$$

The two main families of copulas are called Elliptical and Archimedean. Elliptical copulas are based on elliptical distributions, the two best-known of which are the Gaussian (normal) and Student's t distribution. They are distinct in that the linear correlation fully describes their dependence structure (in contrast to other copula families, where this is false) [39]. In contrast to Elliptical copulas, Archimedean copulas are explicitly defined with so-called generator functions ϕ. They interpolate between dependence structures like independence and comonotonicity, typically using a free parameter θ. The general generator function is continuous and strictly decreasing: $\phi : [0, 1] :\to [0, \infty]$, with $\phi(1) = 0$. In the bivariate case, the copula then has the form

$$C(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2)). \tag{13}$$

Five different copula types have been fitted for the residuals. See Table 2 for their respective formulations. Which copula type is suitable for modeling can be evaluated using measures of dependence and goodness of fit tests.

Table 2. Investigated copula types and mathematical formulations.

Class	Copula	Formulation	Parameters
Elliptical	Gaussian	$C_\rho(u_1, u_2) = \Phi_\Sigma(\Phi^{-1}(u_1), \Phi^{-1}(u_2))$	Correlation ρ, correlation matrix Σ, standard normal CDF Φ
	Student t	$C_{v,\Sigma}(u_1, u_2) = t_{v,\Sigma}(t^{-1}(u_1), t^{-1}(u_2))$	Correlation matrix Σ, t_v the CDF of the one-dimensional t_v distribution with v d.f., $t_{v,\Sigma}$ the CDF of the multivariate $t_{v,\Sigma}$ distribution.
Archimedean	Gumbel	$C_\theta(u_1, u_2) = \exp\left[-((-\ln(u_1))^\theta + (-\ln(u_2))^\theta)^{\frac{1}{\theta}}\right]$	θ
	Clayton	$C_\theta(u_1, u_2) = \left(\max\{u_1^{-\theta} + u_2^{-\theta} - 1, 0\}\right)^{-\frac{1}{\theta}}$	θ
	Frank	$C_\theta(u_1, u_2) = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta u_1} - 1)\cdot(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}\right)$	θ

Goodness of Fit

Model specification and goodness of fit (GoF) tests can be seen as complementary. GoF tests can be limited in their explanatory power and be too weak or too strict to conclude a models suitability.

Model specification tests are a good way to compare different models but do not always help in deciding the validity of a chosen model [24]. For fully parametric models, both the distributions resulting from the marginal models and the copula model are parametric. While this allows to fully specify a log-likelihood for estimation, the commonly used approach is to estimate a model in stages. In that case, the marginal models should not exhibit cross-equation restrictions. For nested models (e.g., comparing a Gaussian copula with a Student's t copula) a likelihood ratio test can be used. An even simpler but very crude method is to rank the model likelihoods (see, e.g., [24]).

The KS and CvM test are two widely used GoF measures to compare an estimated copula with the empirical results. Their statistics adapted to the copula case are

$$\mathrm{KS}_C = \max_t \left| \mathbf{C}(\mathbf{U}_t; \hat{\theta}_T) - \hat{\mathbf{C}}_T(\mathbf{U}_t) \right| \tag{14}$$

$$\mathrm{CvM}_C = \sum_{t=1}^{T} \left\{ \mathbf{C}(\mathbf{U}_t; \hat{\theta}_T) - \hat{\mathbf{C}}_T(\mathbf{U}_t) \right\}^2 \tag{15}$$

and use the empirical copula $\hat{\mathbf{C}}_T$ which is defined as

$$\hat{\mathbf{C}}_T(\mathbf{u}) \equiv \frac{1}{T} \sum_{t=1}^{T} \prod_{i=1}^{n} \mathbf{1}\{\hat{U}_{it} \leq u_i\}. \tag{16}$$

As these tests are based on the empirical copula, they only work for constant, i.e., not time-dependent, copula models [24].

4. Estimation and Simulation Results

4.1. Aggregate Portfolio

4.1.1. Marginal Models

The estimation pipeline for the marginal models and the joint distribution using the copula are now applied to the portfolio of 46 wind power plants. The best fitting model combination is used to bootstrap a Monte Carlo simulation of possible scenarios. Table 3 shows the estimation of the marginal models for both the infeed and prices.

Almost all parameters are significant. The insignificant parameters (marked by italic font) have all p-values under 0.2. Some typical characteristics of electricity markets are visible in the parameters, especially daily patterns. While weekdays (except Monday) have almost identical dummy factors, Saturday and especially Sunday have highly negative dummy parameters, due to the fact that reduced electricity consumption drives prices down.

The infeed data are logit-transformed and de-meaned. Then, a sinusoidal model is applied to account for intra-yearly seasonality. After considering the BIC of prospective lag orders, an ARMA(1,1) model is chosen for the autoregressive process. No significant heteroskedasticity is left in the residuals, which is confirmed with the Ljung-Box test on the squared residuals. Therefore, no GARCH model needs to be estimated. The standardized residuals are fitted to a normal distribution, after comparing the results of the KS and CvM tests to the skew normal distribution. Similar to the infeed, the price data is also de-meaned. Before applying the seasonal models, outliers are filtered with the approach described earlier and a threshold of four standard deviations. Subsequently, a sinusoidal long-term seasonal component and a dummy-based short-term seasonal component with dummies for each weekday are fitted. While the analysis of ACF/PACF plots makes a seasonal model likely, comparing the respective BIC values suggest a simple ARMA(2,2) process. Because the residuals show clear signs of heteroskedasticity, a GARCH(1,1) model is applied to them. A skew Student's t distribution is then fitted to the standardized residuals.

After the marginal models are applied, the resulting standardized residual series exhibit a Spearman's rank correlation of -0.6534. This can be attributed to the marginal models stripping

away independent properties inherent to the two time series. Further, it shows that the correlation of prices and infeed are not (only) due to seasonal effects explaining both price and infeed variations but due to their direct relationship. This result corresponds to the literature [41]. The residuals are transformed to uniformly distributed variables using the empirical CDF. With the resulting residuals, a copula can be estimated.

Table 3. Estimates for model parameters, goodness of fit measures, and distribution of standardized residuals for infeed and price data. *Italic* typeset denotes parameters that are not highly significant. $\hat{d}_1^d$ refers to Monday. The LB test subscripts indicate the lag. For the variance model, the squared residuals are tested.

Modeling	**Parameter Estimates**					
Step	**Logit Infeed Model**			**Price Model**		
Outliers					*Removed 5 outliers*	
LTSC	sinusoidal	$\hat{a}_0$	−1.7147	sinusoidal	$\hat{a}_0$	39.5089
		$\hat{a}_1$	*0.0830*		$\hat{a}_1$	−5.7448
		$\hat{a}_2$	0.6058		$\hat{a}_2$	2.2526
STSC				daily dummies	$\hat{d}_1^d$	1.5066
					$\hat{d}_2^d$	3.3147
					$\hat{d}_3^d$	3.8743
					$\hat{d}_4^d$	3.6809
					$\hat{d}_5^d$	3.5756
					$\hat{d}_6^d$	−5.1595
					$\hat{d}_7^d$	−10.4087
Mean	ARMA(1,1)	$\hat{\varphi}_1$	0.4624	ARMA(2,2)	$\hat{\varphi}_1$	1.35861
					$\hat{\varphi}_2$	−0.36321
		$\hat{\theta}_1$	0.1870		$\hat{\theta}_1$	−0.74176
					$\hat{\theta}_2$	−0.15576
		$\hat{\sigma}^2$	0.6813			
	LB_5	*p*-val.	0.8880	LB_5	*p*-val.	0.0231
	LB_{10}	*p*-val.	0.8003	LB_{10}	*p*-val.	0.0722
Variance				GARCH(1,1)	$\hat{\omega}$	*5.3059*
					$\hat{\alpha}_1$	0.2599
					$\hat{\beta}_1$	0.7293
	LB_5^2	*p*-val.	0.5041	LB_5^2	*p*-val.	0.1131
	LB_{10}^2	*p*-val.	0.4769	LB_{10}^2	*p*-val.	0.3359
Dist.	Normal			Skew Student's t	$\hat{\nu}$	3.7536
					$\hat{\lambda}$	−0.2615
	KS test	*p*-val.	0.4295	KS test	*p*-val.	0.5108
	CvM test	*p*-val.	0.9120	CvM test	*p*-val.	0.8338

4.1.2. Copula Model

Different copula specifications and their estimates are shown in Table 4. The Frank and the Gumbel copula converged to the Independence copula, an unlikely outcome, and are therefore excluded. The best-performing was the Student's t copula, although only marginally better than the Gaussian copula. As was shown in Figure 2b, the residuals of the marginal models for both infeed and prices show close resemblance to Gaussian characteristic. Still, the Student's t copula demonstrates superiority with respect to all relevant GoF measures. It is therefore the most suitable candidate to bootstrap the simulations.

Table 4. Estimates, log likelihoods, BIC and p-values of the KS- and CvM tests for different copula families estimated on standardized residuals. The copula with the lowest BIC is marked bold.

Name	Parameter Estimates		Log $\mathcal{L}$	BIC	GoF Test p-val. KS	CvM
Gaussian	$\hat{\rho}$	−0.6679	−237.1403	7.3958	0.7591	0.7711
Student's t	$\hat{\rho}$	**−0.6816**	**−256.8243**	**7.3957**	**0.9991**	**0.8036**
	$\hat{\upsilon}$	**5.4066**				
Clayton	$\hat{\theta}$	−0.2934	−102.8156	7.3968	0.0000	0.0000

The student's t copula requires symmetric dependency. Therefore, quantile dependence test is carried out (see Figure 3). The test shows a slight asymmetry in the dependency. The chosen copula is retained however, because the comparison of different copula families yields worse results for possible asymmetric copulas. Also, the empirical results are well within the 95% confidence intervals of a bootstrap simulation of the estimated Student's t copula (Performed with $N = 999$).

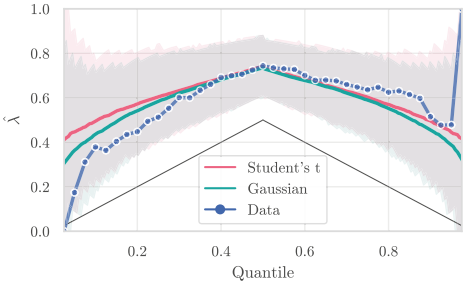


Figure 3. Quantile dependence $\hat{\lambda}$ of the estimated copulas, together with corresponding 95% confidence intervals (depicted as shaded areas).

4.1.3. Simulation and Optimized Hedging

Based on the simulation framework, a routine for determination of the optimal hedging position is developed. In this scenario, the electricity trader's goal is to minimize variance of its returns. As common in power purchase agreements, the trader is obligated to pay a fixed amount $P_{t_0}^{\text{fixed}}$ per unit of energy to the producer, no matter the time of feed-in. With this, the daily revenue R^f of the traders' portfolio can be calculated as

$$R^f(d) = 24c \cdot \sum_{t=h_0}^{h_0+23} Q_t(P_t^{\text{DA}} - P_{t_0}^{\text{fixed}}). \tag{17}$$

Here, Q_t denotes the (stochastic) infeed at time t, with h_0 being the first hour of day d. P_t^{DA} is the stochastic day-ahead price at time t. c is a capacity factor denoting the size of the portfolio to scale the revenue. The optimization routine is, however, scale invariant. Prices are aggregated per day, hence a factor of 24 is included in the expression. Again, this is barely for scaling, but does not affect the optimization procedure. Balancing risk is excluded in our analysis by assuming that the quantities sold on the day ahead market reflect the actual infeed, or $Q_t = \mathbb{E}_{t-1}[Q_t]$. This is done for simplicity and because the issues arising from explicitly modelling balancing risk would call for a detailed explicit consideration and would aggravate assessment of the effect of the hedging position.

Following [19], we are imposing revenue neutral financial instruments. Doing so is common practice (cf. [18,19,21,42]) and enables a focus on the variance effects of hedging instruments in dealing with the price risk. The fair value of $P_{t_0}^{\text{fixed}}$ is then given by

$$P_{t_0}^{\text{fixed}}(d) = \frac{\mathbb{E}_{t_0}^{\mathbb{Q}}\left[\sum_{t=h_0}^{h_0+23} Q_t P_t^{\text{DA}}\right]}{\mathbb{E}_{t_0}^{\mathbb{Q}}\left[\sum_{t=h_0}^{h_0+23} Q_t\right]}. \tag{18}$$

$\mathbb{Q}$ is the risk-neutral measure. Under the rational expectation hypothesis it can be set equal to the physical measure $\mathbb{P}$, which accounts for the uncertainty arising from using historical data for the model (cf. [42,43]). Due to their liquidity, we are focusing on daily futures as hedging instruments. The payoff for one futures contract for day d is

$$R^{\text{Future}}(d) = 24 \cdot \sum_{t=h_0}^{h_0+23} P_t^{\text{DA}} - F_{t_0}(d), \tag{19}$$

with the price of the future at t_0 being $F_{t_0}(d)$ [19]. The price of the future can be defined as the conditional spot price expectation [44,45]. When multiple hedging instruments are used, their payoff can be calculated as a linear combination of the individual contracts, $R^{\text{hedging}}(d) = \sum_{n=1}^{N} \theta_n R^{(n)}(d)$. The total revenue is then

$$R^{\text{total}}(d) = R^f(d) + \sum_{n=1}^{N} \theta_n R^{(n)}(d). \tag{20}$$

Enabling the optimization of the hedging position based on these calculations requires assumptions regarding the financial aspects of the given energy market. Under the rational expectation hypothesis, the expected revenue becomes $E_{t_0}^{\mathbb{Q}}[R^f] = 0$. It is, however, not realistic [46,47], due to incompleteness of the electricity market. Still, it is common practice [17,21,43], therefore we proceed the same way. Further, we are setting the interest rate to zero, allowing for the optimization after simulating from the dependency model (cf. [19]). Because hedging is assumed to take place at time t_0, the optimization is limited to a static hedge (in contrast to a dynamic hedge, where the hedged quantities can be dynamically altered after t_0.) Furthermore, as [18] concede, the problem of timing, i.e., at which t_0 to perform hedging for a contract covering d, is complicated. This renders excluding it from the problem a reasonable option. An example of a model which includes the timing decision can be found in [48].

Since applying all assumptions to calculating prices for hedging instruments means that they are revenue-neutral as well, the optimization problem is reduced to the variance aspect, which can be formulated as

$$\min_{\theta} \text{Var}_{t_0}\left[R^{\text{total}}(d)\right] \tag{21}$$

where n specifies the corresponding hedging instrument out of N different ones. With the corresponding quantities for each hedging instrument and for an arbitrary portfolio size, the CVaR and ES of the minimal variance hedge can be evaluated and compared. We are focusing on 2 exemplary months within our dataset, February and August. Both have different characteristics regarding wind infeed and price behavior. Portfolio size is normalized to a capacity of 100 MW. Table 5 shows the results for the total portfolio.

Table 5. Risk assessment for the overall portfolio with and without hedging. Both hedging methods reduce variance and CVaR significantly, with the minimal variance hedge outperforming the volume hedge with respect to expected shortfall.

Simulation	February	August
Mean price (EUR)	36.88	41.66
Mean infeed factor (%)	24.17	11.29
Mean infeed sum (MWh)	16,242.85	8,399.57
Unhedged case		
5% CVaR (EUR)	−16,615.11	−8,003.63
1% CVaR (EUR)	−45,133.95	−24,039.67
Minimal variance hedge		
Mean hedging quantity (MW)	−646	−323
Variance reduction (%)	96.24	94.56
5% CVaR reduction (%)	46.19	35.11
5% ES reduction (%)	61.68	54.97
1% CVaR reduction (%)	56.45	47.46
1% ES reduction (%)	71.20	66.39
Volume hedge		
Mean hedging quantity (MW)	−580	−270
Variance reduction (%)	57.34	51.75
5% CVaR reduction (%)	53.15	48.15
5% ES reduction (%)	47.52	41.05
1% CVaR reduction (%)	50.88	45.34
1% ES reduction (%)	43.03	5.37

It can be seen that both hedging variants, the simple volume hedge and the variance minimization significantly reduce the CVaR. As is expected, the variance is reduced more strongly for the hedging method defining this as its goal. Interestingly, CVaR is reduced more strongly using the volume hedge. However, expected shortfall is reduced more under the variance minimizing hedging scheme. This means that while a volume hedge reduces the starting point of the revenue distribution's tail more, the mass of the tail is reduced further under the variance minimizing hedge.

4.2. Variance Reduced Portfolio

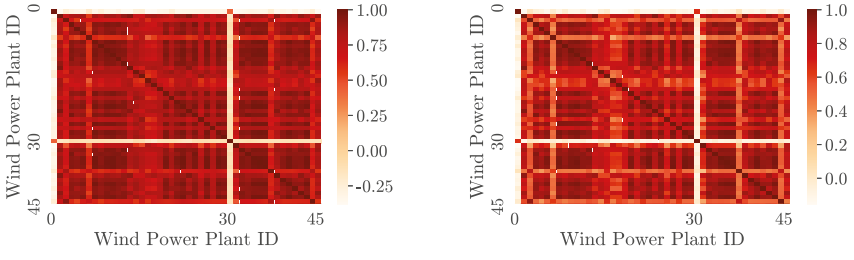
In Section 4, the total, i.e., average, portfolio of NEXT Kraftwerke was subject to the simulation and optimization routine. Now, we are analyzing a portfolio that is constructed based on the goal of reducing cross-correlation of revenue streams of the individual power plants. For this, we use a simple greedy algorithm that picks power plants to add to the portfolio iteratively. It is described in Algorithm 1.

Algorithm 1: Greedy Portfolio Creation.

```
Data: Revenue data of potential wind power plants, portfolio target size
Result: Variance reduced portfolio
Initialize empty portfolio list;
Calculate cross-correlation of revenues;
Add wind power plants with smallest cross-correlation to portfolio;
while Portfolio smaller than target size do
    Calculate cross-correlation of portfolio to remaining wind power plants;
    Add wind power plant with smallest cross-correlation to portfolio;
end
```

The cross-correlation of the infeed and revenue streams from the power plants are depicted in Figure 4. As can be seen, there is a very high correlation between almost all power plants, both with respect to infeed and revenue. Power plants 0 and 30 are clear outliers, with their infeed being practically uncorrelated to the rest of the portfolio. Further, it can be seen that infeed correlation is

much more homogeneous than revenue correlation. This is due to the fact that (total) wind infeed and prices are correlated. In our case study we are analyzing a portfolio that reduces the number of wind power plants from 46 to 10. Doing so, there is a balance between not overemphasizing outliers (such as plants 0 and 30), but also still being able to see a difference from the overall portfolio. The reduced portfolio is representative of a geographically more diversified set of wind power plants. The same estimation and simulation steps as before are applied to the reduced portfolio.



(**a**) (**b**)

Figure 4. Cross-correlation of both infeed and revenues for the portfolio of 46 (indexed 0 to 45) wind power plants. Clear similarities between correlation coefficients can be seen. Wind power plants 0 and 30 are significant outliers in terms of their cross-correlation. (**a**) Infeed correlation of wind power plants; (**b**) Revenue correlation for wind power plants.

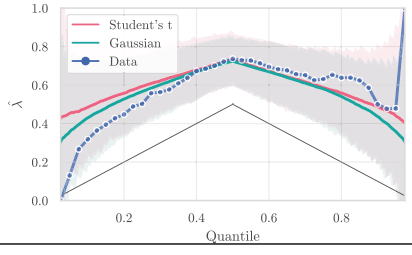
4.2.1. Models

The results of the estimation of the marginal and copula models for the variance reduced portfolio are given in Table 6. As can be seen, both marginal models and copula estimation have significant similarities compared to the overall portfolio. The Student's t copula is still best performing with regard to all defined GoF measures.

Table 6. Fitted marginal model and copula for the variance reduced portfolio together with the quantile dependence plot of the best performing copula. The resemblance to the overall portfolio is uncanny.

Marginal Model			
Model Step	**Parameter Estimates Logit Infeed Model**		
LTSC	sinusoidal	$\hat{a}_0$	-1.7679
		$\hat{a}_1$	*0.0814*
		$\hat{a}_2$	0.5902
Mean	ARMA(1,1)	$\hat{\varphi}_1$	0.4667
		$\hat{\theta}_1$	0.1662
		$\hat{\sigma}^2$	0.6808
	LB_5	p	0.9179
	LB_{10}	p	0.8028
Variance	LB_5^2	p	0.4646
	LB_{10}^2	p	0.4733
Dist.	Std. Normal		
	KS test	p	0.5176
	CvM test	p	0.9731

Copula						
Parameter Name		**Estimates**	**Log $\mathcal{L}$n**	**GoF Test p-val. BIC**	**KS**	**CvM**
Gaussian	$\hat{\rho}$	-0.6493	-219.5089	7.3959	0.7982	0.7804
Student's t	$\hat{\rho}$	$\mathbf{-0.6631}$	$\mathbf{-239.5347}$	**7.3957**	**0.9941**	**0.8192**
	$\hat{v}$	**5.2712**				
Clayton	$\hat{\theta}$	-0.2713	-91.3227	7.3969	0.0000	0.0000



4.2.2. Simulation and Optimized Hedging

Using the estimated marginal model for the portfolio infeed and the Student's t copula the same variance minimization as with the total portfolio is performed. Result of the procedure are

given in Table 7. The general findings are similar to the case with the total portfolio. Both hedging methods reduce variance and CVaR, with the volume hedge yielding a higher reduction in CVaR. Again, the variance minimizing hedge leads to a higher reduction in expected shortfall, i.e., a thinner adverse tail of the revenue distribution. Comparing the hedging volumes with the portfolio in Section 4 shows that smaller hedging volumes are decided upon in the optimal hedging positions (for both cases). In our study, we are resting the calculation of hedging volumes on *fair prices* of futures, i.e., expected spot prices. In reality there are examples of positive risk premiums for longer planning horizons [49], as typical for commodity markets. Reducing the necessity of using financial derivates for any hedging decision reduces the risk of experiencing adverse effects through the payment of risk premiums.

Table 7. Risk assessment for the variance reduced portfolio with and without hedging. Both hedging methods reduce variance and CVaR significantly, with the minimal variance hedge outperforming the volume hedge with respect to expected shortfall. Optimal hedging volumes are reduced compared to the overall portfolio.

Simulation	February	August
Mean price (EUR)	37.08	41.60
Mean infeed factor (%)	22.96	10.81
Mean infeed sum (MWh)	15,428.08	8,041.41
Unhedged case		
5% CVaR (EUR)	−15,870.79	−7,672.37
1% CVaR (EUR)	−41,430.96	−22,549.95
Minimal variance hedge		
Mean hedging quantity (MW)	−611	−308
Variance reduction (%)	90.4	93.12
5% CVaR reduction (%)	48.13	36.5
5% ES reduction (%)	60.27	55.17
1% CVaR reduction (%)	56.39	46.39
1% ES reduction (%)	69.28	67.11
Volume hedge		
Mean hedging quantity (MW)	−551	−259
Variance reduction (%)	66.54	46.3
5% CVaR reduction (%)	54.79	49.39
5% ES reduction (%)	49.36	40.83
1% CVaR reduction (%)	51.59	44.44
1% ES reduction (%)	45.08	34.96

5. Discussion

The key contribution of this paper is the modeling the dependence structure of an actual wind portfolio infeed and German electricity prices with the help of copulas. To enable the estimation, models for cleaning the data of outliers, estimating deterministic seasonal components, and autoregressive models for the mean and variance components of the data are specified. With the standardized residuals of these marginal models, marginal distributions and a suitable copula model are estimated. Following the estimation of marginal models, distributions, and the dependence structure, price-infeed pairs could be simulated. On these values, a model was defined to estimate and optimize the risk arising from the modeled relationship of the variables. This could then be used to minimize the revenue variance by varying the quantity of different hedging products.

In an empirical example, all modeling steps were applied to infeed data from a large German virtual power plant operator and price data from the German market. A yearly seasonal model and an ARMA process was applied to the infeed data, with the residuals conforming to a Normal distribution. The price data was treated using an outlier model, a yearly and a weekly seasonal model and an ARMA-GARCH process.

We show that the revenue variance minimizing hedge using monthly futures contracts strongly reduces the Conditional Value-at-Risk and Expected Shortfall for a market actor facing joint price and volumetric risk. In this respect, the findings are similar to the study by [19] regarding the

Danish market. Additionally, the hedge performs better than a simple volume hedge using the same instrument with regard to expected shortfall. Hence, we conclude that using the minimal revenue variance hedge with monthly futures can significantly reduce the price risk for a volatile electricity producer. Further, we show that a diversified portfolio with low cross-correlation in revenue streams from individual power plants improves risk aspects of the portfolio. Hedging volume can be reduced both with regard to a volume hedge and with regard to the minimum variance hedge. It can be therefore seen that value of an individual power plant does not only depend on the windiness of its location, but also its relationship to the remainder of the portfolio. This is especially true for risk averse decision-makers.

Some limitations remain. The risk model rests on strong assumptions, e.g., enforcing revenue neutrality, not all of which are realistic. Comparing the simulated distributions of both infeed and prices to the empirical ones, there remain differences for the price values. This suggests that there are further price drivers that are unaccounted for in the marginal model (see Appendix A). The empirical example limited itself to only one type of hedging instrument, primarily because illiquid markets preclude an application. Still, accounting for a broader set of derivates, e.g., weather derivates, would enhance the work. An interesting extension of our work is to include (stochastic) risk premiums together with an explicit modeling of the decision-makers risk aversion, in order to develop a decision support system for energy traders seeking to optimize their position.

Despite the limitations, we showed that volatile RES infeed and electricity prices show a complex relationship that is not fully captured by a simple Gaussian model only specifying correlation. Providing an initial method to manage risk subject to this relationship, we are motivating more research on complex risk management in electricity markets with high degree of RES penetration.

6. Materials and Methods

For the technical implementation of the estimation and simulation procedure, the Python programming language and associated statistical software packages are used in conjunction with packages for the R programming language.

Author Contributions: Investigation, J.K.; Supervision, W.K.; Visualization, J.K. and P.A.K.; Writing—original draft, P.A.K.; Writing—review & editing, J.K., P.A.K., and W.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: We would like to thank Next Kraftwerke for providing the data of their wind power portfolio.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

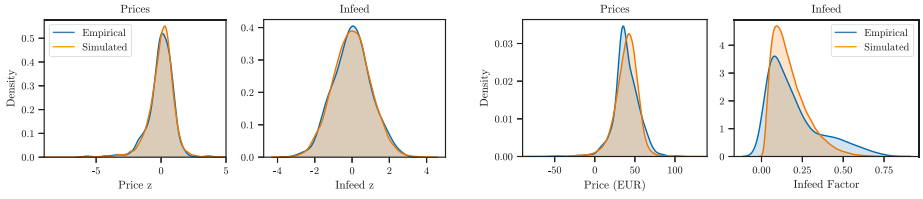
The following abbreviations are used in this manuscript:

AIC	AKAIKE Information Criterion
ARMA	Autoregressive-Moving Average
BIC	Bayesian Information Criterion
CDF	Cumulative Distribution Function
CvM	Cramer-von-Mises
(C)VaR	(Conditional) Value-at-Risk
EPEX SPOT	European Power Exchange
ES	Expected Shortfall
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GoF	Goodness of Fit
iid	independent and identically distributed
KS	Kolmogorov-Smirnov
LTSC	Long-Term Seasonal Component
OTC	Over-the-counter
(P)ACF	(Partial) Autocorrelation Function

PCA	Principal Component Analysis
PPA(s)	Power Purchase Agreement(s)
RES	Renewable Energy Sources
STSC	Short-Term Seasonal Component

Appendix A

Figure A1 shows that the standardized residuals simulated by the copula fit the data well. Their density is barely distinguishable from that defined by a kernel density estimation on the empirical standardized residuals. The deviation is larger when comparing the marginal models with the simulation. Obviously, there are aspects in price and infeed formation that are not accounted for by seasonality and ARMA-GARCH models.



(**a**) (**b**)

Figure A1. Comparison of the empirical and simulated distributions. (**a**) Comparison of standardized residuals of fitted marginal models and results of copula simulation. (**b**) Comparison of empirical distributions and re-transformed results of copula simulation.

References

1. IEA. *Germany 2020*; IEA—International Energy Agency: Paris, France, 2020.
2. Bundesministerium für Wirtschaft und Energie. Sechster Monitoring-Bericht zur Energiewende, Technical Report, 2018. Available online: https://www.bmwi.de/Redaktion/DE/Publikationen/Energie/sechster-monitoring-bericht-zur-energiewende.html (accessed on 9 July 2020).
3. Bundeskartellamt. Monitoringbericht 2018, Technical Report, 2018. Available online: https://www.bundesnetzagentur.de/SharedDocs/Mediathek/Monitoringberichte/Monitoringbericht2018.pdf (accessed on 9 July 2020).
4. Ketterer, J.C. The impact of wind power generation on the electricity price in Germany. *Energy Econ.* **2014**, *44*, 270–280. [CrossRef]
5. Rintamäki, T.; Siddiqui, A.S.; Salo, A. Does renewable energy generation decrease the volatility of electricity prices? An analysis of Denmark and Germany. *Energy Econ.* **2017**, *62*, 270–282. [CrossRef]
6. Mayer, K.; Trück, S. Electricity markets around the world. *J. Commod. Mark.* **2018**, *9*, 77–100. [CrossRef]
7. Li, K.; Cursio, J.D.; Sun, Y. Principal component analysis of price fluctuation in the smart grid electricity market. *Sustainability* **2018**, *10*, 4019. [CrossRef]
8. Bichler, M.; Gupta, A.; Ketter, W. Research commentary—Designing smart markets. *Inf. Syst. Res.* **2010**, *21*, 688–699. [CrossRef]
9. Ketter, W.; Collins, J.; Reddy, P. Power TAC: A competitive economic simulation of the smart grid. *Energy Econ.* **2013**, *39*, 262–270. [CrossRef]
10. Ketter, W.; Peters, M.; Collins, J.; Gupta, A. A multiagent competitive gaming platform to address societal challenges. *Mis Q.* **2016**, *40*, 447–460. [CrossRef]
11. Ketter, W.; Peters, M.; Collins, J.; Gupta, A. Competitive benchmarking: an IS research approach to address wicked problems with big data and analytics. *MIS Q.* **2015**. [CrossRef]
12. Nicolosi, M.; Fürsch, M. The Impact of an increasing share of RES-E on the Conventional Power Market—The Example of Germany. *Zeitschrift für Energiewirtschaft* **2009**, *33*, 246–254. [CrossRef]
13. Woo, C.; Horowitz, I.; Moore, J.; Pacheco, A. The impact of wind generation on the electricity spot-market price level and variance: The Texas experience. *Energy Policy* **2011**, *39*, 3939–3944. [CrossRef]

14. Li, K.; Cursio, J.D.; Jiang, M.; Liang, X. The significance of calendar effects in the electricity market. *Appl. Energy* **2019**, *235*, 487–494. [CrossRef]
15. Mbistrova, A.; Nghiem, A. *The Value of Hedging New Approaches to Managing Wind Energy Resource Risk*; Technical Report; WindEurope: Brussels, Belgium, 2017.
16. McKinnon, R.I. Futures Markets, Buffer Stocks, and Income Stability for Primary Producers. *J. Polit. Econ.* **1967**, *75*, 844–861. [CrossRef]
17. Oum, Y.; Oren, S.S. Optimal Static Hedging of Volumetric Risk in a Competitive Wholesale Electricity Market. *Decis. Anal.* **2010**, *7*, 107–122. [CrossRef]
18. Oum, Y.; Oren, S.; Deng, S. Hedging quantity risks with standard power options in a competitive wholesale electricity market. *Naval Res. Logist.* **2006**, *53*, 697–712. [CrossRef]
19. Pircalabu, A.; Jung, J. A mixed C-vine copula model for hedging price and volumetric risk in wind power trading. *Quant. Financ.* **2017**, *17*, 1583–1600. [CrossRef]
20. Oum, Y.; Oren, S. VaR constrained hedging of fixed price load-following obligations in competitive electricity markets. *Risk Decis. Anal.* **2009**, *1*, 43–56. [CrossRef]
21. Pircalabu, A.; Hvolby, T.; Jung, J.; Høg, E. Joint price and volumetric risk in wind power trading: A copula approach. *Energy Econ.* **2017**, *62*, 139–154. [CrossRef]
22. Hain, M.; Schermeyer, H.; Uhrig-Homburg, M.; Fichtner, W. Managing renewable energy production risk. *J. Bank. Financ.* **2018**, *97*, 1–19. [CrossRef]
23. Bartelj, L.; Gubina, A.F.; Paravan, D.; Golob, R. Risk management in the retail electricity market: The retailer's perspective. In Proceedings of the IEEE PES General Meeting, PES 2010, Minneapolis, MN, USA, 25–29 July 2010.
24. Patton, A.J. Copula Methods for Forecasting Multivariate Time Series. In *Handbook of Economic Forecasting*; Elsevier: Amsterdam, The Netherlands, 2013; Volume 2, pp. 899–960.
25. Chen, X.; Han, J.; Zheng, T.; Zhang, P.; Duan, S.; Miao, S. A Vine-Copula Based Voltage State Assessment with Wind Power Integration. *Energies* **2019**, *12*. [CrossRef]
26. Reboredo, J.C.; Ugolini, A.; Chen, Y. Interdependence Between Renewable-Energy and Low-Carbon Stock Prices. *Energies* **2019**, *12*, 4461. [CrossRef]
27. Weron, R. *Modeling and Forecasting Electricity Loads and Prices*; John Wiley & Sons Ltd.: Oxford, UK, 2006; Volume 33.
28. Meyer-Brandis, T.; Tankov, P. Multi-Factor Jump-Diffusion Models of Electricity Prices. *Int. J. Theor. Appl. Financ.* **2008**, *11*, 503–528. [CrossRef]
29. Manner, H.; Alavi Fard, F.; Pourkhanali, A.; Tafakori, L. Forecasting the joint distribution of Australian electricity prices using dynamic vine copulae. *Energy Econ.* **2019**, *78*, 143–164. [CrossRef]
30. Janczura, J.; Trück, S.; Weron, R.; Wolff, R.C. Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling. *Energy Econ.* **2013**, *38*, 96–110. [CrossRef]
31. Ketter, W.; Collins, J.; Gini, M.; Gupta, A.; Schrater, P. Real-time tactical and strategic sales management for intelligent agents guided by economic regimes. *Inf. Syst. Res.* **2012**, *23*, 1263–1283. [CrossRef]
32. Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *J. Econ.* **1986**, *31*, 307–327. [CrossRef]
33. Ljung, G.M.; Box, G.E.P. On a measure of lack of fit in time series models. *Biometrika* **1978**, *65*, 297–303. [CrossRef]
34. Sklar, A. Fonctions de répartition à n dimensions et leurs marges. *Publ. l'Inst. Stat. L'Univ. Paris* **1959**, *8*, 229–231.
35. McNeil, A.J.; Frey, R.; Embrechts, P. *Quantitative Risk Management: Concepts, Techniques and Tools*; Princeton Series in Finance, Princeton University Press: Oxford, UK, 2005.
36. Manner, H.; Reznikova, O. A survey on time-varying copulas: Specification, simulations, and application. *Econ. Rev.* **2012**, *31*, 654–687. [CrossRef]
37. Embrechts, P.; Lindskog, F.; Mcneil, A. Modelling Dependence with Copulas and Applications to Risk Management. In *Handbook of Heavy Tailed Distributions in Finance*; Elsevier: Amsterdam, The Netherlands, 2003; pp. 329–384.
38. Czado, C.; Schepsmeier, U.; Min, A. Maximum likelihood estimation of mixed C-vines with application to exchange rates. *Stat. Model. Int. J.* **2012**, *12*, 229–255. [CrossRef]
39. Schmidt, T. Coping with Copulas. In *Copulas: From Theory to Application in Finance*; Rank, J., Ed.; Risk Books: London, UK, 2007.

40. Patton, A.J. Modelling asymmetric Exchange Rate Dependence. *Int. Econ. Rev.* **2006**, *47*, 527–556. [CrossRef]
41. Elberg, C.; Hagspiel, S. Spatial dependencies of wind power and interrelations with spot price dynamics. *Eur. J. Oper. Res.* **2015**, *241*, 260–272. [CrossRef]
42. Coulon, M.; Powell, W.B.; Sircar, R. A model for hedging load and price risk in the Texas electricity market. *Energy Econ.* **2013**, *40*, 976–988. [CrossRef]
43. Benth, F.E.; Kettler, P.C. Dynamic copula models for the spark spread. *Quant. Financ.* **2011**, *11*, 407–421. [CrossRef]
44. Benth, F.E.; Meyer-Brandis, T. The information premium for non-storable commodities. *J. Energy Mark.* **2009**, *2*, 111–140. [CrossRef]
45. Benth, F.E.; Jurate, S.B. *Modeling and Pricing in Financial Markets for Weather Derivatives*; World Scientific: Singapore, 2012; Volume 17.
46. Burger, M.; Klar, B.; Müller, A.; Schindlmayr, G. A spot market model for pricing derivatives in electricity markets. *Quant. Financ.* **2004**, *4*, 109–122. [CrossRef]
47. Kolos, S.P.; Ronn, E.I. Estimating the commodity market price of risk for energy prices. *Energy Econ.* **2008**, *30*, 621–641. [CrossRef]
48. Näsäkkälä, E.; Keppo, J. Electricity load pattern hedging with static forward strategies. *Manag. Financ.* **2005**, *31*, 116–137. [CrossRef]
49. Benth, F.E.; Sgarra, C. The risk premium and the Esscher transform in power markets. *Stoch. Anal. Appl.* **2012**, *30*, 20–43. [CrossRef]

energies

MDPI

Article

Evolving Bidding Formats and Pricing Schemes in USA and Europe Day-Ahead Electricity Markets †

Ignacio Herrero [1], Pablo Rodilla [2] and Carlos Batlle [3,4,*]

[1] Citadel LLC, 120 London Wall, London EC2Y 5ET, UK; Ignacio.herrero@citadel.com
[2] Institute for Research in Technology, Comillas Pontifical University, Sta. Cruz de Marcenado 26, 28015 Madrid, Spain; Pablo.Rodilla@comillas.edu
[3] MIT Energy Initiative, 77 Mass. Av., Cambridge, MA 02139, USA
[4] Florence School of Regulation, European University Institute, Via Boccaccio 121, I-50133 Florence, Italy
* Correspondence: CBatlle@mit.edu; Tel.: +1-617-324-4390

† This text solely reflects the analyses and views of the authors. No recipient should interpret this document to represent the general views of the authors' employers or its personnel. Facts, analyses, and views presented herein have not been reviewed by, and may not reflect information known to other professionals of the authors' employers.

Received: 21 June 2020; Accepted: 28 July 2020; Published: 24 September 2020

Abstract: This paper compares the evolution of USA and European power markets and evaluates the suitability and future challenges of their designs in the context of the transition to a low-carbon power system. The analysis focuses on bidding formats (the way in which organized electricity markets allow participants to express their operational constraints) and pricing schemes (how agents recover their short-term costs from market prices). The radical evolution of the power mixes worldwide already experienced in the last decade and the larger one to come, with even greater shares of renewable energy and a more prominent role for storage resources, exposes limitations in current market designs. We develop an in-depth and comprehensive review of best practices from both sides of the Atlantic, and learning from them, we draw recommendations to evolve these market design elements.

Keywords: wholesale electricity markets; market design; bidding formats; pricing rules; renewable energy sources

1. Introduction

In the context of liberalized electric power systems, organized short-term electricity markets (as, for instance, the ones run by power exchanges in the European context (See www.europex.org)) not only help participants manage their risks, but mainly serve as a tool to facilitate an efficient matching of supply and demand, ideally contributing to the goal of maximizing market welfare. While electricity is often defined as a commodity (in the sense that one MWh of electricity is indistinguishable from another), experience has shown that for electricity markets to perform these tasks—aligning the economic utility functions of market agents and the physical constraints conditioning supply, it is more than suitable to allow for some complexity to the bidding and clearing procedures.

In markets for most commodities, only the willingness to buy/sell is relevant, but in the case of electricity, a proper consideration of the physical and economical constraints of agents is instrumental to achieve efficient clearing results. Bidding formats allow agents to express in a complex format their willingness to buy or sell electricity, reflecting how their particular constraints lead to the need to respect quantities and intertemporal links. Two very different approaches have been followed in the USA and Europe as regards to how to design these bidding formats, and in both cases, these formats are experiencing limitations to deal with the new paradigm.

Pricing electricity poses several challenges, mostly derived from the presence of non-convexities (such as non-convex costs). As is well known, in a non-convex context, there may be no linear prices (constant prices that remunerates all quantities) that are able to support a competitive market equilibrium. To deal with this issue, again, USA and Europe have opted for different approaches, and in both cases, these different schemes are being challenged by the penetration of new resources.

Bidding formats and pricing rules are key market design elements to allow for an efficient and massive integration of new resources such as renewables and battery storage systems. The objective of this paper is to develop an in-depth analysis on these elements, including the latest and most up-to-date discussions and challenges in USA ISOs (Independent System Operators) and EU Power Exchanges at the time of this writing. The paper is structured as follows:

In Section 2 we describe how USA and European markets followed very different approaches to the design of bidding formats and pricing rules.

Section 3 describes the context that motivates the evolution of power market design. Power system operation is becoming increasingly complex by the introduction of renewable energy resources, and new market players with new requirements are gradually entering into play, such as storage resources and aggregators. Other studies have discussed the impact of these changes in power markets performance, e.g., Hu et al. [1] and IRENA [2] develop comprehensive reviews on overall system needs, and Anuta et al. [3] focuses on the specific case of storage, but this paper focuses on bidding formats in greater detail and encompasses power system resources with more generality.

Section 4 analyzes performance implications of alternative designs and explores potential improvements for current bidding formats, especially in the European context where more limitations have been identified. We also assess other key elements linked to bidding formats, such as the design of clearing and pricing rules. Section 5 provides final conclusions and recommendations.

2. Bidding Formats and Pricing Rules in USA and the EU: Two Different Approaches

USA- and EU-organized power markets, from their initial implementation, opted for significantly different approaches to design their bidding formats and market clearing rules. The reasons for these differences were diverse, but maybe the most relevant one was the fact that from the very beginning, USA markets run by Independent System Operators (ISOs) were based on a pre-existent integrated structure (the Regional Transmission Operators) who had the responsibility to determine the economic dispatch in a centralized way. Meanwhile, in Europe, the market implementation was focused on prioritizing the economic dispatch (previously, run independently by different utilities) in a single market supported by Power Exchanges (initially, mostly national in scope and mostly non-compulsory). Green [4] and Conejo and Sioshansi [5] develop good descriptions of the fundamental differences of both approaches.

2.1. Markets in the United States: Resource-Specific Bidding Formats

USA markets use multi-part offers, which explicitly reflect generating units' operational and opportunity costs (such as start-up costs), and their technical constraints (e.g., ramp rates). Multi-part offers are clearly motivated by the market clearing approach adopted by ISOs, which is nothing other than the straightforward application of the Security Constrained Unit Commitment and Economic Dispatch optimization models used before the liberalization of the power industry. Table 1 highlights the typical offer parameters that ISOs make available to thermal units (see, for example, exhibit 4–6 at MISO (Midcontinent Independent System Operator) [6]).

Table 1. Typical multi-part offer structure in ISO (Independent System Operator) markets.

Operating Costs		**Technical Constraints**	
Energy offer curve	**MWh, $/MWh**	**Economic min**	**MWh**
Piecewise linear or stepwise linear function with multiple MW/Price pairs		Economic max Ramp rate	MWh MWh/hour
No-load offer	$/hour	Min/Max run time Min down time	hours, min hours, min
Start-up offer	$	Notification time	hours, min
Available for different types of start-ups (hot/intermediate/cold)		Cooling time Start-up time	hours, min hours, min

In some cases, additional parameters allow multi-stage resources to represent their different operating regimes, and transition costs and constraints between modes (i.e., combined cycle gas turbines, which allow multiple configurations of their gas and steam turbines, and therefore have multiple commitment decisions to take). The bidding parameters highlighted here focus on the energy market, but USA markets also optimize operating reserve provision, and other bidding parameters are provided to this end. So-called flexibility products (which are close in nature to reserve products) are also out of the scope of this paper. Jacob [7] presents a good review of current discussions around flexibility products.

The archetypical multi-part offer is the thermal unit bidding format (predominant type in USA systems), but other bidding formats have also been implemented for different types of resources. For instance, in Section 3.2, we describe recent developments to improve bidding formats available to pumped-storage hydro and other storage units. Not all market agents require complex multi-part offers, and it is possible to submit only price-quantity bids, which could be the case for renewable generators and load serving entities.

In summary, ISOs attempt to represent the power system with the highest detail possible in their clearing algorithm, including the technical characteristics of each generator individually, apart from all the constraints required to ensure reliability. This complex model allows ISOs to optimally schedule resources, while enabling competition in the provision of energy and electricity services.

Pricing Approach

These multi-part bidding formats make the clearing problem non-convex, causing well-known challenges in the computation of marginal prices [8–15]. The basic matter of this non-convexity is that the marginal cost of the system may be lower than the average cost of some units. For example, a power plant may be block-loaded, meaning its minimum output constraint is equal to its maximum capacity. These units are usually fast-start gas turbines that only operate economically at full load. A block-loaded unit may be committed by the clearing algorithm but, because it cannot supply the next marginal increment of load, it cannot set marginal prices. In this case, a lower-cost unit could set the marginal cost of the system, making the market price lower than the average cost of the block-loaded unit. For this reason, marginal prices are complemented with uplift payments, which are separate payments that compensate generators for the costs incurred above the revenue earned through market prices. Uplift payments are also referred to as side-payments or make-whole payments.

Uplifts are unavoidable in an optimal dispatch-based market, required to support the welfare, maximizing commitment and dispatch (i.e., to provide a remuneration aligned with dispatch orders). The underlying problem with uplift payments is that they create a discriminatory pricing regime, where not all agents benefit from the same prices, potentially creating misaligned incentives. This means price signals do not fully reflect operational costs, which can also have an effect in long-term investment decisions [16].

Pricing in USA markets, especially in recent years, has deviated from pure marginal costs in an attempt to reduce the weight of uplift, and to internalize, as much as possible, all operational costs into market prices. A notable example is the "hybrid pricing" approach first implemented in 2001 by NYISO (New York Independent System Operator) [17], and that is continuously updated and under review [18–20]. The NYISO pricing approach allows block-loaded units (as the one in the previous example) to artificially become marginal in an ex-post run of the dispatch problem, where the inflexible bid is treated as flexible (as if it could be dispatched at any level between zero and its maximum power output). This way, block-loaded units can set prices, although NYISO only applies this method for a subset of fast-start units.

The more general term used for this practice is Integer Relaxation (IR), which involves relaxing binary constraints in an ex-post pricing run of the commitment and dispatch problem (see [12] or [21] for more detailed discussions). However, the exact method is more nuanced and varies from one ISO to another. Indeed, most ISOs apply some type of IR, but they differ in which units can set prices, and whether they consider start-up and no-load cost in the pricing problem. In some cases (for instance, in the original NYISO hybrid pricing), only the minimum output constraint is relaxed in the pricing run, so only variable costs can impact prices; this practice is frequently called "EcoMin relaxation".

In addition, "fast-start pricing" is also a common term in practice because the relaxation often involves only fast-start units. Allowing fast-start units to set marginal prices can have positive effects, such as sending efficient signals to price-responsive load [22], or incentives to fast-start units to improve their performance or bid their true cost [23]. Fast-start pricing has been a hot topic in USA markets in the last years, with some relevant ISOs not fully allowing prices to reflect the short-term true costs. This led the Federal Energy Regulatory Commission (FERC) on 2019 to, for example, require PJM Interconnection and New York Independent System Operator (NYISO) to "implement tariff changes to ensure their pricing practices for fast-start resources were just and reasonable" [20]. The measures included, among others, using the same time granularity in the ex-post pricing run of the model and in the previous commitment and dispatch problem, or further relaxing the capacity of the fast-start units that are flexible in the price computation.

Fast-start block-loaded resources are certainly a very relevant part of the uplift problem, but this is not the only non-convexity causing price distortions. Start-up and no-load costs of thermal units can potentially cause uplift, both in the real-time and day-ahead markets. A more inclusive approach is applied in MISO (based on a simplified version of Convex-Hull pricing, see [24]); called approximated ELMP (extended locational marginal pricing). This approach is essentially an IR, but it is more comprehensive than NYISO's hybrid pricing. MISO includes start-up and no-load costs in pricing and applies ELMP to all fast-start resources (not only to block-loaded ones). Indeed, MISO broadened the definition of fast-start resources to allow more peaking units to set prices [25]. MISO continues to search for improvements to its pricing approach; during 2019, MISO studied the practical application of new formulations of the convex envelope for ELMP [26], with the objective of changing the methodology when the cost-benefit analysis was clear.

2.2. Markets in Europe: Abstract Bidding Formats

European power exchanges follow a completely different approach; their main goal is to provide a platform for market agents to trade electricity, simplifying as much as possible the consideration of physical constraints, under the presumed objective to facilitate trading and maximize market clearing replicability and transparency. System operation is decoupled from the market, and left to transmission system operators, which eventually enforce reliability constraints. This vision shifts part of the responsibility in optimizing the operation of generation resources to market agents and expects them to express their willingness to buy or sell power in a simpler way. For instance, most European power exchanges allow portfolio bidding, i.e., generation companies that own several generation units in the same pricing area can submit combined offers, and then internally decide the operation of each unit to reach the required production.

The basic bidding format in Europe is the price-quantity pair; however, a set of more complex bidding formats (or order types or market products, in the European terminology) is also available, as shown in Table 2.

Table 2. Bidding formats in EUPHEMIA (acronym of Pan-European Hybrid Electricity Market Integration Algorithm).

Bid Format	Description
Simple orders:	-
Hourly step orders	Buy or sell orders for a given volume and a limit price. It can be partially accepted if the market clearing price is equal to the bid price.
Hourly linear piecewise order	Buy or sell order for a given volume and a pair of prices: An initial price at which the orders begin to be accepted and a final price at which the order is totally accepted.
Block orders:	-
Regular block order	Buy or sell order for a single price and volume and a period of consecutive hours that can only be totally accepted.
Profile block order	Regular block order that can be partially accepted, it includes a minimum acceptance ratio condition.
Exclusive block orders	Set of block orders in which the sum of accepted ratios cannot exceed one.
Linked block orders	Set of block orders where the acceptance of some blocks (children) is conditioned to the acceptance of others (parents).
Flexible block order	Price and volume combination that can be accepted in several consecutive periods within a defined delivery range.
Complex conditions:	-
Minimum Income	Condition to reject all hourly orders of a resource if its daily remuneration does not reach the minimum income amount, defined by a fixed and a variable component.
Load gradient	Limit to the variation between the accepted volume at a period and the accepted volume at the adjacent periods

Hourly step and linear piecewise orders resemble the variable cost component in USA multi-part offers, but in this case, all operational costs must be internalized in the offered price (no additional components such as start-up cost are explicitly considered).

Complex conditions can be added to hourly orders to reflect more sophisticated constraints [27]. The minimum income condition available in the Iberian market can constrain the hourly orders of a unit, so they are only accepted if the income of the resource throughout the day reaches a fixed amount (representing, for example, the start-up cost) plus a variable cost component. The minimum income condition, combined with the load gradient condition, represents some, but not all, of the features of multi-part offers. However, the fixed and variable cost components are not considered for the maximization of market welfare, only to reject some hourly orders when the minimum income condition is not met.

Alternatively, block orders are bids that apply to multiple consecutive periods simultaneously, instead of a single hourly period, and are accepted or rejected based on the average price for those periods [28]. Resorting back to the example of the thermal unit, this could allow offering to start a power plant for a given set of hours, internalizing the start-up cost in the average price. Block orders can be combined using exclusive or linking conditions to represent more complex possibilities.

All order types in Table 2 are implemented in the single clearing algorithm EUPHEMIA (acronym of Pan-European Hybrid Electricity Market Integration Algorithm) [28]. However, the orders available

in the power exchange designated in each country (the Nominated Electricity Market Operator, NEMO) differ. The integration of power exchanges through the Price Coupling of Regions (PCR) initiative has achieved some standardization of market products, but for the moment, NEMOs have not fundamentally modified the orders available in their territory. For instance, complex conditions were, and still are, only available in the OMIE exchange (for Spain and Portugal), while Nord Pool (Nordic-Baltic region) and EPEX SPOT (central Europe) allow the use of block orders.

Bidding formats are now regularly reviewed; the first proposal being submitted jointly by all NEMOs dates back from November 2017 [29]. This proposal did not include significant changes besides updating some definitions. For instance, hourly orders were defined as Market Time Unit (MTU) orders, and any references to hourly periods were modified accordingly; this was to allow changes in the definition of MTU in the future (the goal is to move from hourly to 15-min periods). In addition, it generalized some definitions to allow the use of all orders as both supply and demand. For example, the "maximum payment condition" was introduced as the demand-side version of the minimum income condition. It is worth noting that, although the 15-min change has not been implemented at the time of this writing (it is expected for 2021), it currently focuses most of the efforts in EUPHEMIA developments [30].

The day-ahead products were approved by all Regulatory Authorities and agreed to at the Energy Regulators' Forum on 23 January 2018. Every two years, NEMOs shall consult the products that should be included in the day-ahead market. The last consultation proposal in April 2020 [31] includes a number of amendments to the current list of day-ahead market products, with the most relevant being the inclusion of a new complex condition: the Scalable Complex Orders (SCO). Unlike the classical MIC order that imposes a minimum income condition (MIC) expressed using a fixed cost and a variable cost, the Scalable Complex Order does not use the former variable cost, and instead uses the prices of the hourly suborders as variable cost on top of a fixed cost. As pointed out by EIRGRID et al. [32], the theoretical merit of SCO over MIC is to improve EUPHEMIA performance, but this merit can only materialize if the SCO eventually replaces (not complements) the classical complex orders.

2.2.1. Pricing Approach

European bidding formats, although seemingly simpler than their USA counterpart, also create non-convexities in the clearing problem with similar implications in the determination of market prices. However, the market clearing approach is not a pure welfare-maximization; the essence of European markets is to determine the highest welfare solution possible that also meets these two constrains:

- The market must be cleared with uniform prices; this entails market prices (without uplifts) must suffice to compensate all accepted bids.
- Simple bids (hourly orders) must be fully accepted if in the money, i.e., if the market price is above a generation offer, this offer must be fully accepted.

The uniform-pricing principle is often rephrased as a restriction that does not allow the existence of paradoxically accepted bids (PAB). PABs would be offers accepted in the market which are unprofitable at market prices, so it can be likened to units that require uplift payments in the USA context. Since PABs are not allowed in European markets, uplifts are consequently not needed either. The second principle refers, in the European terminology, to paradoxically rejected bids (PRB), which are those rejected bids that would apparently be profitable at given market prices [28]. As stated previously, simple bids cannot be paradoxically rejected, but PRBs are allowed in European clearing rules for complex conditions and block orders.

These two conditions applied simultaneously constrain the welfare-maximization problem, leading (by definition) to a generally sub-optimal market welfare. This is a matter of trade-offs; in the European context, uniform-pricing is, as an objective worth, the potential loss in short-term efficiency. Among the advantages of uniform-pricing is that demand and generation interact in the market in equal terms, and it is not necessary to define rules to allocate uplift that would inevitably send inefficient signals.

3. The Increasing Need for More Complex Bidding Formats

Both USA and European markets feature different kinds of complex bidding formats, revealing the higher complexity of electricity markets compared to other commodities. Both USA and EU day-ahead markets were originally designed for a predominantly thermal generation mix (with some notable exceptions), and most complex bidding formats were justified by the operational characteristics of thermal generation resources. While this is quite clear in USA multi-part bids (see Table 1), complex European bidding formats tackle the same problem from a different angle. Section 3.1 explains how the needs of thermal resources are addressed (with some limitations) by European bidding formats, and how the penetration of renewable energy sources makes these complex bidding formats even more necessary.

The transition to a low-carbon power system will most likely necessitate from new energy resources (such as batteries), which will bring their own operational constraints, requiring, as well, new bidding formats. This further justifies the need for complex bidding formats, as discussed in Section 3.1.3.

3.1. Operation of Thermal Resources

The following example describes the bidding requirements of thermal resources, building from the simplest bidding format possible, to progressively introducing more complexity as the limitations of the simpler formats arise.

This section focuses especially on the challenges derived from the variability and uncertainty of renewable energy sources. Each step of this sequence faces a trade-off between the operational efficiency lost in day-ahead dispatch decisions from using too simple bidding formats, and the additional computational burden required to introduce complex bids.

3.1.1. Initial Setting

As a starting point, the simplest design possible is a single-period simple auction, where market agents submit price-quantity pairs to express their available production or desired consumption, and their production cost or demand utility. This design takes the assumption that producers' cost structure consists fundamentally only of variable costs, and/or producers are able to properly predict how their plants will be committed (so they can efficiently internalize their non-convex costs in their bids). This could be a reasonable proxy in power systems dominated by thermal power plants and characterized by a rather flat net thermal demand (i.e., hydrothermal systems) or at least characterized by a highly predictable one. This is the foundation of European power exchanges, and, for instance, the Italian day-ahead market which still uses only simple orders (still, the original designs were quickly complemented with intraday markets to allow market agents to rectify dispatch decisions, see, for example, [31]) [33]. If they ever did, these assumptions do no longer hold in practice in the vast majority of European markets, so there is a severe risk that this approach does not provide the most efficient, or even a feasible dispatch.

Traditionally, another way of facilitating dispatch decisions, despite the simplicity of the power exchange, was to allow portfolio bidding. The lack of complexity can be compensated by managing a large portfolio instead of an individual power plant. Large generation companies with sufficiently diverse portfolios mitigate the impact of an infeasible outcome of the market clearing algorithm, since a large portfolio allows "absorbing" potential inefficiencies. However, in this context, this approach is nothing but a potential market barrier for potential new entrants, and eventually an alternative to exercise some degree of market power. By using simpler bids, generation companies benefit from disclosing the minimum amount of information about their operating cost structures, as limiting the amount of information contained in bids complicates monitoring of market power, since it is very difficult to link bid parameters to actual operating costs [27].

3.1.2. Variability

One of the reasons why this simple model can lead to inefficiencies is that, in the real multi-period problem, it cannot reflect constraints coupling different periods. For example, thermal power plants have ramping constraints that make the production available in one period dependent on the production in the preceding and following periods. One of the effects of the introduction of renewable energy sources is an increase of the cycling regime of thermal units [34]; in summary, ramping constraints are expected to be binding more frequently, and not incorporating this constraint in the day-ahead schedule can significantly degrade the efficiency of the dispatch.

Dispatch efficiency could be improved by incorporating ramping constraints in the optimization model (as in USA markets, or using the load gradient condition), at the expense of some market transparency, but this is not the only way to face this problem. For instance, block orders allow bypassing this problem, if used to offer a predefined production profile (the so-called profiled block orders), as shown in Figure 1.

Figure 1. Simple block order representing a ramp-constrained production profile.

3.1.3. Uncertainty

Using a block order requires the producer to take, prior to the market clearing, a decision about what would be the best production profile to offer into the market. The underlying assumption has been that producers can easily forecast market conditions (not only market prices, but also the resulting unit commitment), and offer the most efficient production profiles. In reality, the market outcome is uncertain, and generators need to account for the uncertainty of demand forecasts, competitors' bids, and renewable production schedules [35].

Linked and exclusive block orders can mitigate some of this uncertainty, allowing producers to express a wider range of potential operating profiles for the clearing algorithm to choose. For example, Figure 2 shows how two additional block orders (orange and green) can be linked to the previous order to potentially extend the range of hours where the unit is operating. Linked orders can only be accepted if the previous (parent) order is accepted.

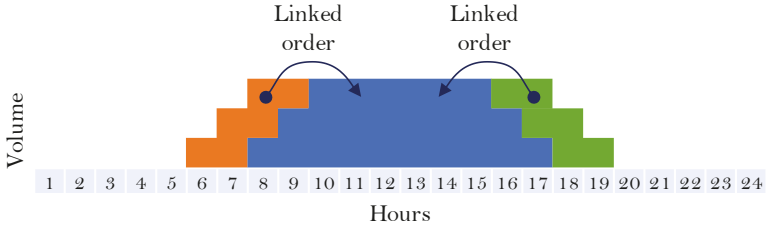


Figure 2. Linked block orders representing multiple possible production profiles.

The uncertainty associated to renewable production has greatly increased the need to model the complexities of power system operation. In USA markets, day-ahead bidding formats already represent operational constraints with detail, and renewable energy sources (RES) does not involve either a change in the way agents bid in the market nor an increase in the number of bids. However, in European markets, the use of block orders has increased significantly in recent years. Vázquez et al. [36] analyzed this effect for the Spanish case.

Figure 3 shows the average and maximum number of block orders used in the PCR region (data from European Stakeholder Committee of the Price Coupling of Regions [37,38]. Only annually aggregated data was available for the period 2011–2013, and monthly for 2014–2017). Not only has the total number of block orders almost tripled from 2011 to 2017, the use of the most complex block types has had even greater growth. As discussed in Section 4.1.2, the use of block orders and complex conditions is expected to keep on rising in the following years. This represents a major challenge from the computational point of view.

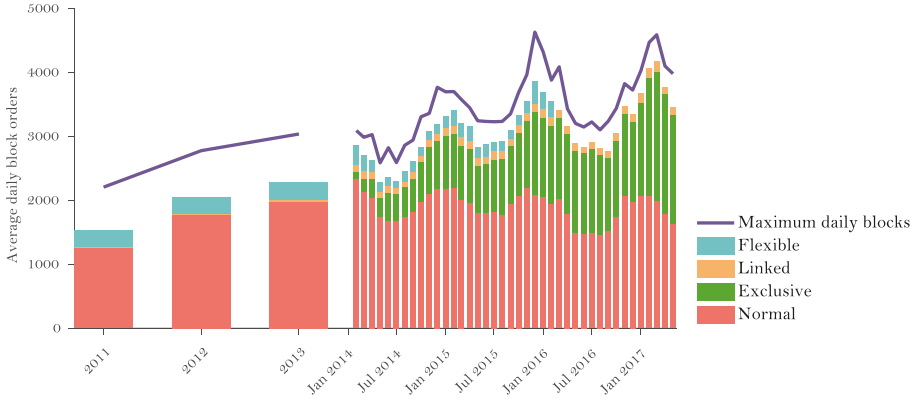


Figure 3. Average and maximum daily number of block orders in PCR (Price Coupling of Regions) region. (Flexible orders reported before March 2016 correspond to a definition phased out by Nord Pool (Flexible Hourly Block Order); no data was available for the new flexible orders).

The use of exclusive orders remarks the fact that, in the uncertain context resulting from renewable production, producers cannot easily plan the operation of their units. Exclusive orders allow expressing multiple possible production profiles of which only one can be accepted by the market, therefore, it makes it easier for market agents to make offers for different scenarios. For example, the orders shown in Figure 2 express three different production profiles, which could also be represented by three exclusive orders. Exclusive orders can sometimes express a wider range of possibilities than linked orders, since exclusive orders do not need a common parent block. In the example in Figure 4, a unit does not know what the best time to sell its production is, so it offers three different blocks and the clearing algorithm will select the best one (maximizing market welfare) only.

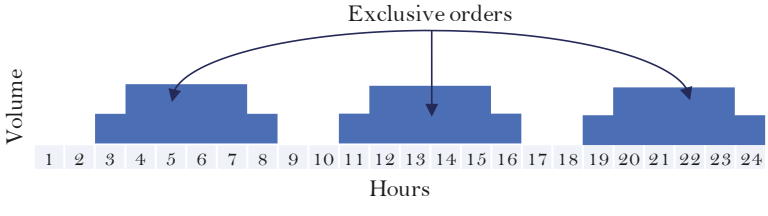


Figure 4. Exclusive block orders expressing multiple production profiles.

3.2. New Bidding Formats for New Resources

The development of bidding formats has been clearly influenced by the needs of thermal power plants, not only in the USA where multi-part bids are used, as previously described complex European bidding formats have also been tailored to the needs of thermal units.

The transition to low-carbon power systems will likely reduce reliance on thermal resources as other flexible energy technologies enter into play to compensate for the variability and uncertainty of

renewables. It is difficult to predict what resources will play that role in the future, but whether it is in the form of batteries, hydrogen or other storage resources, all are clear candidates.

Pumped-hydro storage has been present in power systems for many decades, and for the same reasons that complicate the operations of thermal power plants, the participation of resources with (limited) storage capabilities now requires more complex bidding formats.

The key challenge for storage arises when bidding formats force these resources to decide, in advance, when (in which time intervals) to bid as a generator and when to bid as demand. The volume of generation and consumption possible from a storage resource are interdependent. Bidding is especially difficult for new electro-chemical storage resources (such as Lithium-Ion batteries) because of their limited storage capacity. Grid-scale batteries, due to their high cost, are usually sized to store energy only for a few hours at nominal capacity, while pumped-hydro resources can have up to weekly or monthly planning cycles. Although both resources have limitations to participate in power markets with current bidding formats, small storage resources have clearly more limitations.

3.2.1. Developments in USA Markets

Resource-specific bidding formats in USA markets clearly provide almost perfectly-adapted bidding parameters for a selection of resources, but on the downside, discriminate potential new resources which cannot enter the market with ease until specific bidding formats are designed for them. For instance, pumped-hydro resources have participated in ISO markets for many years, but smaller storage resources (e.g., batteries, flywheels) have different constraints that cannot be represented with existing bidding formats.

The abovementioned created concerns that unnecessary barriers to storage resources were limiting competition, and the FERC initiated a consultation in November 2015, which culminated in Order 841 in February 2018, entitled "Electric Storage Participation in Markets Operated by Regional Transmission Organizations and Independent System Operators" [39]. The requirements most relevant to the topics discussed in this paper are the following [40]:

- ISOs must include a participation model for electric storage resources (ESRs) that allows them to participate in energy markets (also in ancillary service and capacity markets) when technically capable of doing so. This participation model has to prevent conflicting dispatch signals in the same market interval (charging and discharging at the same time).
- ISOs must allow self-management of state of charge (SOC).
- ESRs must be able to set the wholesale price both as seller or as a buyer, when it turns out to be the marginal resource.
- ISOs must account for physical parameters of ESRs through bidding or otherwise.

As it is usually the case, the FERC order allows a high degree of freedom for ISOs to customize rules to their specific context. Therefore, these requirements will not be homogeneous across ISOs, but the rule provides an interesting judgement on what are the most relevant constraints of batteries. Table 3 summarizes the potential bidding format for storage following ISOs' resource-specific approach.

The first characteristic that makes these bidding parameters different from traditional multi-part bids is that it allows for both charging (consuming) and discharging (generating) regimes, in a single bid. Before, storage resources needed to present separate offers as generators and consumers.

Although many of the bidding parameters are equivalent to the ones used in multi-part bids—maximum/minimum operating limits, ramp rates and maximum/minimum run times— a new participation model is necessary because the constraints are applied simultaneously for charging and discharging. Furthermore, new constraints are necessary to represent the limited energy storage. The state of charge represents how much energy is stored in a battery with respect to its maximum capacity. The definition of models to manage the state of charge has been one of the most open design elements, and has led to different approaches [41]. All ISOs have implemented, as an option,

the so-called self-schedule model, where storage operators are responsible for dispatching the output (independently from the market clearing algorithm).

Table 3. Multi-part bid for storage.

Charging	Discharging	Unit
Max charge limit	Max discharge limit	MWh
Min charge limit	Min discharge limit	MWh
Charge ramp rate	Discharge ramp rate	MWh/min
Max charge time	Max run time	hours, min
Min charge time	Min run time	hours, min
Energy bid curve	Energy offer curve	MWh, $/MWh
State of charge parameters		
Initial state of charge	-	p.u.
Max state of charge	-	p.u.
Min state of charge	-	p.u.

On the other extreme, storage operators may prefer to entirely leave to the ISO the responsibility of achieving a feasible dispatch. In this model (known as ISO-state of charge-management), storage operators would not submit hourly price-quantity offers, but rather, the technical parameters allowed in the market. This model has been implemented in CAISO (California ISO), NYISO, and PJM (but only for pumped-hydro storage).

Other models are possible between these two approaches. For instance, the so-called "self-state of charge management", where the operator of the storage facility is still responsible to achieve a feasible dispatch, but may present simple bids in the market to optimize its schedule. This model has been implemented in CASIO, NYISO, MISO, SPP (Southwest Power Pool) and PJM. This model is fundamentally aimed at the real-time market, where the operator can monitor the state of charge and update market bids.

The deadline for ISOs to implement Order 841 was December 2019 Most ISO/RTOs have achieved compliance with the order, but two are still on track to meet the requirements. The New York ISO (NYISO) requested an extension to 2020 (accepted), and Midcontinent ISO (MISO) to 2022 (also accepted).

3.2.2. Developments in European Markets

Arguably, the approach implemented in EU power exchanges provides a general set of "abstract" bidding formats that can be used by any type of resource. However, the current design was not conceived for storage resources.

As described for the USA context, the main bidding requirement of storage resources is to represent the physical link between supply offers and demand bids. In this regard, linked block orders provide a limited way to represent this constraint. For example, as shown in Figure 5, a storage resource could submit a purchase block order and a linked sell block order. This way, the sell order will not be accepted if the purchase order has not been accepted as well. In other words, the battery will only be discharged if it has been charged before.

This use of linked orders has two main limitations for storage resources. First, market agents must decide, in advance, two potential periods to buy and sell power, so the use of storage is not fully optimized. Potentially, this limitation could be addressed by submitting multiple pairs of linked buy-sell orders, including an exclusive constraint so only one of the pairs is accepted. However, linked and exclusive orders cannot be combined.

Second, the linked order guarantees a feasible schedule (since the battery will not be discharged if it has not been charged before), but because the link can only go in one direction, the parent purchase order could be accepted without the sell order. This would produce a feasible but clearly suboptimal

schedule, leaving the battery charged without a commitment to sell its energy. This creates a risk to incur losses if using this bidding format. To address the latter limitation, EPEX Spot introduced a new type of bidding format called "loop order" [42]. Loop orders allow submitting two (and only two) blocks which will be executed or rejected together by the clearing algorithm. The introduction of this new order type highlights that bidding formats need to be continuously updated as the needs of market agents evolve; and shows a shift from the "abstract" bidding format approach to much more resource-specific products.

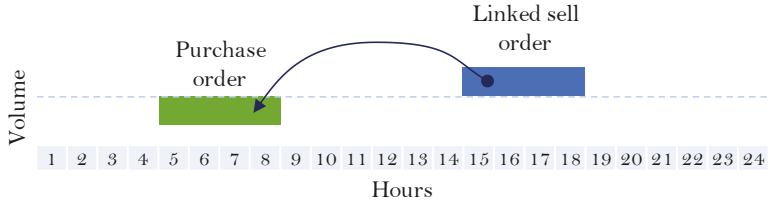


Figure 5. Use of linked block orders by storage resources.

It is worth noting the Nordic power exchange features flexible orders, which is an extremely useful bidding format for limited-energy resources, although it is aimed mainly at storage resources with a longer than one-day planning cycles (such as hydro storage). Flexible orders allow participants to express the maximum volume of energy they are willing to sell, and the clearing algorithm will select its optimal allocation. However, flexible orders share some of the abovementioned limitations for small storage resources. The problem of coordinating the sale and purchase of power within a single day remains because it is not possible to combine flexible and linked or exclusive orders.

The current discussion on the products' definition (the aforementioned consultation process launched in April 2020) includes a number of minor amendments to the products, none of them affecting the capability to properly bid storage resources in the day-ahead market.

4. Challenges Ahead

As described in the previous section, the operational needs of thermal units in a context of increased renewable penetration, and the potential needs of new energy resources, call for rethinking bidding formats. However, reforming such a fundamental market design element opens some additional debates. Section 4.1 discusses some of the computational challenges that could arise from the introduction of complex bidding formats, and Section 4.2 elaborates on the implications in Europe of introducing complex bidding formats similar to those in the USA.

4.1. Computational Tractability

In both USA and European markets, the computational complexity of the clearing problem is an instrumental factor that conditions what bidding formats can be implemented in practice. USA markets use market welfare-maximizing optimization models to clear markets—which despite their large size are reasonably tractable problems—followed by separate pricing models. European markets, despite including less detail in modeling the physical constraints of the system, combine clearing and pricing in a single more computationally complex model.

4.1.1. Challenges in USA Markets

As previously discussed, USA markets use detailed multi-part bids, which capture most of the complexity of thermal generating units. This model is well prepared to face the introduction of larger shares of renewable production. ISOs have progressively increased the modeling detail in their markets [43], as made possible by optimization software improvements and developments in computing technology. This does not mean the USA model is not constrained by its computational

tractability, but for the moment, computational improvements have continued to allow for incremental modeling enhancements. For instance, some ISOs have already implemented new bidding formats for storage, similar to the one previously described, see e.g., [44,45]. However, computational problems could arise, not because of the complexity of these bidding formats, but due to the larger number of participating resources. New storage resources could be 1 MW or less in size, which is orders of magnitude smaller than traditional resources, meaning the number of market participants could be hundreds of times the current amount. The size of the resulting commitment and dispatch problem could become intractable, and indeed, FERC Order 841 [39] included provisions to allow increasing minimum bid size requirements:

> *We are also not concerned about the potential availability of software solutions as multiple RTOs/ISOs already provide a minimum size requirement of 100 kW for all resources and have not expressed similar concerns regarding the minimum size requirement. While establishing a minimum size requirement of 100 kW for the participation model for electric storage resources will result in some smaller resources entering the markets in the near term, we do not expect an immediate influx of these smaller resources or any resulting inability to model and dispatch them. However, we recognize this finding is based on the fact that there are currently fewer 100 kW resources than there may be in the future. Therefore, in the future, we will consider requests to increase the minimum size requirement to the extent an RTO/ISO can show that it is experiencing difficulty calculating efficient market results and there is not a viable software solution for improving such calculations.*

This computational problem in ISO markets results from the combination of two factors: complex bidding formats and the number of resources. Therefore, the scalability issue can be confronted from both sides. Increasing the minimum size requirement is a way to reduce the number of resources, but this also limits competition, so it is only acceptable as a short-term solution. This measure should be accompanied by the development of rules for the participation of aggregations of resources, which opens all sorts of new questions. For instance, defining bidding parameters for aggregators cannot follow the usual resource-specific approach in ISO markets, since this participation model calls instead for general bidding parameters.

An alternative approach would be to simplify existing bidding formats. Just as creating participation models for aggregations rather than individual resources, this approach would reduce the ability of ISO markets to model the physical system accurately. Taking any of these solutions would significantly change the current modus operandi in ISO markets, but there are several reasons why ISO markets will not need to simplify its approach all the way to European-like bidding formats. As already discussed, the welfare-maximizing clearing approach allows for much more complex bidding formats than the uniform-pricing clearing approach, now and in the future. Furthermore, ISOs have not shown interest in facing one of the greatest challenges of European markets (see next section), which is to integrate several states/countries in a continent-scale market. For now, each ISO market has a well-defined footprint, and although some markets are expanding their geographical scope (for example, California ISO's real-time market has been opened to neighboring balancing authority areas through the Energy Imbalance Market (See www.westerneim.com)), no plausible plans exist to further integrate all North American ISOs. Such an objective in the future, however, would most likely require taking modeling simplifications.

4.1.2. Challenges in European Markets

As discussed in Section 2.2.1, the uniform-pricing rule conditions the clearing problem in European markets. This approach combines clearing and pricing in a single, more computationally complex problem. Computational complexity becomes especially relevant when considering the ultimate goal of European markets is to integrate all European member states in a single clearing algorithm. Computational problems have already surfaced during the first years of operation of the PCR, and Eastern European markets are to join the common platform in the upcoming years. Probably, the most

relevant concern nowadays in European markets is the existence of PRBs. As previously described, PRBs are unavoidable under uniform price-based clearing; however, in certain cases, bids may be incorrectly rejected due to the computational complexity of the algorithm. As pointed out by the Market Parties Platform (MPP) in the European Stakeholder Committee of the Price Coupling of Regions [46]:

> *There may exist false PRBs: rejected in-the-money blocks that could have been accepted and result in a better (higher welfare) solution. MPP asks for more transparency on optimality, to prove the absence of false PRBs.*

The reason behind this matter is that, mathematically, the clearing problem is a non-linear and non-convex problem, for which it is difficult to prove the optimality of a solution, or to take a quantitative measure of the quality of a solution. This may hinder the confidence of market participants, together with a lack of clarity in the public documentation of the clearing algorithm [29]. The joint response of ACER (Agency for the Cooperation of Energy Regulators) and CEER (Council of European Energy Regulators) [47] to the European Commission's Consultation on a new Energy Market Design claims that: "We would particularly like to see clearer rules and greater transparency around the market coupling algorithm (EUPHEMIA)".

Computational complexity may limit the scalability of European markets, not only to integrate more Member States in the PCR, but also to cope with the increasing trend in the number of block bids submitted to the market.

Performance and scalability are two pillars that need to be reinforced in the algorithm, and as a consequence, ACER Decision 08/2018 on the Algorithm Methodology on 26th July 2018 established that NEMOs have to report regularly on the following aspects regarding day-ahead market coupling: Operations (incidents and corrective measures), Performance Monitoring (performance indicators, including paradoxically rejected blocks and social welfare), Scalability and R&D.

As regards to the scalability concern and the growing use of block order and complex conditions, the first report, published in November 2019 [48], confirmed the expectation is for the use of both complex and block orders to keep growing in the coming years. Figure 6, from said report, shows the expected usage of complex and block orders in 2021 (expressed as percentage of 2018 usage), with the use of complex orders and linked block orders expected to grow to 170% of its 2018 value. Those estimates did not take into account the potential impact of the forthcoming 15/30-min products, for which proper data sets and specifications were still missing.

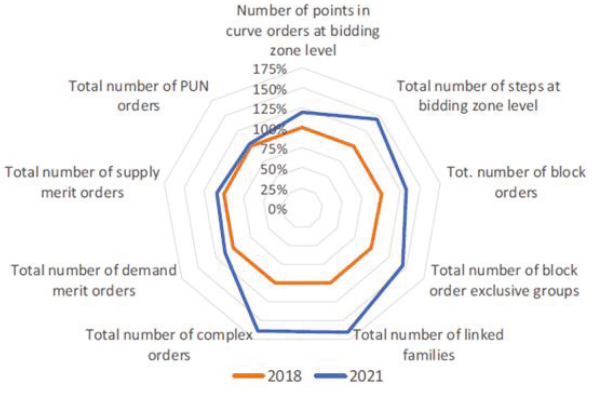


Figure 6. Usage of complex and block orders in 2021 (expressed as percentage of 2018 usage) [48].

EU power exchanges have traditionally limited the amount of block orders submitted by portfolio (i.e., by market agent and market area), as shown in Table 4. A potential solution would be to further reduce block order limits, but this is clearly not a desirable outcome.

Table 4. Limits to block orders in main European power exchanges [49,50].

	EPEX SPOT							NORDPOOL		
Area	**AT, DE-LU**	**FR**	**BE, NL**	**DK, FI**	**NO, SE**	**CH**	**GB**	**NO, SE, FI, DK, EE, LV, LT**	**AT, DE-LU, FR, BE, NL**	**GB**
Block order										
Max. volume (MW)	600	600	400	500	500	150	500	500	800	500
Max. Blocks/Portfolio	100	40	40	100	100	40	80	50	100	80
Linked block order										
Max. Levels	7	7	7	7	7	7	7	3	7	
Max. Children	6	6	6	6	6	6	6	6	6	
Max. Blocks/Family	7	7	7	7	7	7	40	7	13	
Max. Families/Portfolio	5	5	5	5	5	1	7			
Exclusive group block orders										
Max. Blocks in group	24	24	24	24	24	24	24	15	24	15
Max. Groups/Portfolio	5	5	5	5	5	5	5	3	5	3
Loop blocks										
Max. Volume/Family (MW)	800	800	500	800	800	300	800			
Max. Blocks/Family	2	2	2	2	2	2	2			
Max. Families/Portfolio	3	3	3	3	3	3	3			
Flexible orders										
Max. Orders/Portfolio								5		3

Furthermore, current rules require the clearing algorithm to obtain a solution in less than ten minutes [51]—A much more demanding timeline than USA markets, so the quality of the solution could also be improved by allowing additional time for the clearing process. However, European markets do not incorporate many physical constraints, making additional corrections by Transmission System Operators necessary, so it may not be possible to extend this timeline by a large margin.

Both reducing the number of block orders, and extending the time available to compute the solution, are temporary fixes. In the long term, it is necessary to focus on the root causes. The number of block orders submitted has greatly increased and will keep increasing because no single bidding format properly addresses the needs of market agents, therefore, agents combine orders in an effort to hedge against all the possible market outcomes.

A more permanent alternative would be to create resource-specific bidding formats that would only require one (multi-part) bid per resource. However, since such resource-specific bidding formats would likely be more complex, they should be carefully designed to ensure they indeed reduce the number of orders and overall problem complexity. A currently discussed alternative in this line is the introduction of thermal orders, which is nothing more than multi-part bids like the ones used in ISO markets.

The European Stakeholder Committee of the Price Coupling of Regions [38] suggests that thermal orders can be preferable (from a computational efficiency point of view) if an agent can submit a single multi-part bid replacing multiple block orders. However, in the same stakeholder committee, the EUPHEMIA software provider (N-Side) has also pointed out that including such a bid would need to consider a significant change in the market design, and pricing and clearing rules, which are discussed in the next section.

4.2. Clearing and Pricing Rules

The same fundamental problem—pricing in non-convex markets—has surfaced in different ways in USA and European markets because of differences in clearing rules. The main challenge in USA markets is to reduce uplift payments, which may be a clearer symptom of inefficient pricing, but similar issues should be expected in the European context. The uniform price-based clearing rule in EUPHEMIA relies on marginal pricing, even if the dispatch solution is not fully welfare-maximizing. This means that, in the EU context, inflexible bids cannot set market prices (similarly to block-loaded units in USA markets). As described by Eirgrid et al. [51]:

> *The effect of defining an order as a block is that the order cannot then be a full price maker. Rather, block orders may impose a bound on the range of prices possible while the price being set would still need to come from the simple order or complex order curves. This is because the decision to execute the order is an integer decision (i.e., the order is executed or not executed) and the decision on whether to accept a block occurs before the price determination sub-problem. The bound created by the last accepted block order would function to affect the price (by limiting possible values) but could not directly set this price.*
>
> *This was discussed with the PCR ALWG representative, APX, who confirmed that without the blocks setting the price, the price could only be set by other price makers, i.e., simple orders or complex orders, or the price indeterminacy rules of EUPHEMIA.*

Another similarity with USA markets is in the way some units may represent their start-up costs using the minimum income condition. This constraint guarantees that the offers of a unit will only be accepted if the price is high enough to compensate a given fixed cost (representing the start-up cost). Although this fixed cost influences the clearing of the market—triggering the rejection of the offer if the price is not high enough—it does not directly set the market price, which is always set by a simple (marginal) bid.

In a scenario where most bids are inflexible, or where they use any non-convex order type, EUPHEMIA can lead to market prices that do not accurately reflect system costs. Note that, even if

uniform prices necessarily include all operational costs, there is a nuanced difference between market prices being high enough to cover costs, and market prices being reflective of system costs. This is, in essence, the same problem described for USA markets, where inflexible units could not set the price. Fortunately, this also means that European markets could benefit from the solutions developed for the USA context. However, the only way to allow alternative pricing rules is to modify clearing rules as well.

Clearing rules in European markets are based on uniform pricing, but this is not imposed by bidding formats. Even though market orders always express their constraints with respect to market prices (i.e., if price is below X, reject order), the uniform-price clearing approach is mostly the result of a market that has evolved from a simple auction. The basic information contained in market orders is the necessary remuneration, but said remuneration could include uplift payments if such a policy choice was made. Obviously, this requires a significant change, but current bidding formats do not prevent using a market welfare-maximization clearing approach and an alternative pricing rule.

Indeed, a welfare-maximization clearing approach would greatly simplify the clearing algorithm, helping European markets cope with the current increase in the use of block orders. Additionally, it would enable more complex bidding formats, such as thermal orders; and more complex combinations of block orders, as made necessary by storage resources.

This discussion had a greater momentum back in 2016 [52] and today is not high up on the agenda of NEMOs. However, the possibility has not been completely discarded, and it is still included among the different R&D activities to be explored [53].

5. Conclusions

The penetration of renewable energy resources has significantly altered power systems. In light of these changes, wholesale electricity markets, and, in particular, day-ahead markets, in their role to facilitate planning and operating decisions, require increasingly complex bidding formats. While USA markets already provide detailed multi-part bids to reflect the most relevant constraints of thermal generators, European markets provide a limited choice of block orders and complex conditions. These orders may be falling short to facilitate an efficient participation of all resources into electricity markets, as evidenced by the ever more complex combinations of orders submitted by market agents to achieve an adequate representation of their constraints.

The energy transition will also bring about the introduction of new energy resources, for example, batteries and other types of storage, making it necessary to address their needs and remove barriers for effective competition. The definition of participation models for storage is underway in USA markets, but European markets lack specific bidding formats for these types of resources. Although European markets use abstract bidding formats that should not discriminate any resource (meaning, they are equally limited for all types of resources), storage resources face significant barriers.

Since most difficulties have been identified in the European context, this is where the following conclusions focus. Regarding the limits of current bidding formats to represent both thermal and storage resources, a potential solution is to increase the range and complexity of the orders available. However, European markets are already facing computational difficulties, and this approach would most likely fall into scalability issues.

Therefore, the most sustainable approach in the long term would need to both reduce the computational complexity of the clearing problem, and allow more complex bidding formats. This may seem an impossible puzzle, but there are ways in which it can be achieved. First, resource-specific bidding formats, similar to USA's multi-part bids, can, in some cases, reduce the computational burden if one multi-part bid substitutes a complex combination of block orders. At the same time, resource-specific bidding formats would remove barriers for small market players (current portfolio bidding is advantageous for large players), and facilitate market monitoring.

However, the primary cause for the limitations of European bidding formats is the clearing approach. European markets are based on uniform prices; clearing the market under uniform-pricing

constraints complicates the computation of market programs, so the range of bidding formats available is limited to keep the computational burden under control. Alternatively, USA markets use computationally simpler clearing algorithms based on the maximization of market welfare (without pricing constraints). As reviewed in previous chapters, this approach requires an ex-post price computation, with its own challenges, but it would enable the use of increasingly necessary complex bidding formats in the European context.

In conclusion, resource-specific bidding formats are most advantageous, especially when combined with welfare-maximization clearing rules. However, their design has to be regularly reviewed to ensure no resources are discriminated. This is especially challenging when considering future potential energy resources, of which their technical characteristics are yet unknown. Nonetheless, this cannot be strictly considered a disadvantage over European (resource-independent) bidding formats, since these are equally influenced by current resource needs, and also become outdated. For instance, European-bidding formats present several limitations to represent the constraints of storage resources, so their resource-independence cannot be unequivocally considered a positive feature, unless full generality is achieved.

Author Contributions: I.H. conceived and designed the analysis, collected the data, performed the analysis and wrote the paper. P.R. conceived and designed the analysis, collected the data, performed the analysis and wrote the paper. C.B. conceived and designed the analysis, collected the data, performed the analysis and wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Hu, J.; Harmsen, R.; Crijns-Graus, W.; Worrell, E.; Van Den Broek, M. Identifying barriers to large-scale integration of variable renewable electricity into the electricity market: A literature review of market design. *Renew. Sustain. Energy Rev.* **2018**, *81*, 2181–2195. [CrossRef]
2. IRENA. *Adapting Market Design to High Shares of Variable Renewable Energy*; International Renewable Energy Agency: Abu Dhabi, UAE, 2017; pp. 1–168.
3. Anuta, H.O.; Taylor, P.; Jones, D.; McEntee, T.; Wade, N. An international review of the implications of regulatory and electricity market structures on the emergence of grid scale electricity storage. *Renew. Sustain. Energy Rev.* **2014**, *38*, 489–508. [CrossRef]
4. Green, R. Electricity wholesale markets: Designs now in a low-carbon future. *Energy J.* **2008**, *29*, 95–124.
5. Conejo, A.J.; Sioshansi, R. Rethinking restructured electricity market design: Lessons learned and future needs. *Int. J. Electr. Power Energy Syst.* **2018**, *98*, 520–530. [CrossRef]
6. Addepalle, P.; Hartman, A.; LiB, S.; Borissov, B. *Energy and Operating Reserve Markets Business Practices Manual, BPM-002-r18*; Midcontinent Independent System Operator: Carmel, IN, USA, 2019; pp. 1–274.
7. Jacob, M. Missing Incentives for Flexibility in Wholesale Electricity Markets; USAEE Working Paper No. 20-453. Available online: https://ssrn.com/abstract=3623962 or http://dx.doi.org/10.2139/ssrn.3623962 (accessed on 15 July 2019).
8. Hogan, W.W.; Ring, B. *On Minimum-Uplift Pricing for Electricity Markets*; Harvard Electricity Policy Group and Harvard-Japan Project on Energy and the Environment: Melbourne, Australia, 2003; pp. 1–30.
9. O'Neill, R.; Sotkiewicz, P.M.; Hobbs, B.F.; Rothkopf, M.H.; Stewart, W.R. Efficient market-clearing prices in markets with non-convexities. *European J. Oper. Res. Num.* **2005**, *164*, 269–285. [CrossRef]
10. Gribik, P.; Hogan, W.W.; Pope, S.L. *Market-Clearing Electricity Prices and Energy Uplift*; Harvard Electricity Policy Group: Cambridge, MA, USA, 2007; pp. 1–46.
11. Pope, S.L. *Price Formation in ISOs and RTOs, Principles and Improvements*; FTI Consulting: Boston, MA, USA, October 2014; Available online: http://lmpmarketdesign.com/papers/Pope.EPSA_Price_Formation_Oct_29_2014_FINAL.pdf (accessed on 5 May 2019).
12. Liberopoulos, G.; Andrianesis, P. Critical Review of Pricing Schemes in Markets with Non-Convex Costs. *Oper. Res.* **2016**, *64*, 17–31. [CrossRef]

13. Hua, B.; Baldick, R. A convex primal formulation for convex hull pricing. *IEEE Trans. Power Syst.* **2017**, *32*, 3814–3823. [CrossRef]
14. Eldridge, B.; O'Neill, R. Revisiting MIP Gaps and Pricing in RTO-scale Unit Commitment, FERC Software Conference. 2018. Available online: https://cms.ferc.gov/sites/default/files/2020-05/20180627082016-W1%2520-%25201%2520-%2520Eldridge%2520-%2520FERC_Software_June2018_final.pdf (accessed on 10 May 2019).
15. Eldridge, B.; O'Neill, R.; Hobbs, B.F. Near-Optimal Scheduling in Day-Ahead Markets: Pricing Models and Payment Redistribution Bounds. *IEEE Trans. Power Syst.* **2019**, *35*, 1684–1694. [CrossRef]
16. Herrero, I.; Rodilla, P.; Batlle, C. Electricity market-clearing prices and investment incentives: The role of pricing rules. *Energy Econ.* **2015**, *47*, 42–51. [CrossRef]
17. INC's. *Motion for Permission to Implement Hybrid Fixed Block Generation Pricing Rule*; New York Independent System Operator: New York, NY, USA, 2001.
18. FERC. Order Instituting Section 206 Proceeding and Commencing Paper Hearing Procedures and Establishing Refund Effective Date, 161 FERC 61,294. Docket EL18-33. Issued 21 December 2017. Available online: https://elibrary.ferc.gov/idmws/docket_search.asp (accessed on 10 May 2019).
19. NYISO. Initial Brief of the New York Independent System Operator in the Section 206 proceeding. Docket EL18-33, 20180213-5203. Filed 2 December 2018. Available online: https://elibrary.ferc.gov/idmws/docket_search.asp (accessed on 5 May 2019).
20. FERC. Order on Paper Hearing re New York Independent System Operator, Inc. under EL18-33. 167 FERC 61,057. Docket EL18-33. Issued 18 April 2019. Available online: https://elibrary.ferc.gov/idmws/docket_search.asp (accessed on 10 May 2019).
21. Ruiz, C.; Conejo, A.J.; Gabriel, S.A. Pricing non-convexities in an electricity pool. *IEEE Trans. Power Syst.* **2012**, *27*, 1334–1342. [CrossRef]
22. Hogan, W. Electricity Market Design and Efficient Pricing: Applications for New England and Beyond. *Electr. J.* **2014**, *27*, 23–49. [CrossRef]
23. Harvey, S. *Is the California ISO Becoming an Uplift Market? Pricing, Uplift and Commitment*; FTI Consulting: Folmson, CA, USA, 2014; pp. 1–24.
24. Schiro, D.A.D.A.; Zheng, T.; Zhao, F.; Litvinov, E. Convex Hull Pricing in Electricity Markets: Formulation, Analysis, and Implementation Challenges. *IEEE Trans. Power Syst.* **2015**, *31*, 1–8. [CrossRef]
25. Economics, P. State of the Market Report for the MISO Electricity Markets. 2016. Available online: www.potomaceconomics.com (accessed on 7 June 2017).
26. MISO. ELMP III White Paper I R&D Report and Design Recommendation on Short-Term Enhancements. 2019, pp. 1–31. Available online: https://www.misoenergy.org/stakeholder-engagement/stakeholder-feedback/msc-elmp-iii-whitepaper-20190117/ (accessed on 12 February 2020).
27. Schittekatte, T.; Reif, V.; Meeus, L. The EU Electricity Network Codes. 2019. Available online: https://fsrglobalforum.eu/wpcontent/uploads/2019/03/FSR_2019_EU_Electricity_Network_Codes.pdf (accessed on 5 February 2019).
28. EPEX SPOT, GME, Nord Pool, OMIE, OPCOM, OTE, TGE EUPHEMIA Public Description, PCR Market Coupling Algorithm. Version 1.5. December 2016. Available online: https://www.europex.org/wp-content/uploads/2016/11/Euphemia-Public-Description.pdf (accessed on 8 June 2019).
29. NEMOs. All NEMOs' Proposal for Products That Can be Taken into Account by NEMOs in Single Day-Ahead Process in Accordance with Article 40 of Commission Regulation (EU) 2015/1222 of 24 July 2015 Establishing a Guideline on Capacity Allocation and Congestion Management. 13 November 2017. Available online: www.europex.org/all-nemos (accessed on 8 June 2019).
30. All NEMO Committee. Status Update on Ongoing Activities. MESC Meeting. December 2019. Available online: http://www.nemo-committee.eu/assets/files/MESC%2018%20DEC%202019%20NEMO%20COMMITTEE%20presentation.pdf (accessed on 29 February 2020).
31. All NEMO Committee. Products That Can be Taken Into Account by NEMOs in Single Day-Ahead Coupling in Accordance with Article 40 of Commission Regulation (EU) 2015/1222 of 24 July 2015 Establishing a Guideline on Capacity Allocation and Congestion Management. Available online: http://www.nemo-committee.eu/assets/files/200408_Products%20Proposal_DA_CLEAN-e868fe7446813b04398c1023a3496723.pdf (accessed on 9 May 2018).

32. Eirgrid, SEMO, SONI Market Operator User Group Presentation, Dublin. Available online: https://www.sem-o.com/documents/general-publications/Market-Operator-User-Group-Presentation-16-January-2020.pdf (accessed on 5 January 2020).
33. GME (Gestori Mercati Energetici). Guidelines Facilitating Access to and Participation in GME's Electricity Markets. Available online: www.mercatoelettrico.org (accessed on 1 December 2016).
34. Rodilla, P.; Cerisola, S. Modeling the major overhaul cost of gas-fired plants in the Unit Commitment problem. *IEEE Trans. Power Syst.* **2013**, *29*, 1001–1011. [CrossRef]
35. Neuhoff, K.; Ritter, N.; Schwenen, S. *Bidding Structures and Trading Arrangements for Flexibility across EU Power Markets, Report from the Future Power Market Workshop 2015*; DIW–Deutsches Institut für Wirtschaftsforschung: Berlin, Germany, 2015.
36. Vázquez, S.; Rodilla, P.; Batlle, C. Residual demand models for strategic bidding in European power exchanges: Revisiting the methodology in the presence of a large penetration of renewables. *Electr. Power Syst. Res.* **2014**, *108*, 178–184. [CrossRef]
37. European Stakeholder Committee of the Price Coupling of Regions PCR Stakeholder Forum. January 2016. Available online: https://www.entsoe.eu/network_codes/esc/ (accessed on 11 April 2019).
38. European Stakeholder Committee of the Price Coupling of Regions PCR Status Update. June 2017. Available online: https://www.entsoe.eu/network_codes/esc/ (accessed on 11 April 2019).
39. FERC. Electric Storage Participation in Markets Operated by Regional Transmission Organizations and Independent System Operators. Order No. 841, 162 FERC 61,127. 15 February 2018. Available online: https://www.federalregister.gov/documents/2018/03/06/2018-03708/electric-storage-participation-in-markets-operated-by-regional-transmission-organizations-and (accessed on 5 January 2019).
40. Singhal, N.G.; Ela, E.G. *Incorporating Electric Storage Resources into Wholesale Electricity Markets While Considering State of Charge Management; Options*; EPRI Institute: Palo Alto, CA, USA, 2019.
41. Singhal, N.G.; Ela, E.G. *Storage Integration Efforts in the U.S. Wholesale Electricity Markets IESO Energy Storage Design; Project*; EPRI Institute: Palo Alto, CA, USA, 2020.
42. EPEX SPOT. Press Release. EPEX SPOT Introduces Curtailable Blocks and Loop Blocks on all Day-Ahead Markets. Available online: https://www.epexspot.com/en/news/epex-spot-introduces-curtailable-blocks-and-loop-blocks-all-day-ahead-markets (accessed on 8 December 2018).
43. O'Neill, R.P.; Dautel, T.; Krall, E. Recent ISO Software Enhancements and Future Software and Modeling Plans. 2011; pp. 1–42. Available online: https://www.ferc.gov/industries/electric/indus-act/rto/rto-iso-soft-2011.pdf (accessed on 7 April 2018).
44. Caiso. Energy Storage and Distributed Energy Resources (ESDER) Phase 3. Revised Straw Proposal. Comments by Department of Market Monitoring. Available online: http://www.caiso.com/Documents/DMMComments-EnergyStorage-DistributedEnergyResourcesPhase3-Jun252018.pdf (accessed on 6 June 2019).
45. McDonough, C. Enhanced Storage Participation. Enhance The Ability of Electric Storage Facilities to Participate in The New England Wholesale Electricity Markets. NEPOOL Markets Committee. 6–7 June 2018. Available online: https://www.iso-ne.com/static-assets/documents/2018/05/a10_presentation_enhanced_storage_participation.pptx (accessed on 9 July 2019).
46. European Stakeholder Committee of the Price Coupling of Regions Euphemia Performance. PCR Stakeholder Forum, September 2015. Available online: https://www.opcom.ro/opcom/uploads/doc/PCR/PCR%20Stakeholder%20Forum.pdf (accessed on 5 April 2019).
47. ACER-CEER. Joint ACER-CEER Response to The European Commission's Consultation on a New Energy Market Design. 2015, pp. 1–34. Available online: https://www.acer.europa.eu/Official_documents/Position_Papers/Position%20papers/ACER_CEER_EMD_Response.pdf (accessed on 15 July 2018).
48. All NEMO. Committee CACM Annual Report 2018, A Report Carried out in Coordination with ENTSO-E. November 2019. Available online: http://www.nemo-committee.eu/assets/files/cacm-annual-report-2018.pdf (accessed on 4 February 2020).
49. EPEX SPOT. Operational Rules. Available online: https://www.epexspot.com/en/downloads#rules-fees-processes (accessed on 10 May 2020).
50. Nord Pool. Rules and Regulations. Section 3. Product Specifications. Online. Available online: https://www.nordpoolgroup.com/trading/Rules-and-regulations/ (accessed on 30 June 2020).

51. Eirgrid, SEMO, APX Power & SONI I-SEM Trialling of EUPHEMIA: Initial Phase Report. 2015, pp. 1–68. Available online: www.sem-o.com (accessed on 7 April 2017).
52. N-SIDE. Non-Uniform Pricing and Thermal Orders for the Day-Ahead Market. PCR Stakeholder Forum, Brussels. 11 January 2016. Available online: http://www.enexgroup.gr/fileadmin/groups/PCR/Non-Uniform-Pricing_and_Thermal_Orders.pdf (accessed on 15 June 2018).
53. SDAC. Joint Steering Committee. Minutes SDAC Joint Steering Committee Meeting. 2020. Available online: http://www.nemo-committee.eu/publications (accessed on 4 April 2020).

energies

MDPI

Article

Short-Term Electricity Price Forecasting with Recurrent Regimes and Structural Breaks

Rodrigo A. de Marcos [1,*], Derek W. Bunn [2], Antonio Bello [1] and Javier Reneses [1]

[1] Institute for Research in Technology, Technical School of Engineering (ICAI), Universidad Pontificia Comillas, 28015 Madrid, Spain; abello@comillas.edu (A.B.); javierr@comillas.edu (J.R.)
[2] Management Science and Operations, London Business School, London NW1 4SA, UK; dbunn@london.edu
* Correspondence: demarcos@alu.comillas.edu; Tel.: +34-91-540-2800

Received: 25 September 2020; Accepted: 15 October 2020; Published: 19 October 2020

Abstract: This paper develops a new approach to short-term electricity forecasting by focusing upon the dynamic specification of an appropriate calibration dataset prior to model specification. It challenges the conventional forecasting principles which argue that adaptive methods should place most emphasis upon recent data and that regime-switching should likewise model transitions from the latest regime. The approach in this paper recognises that the most relevant dataset in the episodic, recurrent nature of electricity dynamics may not be the most recent. This methodology provides a dynamic calibration dataset approach that is based on cluster analysis applied to fundamental market regime indicators, as well as structural time series breakpoint analyses. Forecasting is based upon applying a hybrid fundamental optimisation model with a neural network to the appropriate calibration data. The results outperform other benchmark models in backtesting on data from the Iberian electricity market of 2017, which presents a considerable number of market structural breaks and evolving market price drivers.

Keywords: day-ahead electricity markets; electricity price forecasting; fundamental-econometric models; market structural breaks

1. Introduction

Price forecasting in electricity markets is facing frequent, and perhaps increasing, structural changes in the market. Apart from new entrants and corporate restructuring affecting market conduct, the technology mix is going through a transition to intermittent renewables and end-user engagement is becoming substantial. In addition, policy interventions are increasing as governments seek to achieve a balance of decarbonisation, security and affordability. All of this creates a modelling challenge for price forecasting. Time series estimation, therefore, has to take account of structural breaks and evolving parameters as market circumstances change. A simple response is to work with short time series to reflect only recent conditions that may be representative of the intended forecast horizon, but that limits the complexity of model estimation. In contrast, econometric methods often seek to include estimated structural break terms. These, however, tend to be limited to a few distinct interventions and do not capture more complex evolutions. Hybrid methods, alternatively, that link time series analyses to underlying market simulation models can be more effective [1], but even with a hybrid method, the choice of an appropriate time series calibration length still remains.

Surprisingly, despite its crucial role, research on how to select the appropriate data window for model estimation is an under-researched topic in forecasting. Whilst we have seen time series methods increase in complexity to capture the distinctive features of power price formation, going from ARIMA and its variants [2–4], neural network and other AI approaches [5–10] to wavelets [11–13] and various combinations, these procedures all rely on the presumption that the time series model, as estimated,

can be projected forward, which may not be so appropriate in the more evolving power systems of today. In one of the few research papers to look at this aspect, the sensitivity of forecast errors to the estimation window has been analysed in [14] and based upon this, the research in [15] presented a pragmatic averaging of forecasts of individual ARX models estimated upon different data calibration windows. However, only heuristic suggestions were made for the choice of windows for calibration.

To complicate the specification further, apart from permanent abrupt and gradual structural changes affecting the window of relevant history, power price formation is known to manifest recurrent regime changes and exhibit multi-seasonal behaviour [16], according to the interactions of periods of scarcity, input prices, weather conditions and behavioural dynamics. Thus we have seen Markov and factor-based regime-switching methods outperform single regime models in several comparative studies [17,18]. The implication of this is that there are recurrent episodes in the time series when one specification is more relevant than another. So, if we are seeking to find the most relevant window of data for model calibration, it may not be the most recent. For instance, if the power system is expected to experience a sharp and sudden increase in wind generation based on weather forecasts, it may prove advantageous to disregard for predictive estimation any periods in the past that do not present significant wind outputs. Our research, therefore, seeks to make a contribution by developing a methodology to select the appropriate calibration window for a hybrid fundamental/timeseries forecasting approach based upon considerations of structural breaks and recurrent regimes. It is evidently important to have an integrated method to select the calibration window both with respect to considerations of recurrent regimes as well as respecting structural changes, and we are not aware of this joint specification being considered in previous research. This is therefore the focus and main contribution of this research.

Various aspects of calibration window selection have appeared in previous research but without the full specification being sought in this paper. For example, in neural networks, [11] uses a training set involving the seven days prior to the forecasting day and adds three extra days based on the similarity with respect to the day immediately prior to the forecasting day in terms of daily price patterns. In contrast, [19] utilises a modified version of the similar days method proposed in [20] in order to select the 12 most similar days in a predefined 4-month calibration period according to exogenous variables available at the moment of the forecast, such as expected demand and temperature. However, these methods are motivated more by considerations of neural network overfitting issues rather than by market regime changes.

In providing a more formal method for calibration window selection, according to robust criteria for identifying both recurrent regimes and structural changes, we undertake this in the context of advocating a hybrid fundamental/econometric approach. We consider the inclusion of a fundamental market simulation model crucial for forecasting with structural changes since it can explicitly represent price formation under new market conditions. For example, the impending decarbonisation of power systems is not a recurring event that econometric approaches can interpret and project in the future. Indeed, that is why hybrid methods have become widely applied in medium-term applications [21–23]. But only a few researchers have considered applying them to the short term [19,24], perhaps because of the high computational requirement of running hourly fundamental models. Nevertheless, this issue can be dealt with by means of simplification methods, such as aggregating similar generation units [24,25]. Furthermore, given that these hybrid models explore most of the drivers of electricity prices, an immense volume of information must be handled by the models. For accuracy, however, this is worthwhile, but only if the time series calibration window is appropriately chosen.

In summary, therefore, this work attempts to provide a novel forecasting method that properly addresses the joint problem of recurrent regimes and market structural breaks in selecting the calibration window, in order to support a state-of-the-art hybrid model. The hybrid model is similar to one of the models proposed in [19] and involves a short-term model that is composed of an hourly cost-production optimisation model whose outputs provide market-related information to a neural network (NN)

model. However, there are several distinctive modelling features of this new work that add to its novel research contributions:

- Prior to the NN forecast, the NN training period, which is initially set to a very large window, is filtered by means of a structural break analysis method and periods where prices significantly differ from those prior to the forecasting period (i.e., most recent prices) are discarded.
- Furthermore, the hourly trends in the actual forecasting period according to market regime related variables are evaluated via a K-means clustering procedure. The hours of the initial NN calibration period where the assigned cluster coincides with that of the hours in the forecasting period are included in the previously filtered calibration period by the structural break analysis method. This combination of training window selection techniques is carried out ex-ante and therefore provides a dynamic calibration dataset.
- The proposed set of methodologies is backtested on the real and full-scale Iberian electricity market of 2017. The performance of this approach is compared with that of other well-recognised forecasting models.

The remainder of this manuscript is organised as follows: the methodology is described in Section 2; Section 3 presents the case studies in which the proposed forecasting method, as well as other comparative forecasting methods, have been tested; and Section 4 contains the conclusions, including suggestions for potential extensions.

2. Proposed Methodology

Essentially, this work's proposed methodology is comprised of the methods displayed in Figure 1, all of which have been tested on a real-size power exchange with complex price dynamics: the Iberian (Spain and Portugal) electricity market. The first phase of the methodology represents its fundamental component, the cost-production optimisation model. The next stage involves several data pre-processing approaches that aim to enhance the final step of the methodology, which is an artificial neural network model. Each element of the proposed methodology is explained below.

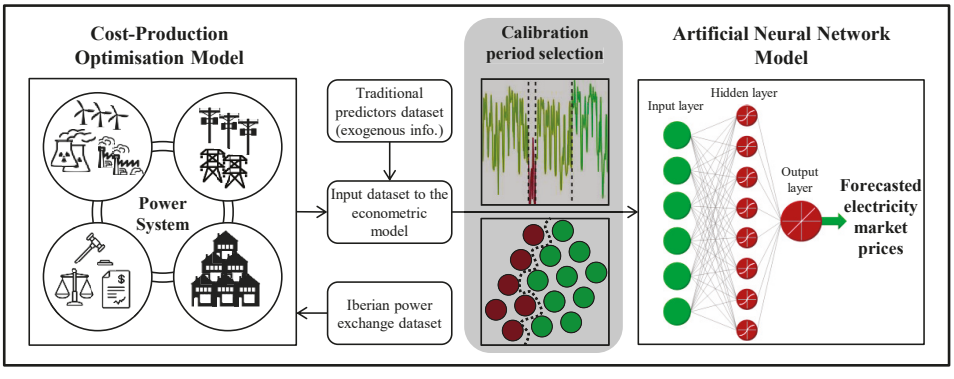


Figure 1. Overview of the proposed hybrid forecasting methodology.

2.1. Cost-Production Optimisation Model

In order to consider physical elements, regulatory limits and the operation of the market, a cost-production optimisation model, which is based on the Iberian power exchange, is specified. The required information is obtainable from the transparency platforms of the Spanish System Operator [26] and of the ENTSO-E [27]. This fundamental model seeks to reduce total system costs under perfect competition assumptions by setting the production outputs of the system's power units to optimal values. The mathematical formulation of this optimisation model is similar to the one presented

in [19,25] and estimates the electricity market price as a result of the market-clearing according to competitive fundamentals. These prices are known as system marginal prices, and they are derived from the dual variables of the demand and generation balance constraints. Furthermore, it was observed in [19] that considering thermal units separately is worthwhile in terms of accuracy. Specifically, although a week is solved by minimising system costs simultaneously throughout the 168 h in 7.4 s (up from 3.91 if the thermal units are aggregated), the forecasting error is reduced by approximately 33% when compared to the optimisation model of [25]. The optimisation is solved via relaxed mixed-integer programming (RMIP) in order to consider all units' variable costs and not only those of the committed units.

2.2. Period Selection

The main contribution of the work, however, is an improvement in model performance by achieving an appropriate calibration data selection procedure. The calibration period selection methodology provides a suitable and novel solution to this issue, allowing the subsequent NN model to handle only the necessary data by focusing upon the relevant circumstances or regimes present in the power system at the moment of the forecast. This methodology is split into three steps.

2.2.1. Structural Breaks

Before applying any filtering method, the initial dataset period needs to be oversized in order to find an appropriate subset. In this case, 13 months prior to the forecasting period are taken (i.e., a 13-month rolling window dataset), which is too large a calibration dataset for NN models if hourly precision is considered. The fact that structural patterns change throughout a 13-month period is not in question. Not only due to several seasonal effects that occur in the system but also abrupt market condition fluctuations or other structural breaks. An example can be seen in Figure 2, which shows the evolution of the Iberian electricity market prices during the autumn of 2016. It can be observed that early autumn is significantly different from late autumn. When it comes to forecasting late autumn prices (e.g., shortly after 6 December), it is evident that one should consider discarding the previous periods with the lowest prices, as they clearly correspond to other market circumstances. It should be noted that the structural breaks depicted in Figure 2 serve as an illustrative example and this work's case study does not involve forecasting prices during late 2016. The different market circumstances are separated by the vertical lines, which correspond to the structural breaks. These structural breaks have been computed by means of the "strucchange" package in R that is based on the work presented in [28].

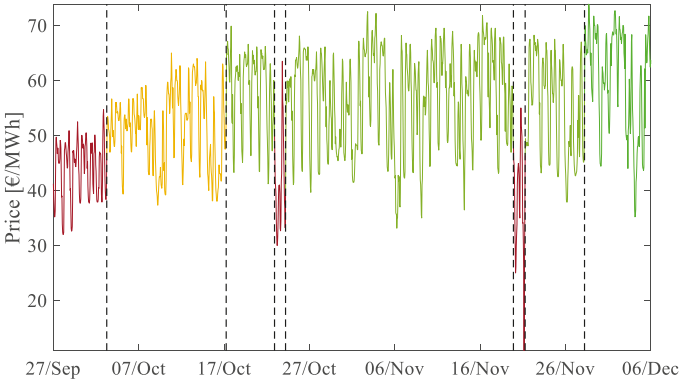


Figure 2. Iberian electricity market prices during autumn 2016.

In theory, structural breaks split a time series into several segments that feature significantly different coefficients and perhaps different model specifications. In this application, we test the baseline model that the electricity price equals a constant. Thus, the "constant" becomes the varying element in the segments that are separated by the structural breakpoints. The purpose of the methodology presented in [28] is the determination of these breakpoints whose corresponding segments provide the least total residual sum of squares of the models associated with each segment.

Evaluating a 13-month hourly dataset is cumbersome if high precision is desired and therefore, the number of candidate breaks should be limited. In order to capture most of the structural breaks in the 13-month price dataset, the breakpoints were computed in two sequential runs. The first run involves a daily arrangement of the 13-month dataset with a minimum breakpoint distance of one week. The second run involves an hourly arrangement of the remaining days as a result of the first run. After computing the breakpoints in a run, the input dataset is divided into periods, which are compared to the most recent period in terms of the average price. In order to discard sufficiently dissimilar periods that belong to other market circumstances, the periods where the price average falls outside the interval $\mu \pm \sigma$, where μ and σ represent the most recent period's price average and standard deviation respectively, are discarded. While larger thresholds than $\mu \pm \sigma$ (e.g., $\mu \pm 2\sigma$, $\mu \pm 3\sigma$, etc.) are chosen in other contexts to discard outliers, two periods in time may belong to different market conditions even with a difference of one standard deviation. As a result, this unique manner of performing the methodology of [28] provides an efficient way of detecting structural breaks in a 13-month dataset with hourly precision, as well as discarding significantly different periods as per price behaviours.

Figure 3 depicts the resulting calibration period selection according to the structural break analysis. Whilst the left y-axis is related to Iberian electricity market prices from December 2015 up to December 2017, the shaded shape indicates the calibration periods (x-axis) selected for a certain forecasting day (right y-axis). For example, if the first day of 2017 is selected by drawing an imaginary horizontal line that crosses said day in the right y-axis (which, in this case, the line coincides with the upper border of the graph), the shaded area overlaps this imaginary line during 3 months in late 2016 and part of December 2015 according to the x-axis, which represents the calibration periods that are selected if the forecasting day is the first day of 2017. Given that early 2017 was characterised by uncommonly high prices, the selected calibration periods were much shorter than those of late 2017. Furthermore, January's peak is generally discarded from calibration datasets when forecasting days later in that year. Moreover, summer 2016 is considered while forecasting summer 2017. Therefore, the result of this algorithm eliminates periods in the past that are expected to be highly dissimilar to the forecasting period.

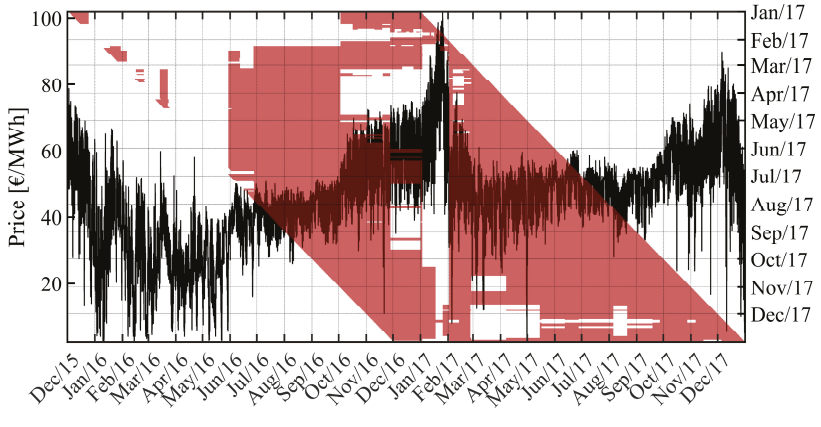


Figure 3. Structural breaks algorithm period selection.

2.2.2. Hourly Clustering

This stage seeks to determine the most relevant factors regarding the market conditions during the forecasting period. The variable with the most predictive content is the estimation of the actual price from the fundamental model, which reflects several aspects of the operations and the dynamics of the market. Although futures prices are often used as predictive variables, they were less useful here than the market-clearing prices, since they do not specify intraday effects. A variable that responds well to sudden market condition disruptions is the expected thermal gap, which represents the difference between the expected demand and the expected renewable generation from wind and solar facilities. Prices are bound to fall if the gap is low. Although the expected market-clearing prices also capture this effect, the expected thermal gap contains a higher level of short-term dynamic information and thus indicates intraday effects with higher definition. The expected temperature is also useful in order to remove periods with significantly different temperature effects.

A K-means clustering method was applied to take these three exogenous variables into account (estimated market-clearing prices, expected thermal gap and expected temperature) and relate the hours in the forecasting period to those of the training period. Given that these exogenous variables are expressed in different units and orders of magnitude, they were standardized before applying the clustering procedure. The K-means clustering application involves the identification of centroids of the values of those three variables throughout the 13-month initial dataset. Consequently, each hour in the dataset belongs to the closest centroid in terms of squared Euclidean distances in the 3D plane formed by the three variables. Depending on the predefined number of clusters, the centroids are placed so as to minimise the total quantisation error or the sum of squared Euclidean distances. Thus, a greater number of clusters lead to lower quantisation errors and higher complexity levels. In order to appropriately set the number of clusters, the K-means algorithm is computed for several numbers of clusters and, by means of a Pareto optimal frontier procedure [29], a suitable compromise between complexity level and total quantisation error is obtained. Finally, the clusters that include the hours of the forecasting period are deemed relevant and thus the hours of the input dataset that do not belong to said clusters are discarded.

The combination of these period selection algorithms is intended to discard the information pertaining to dissimilar market regimes, according to recent price behaviours and forecasted market regime indicators. Therefore, the hours that were not discarded by the structural breaks method were combined with those included by the K-means procedure, as displayed in Figure 4. The difference in Figure 4's shape with respect to that of Figure 3 is related to the hours that are only selected by the K-means method (i.e., there are hours that were chosen by both techniques). This new shaded shape is somewhat hollow given that the clustering has been performed hourly. This provides useful information as to what intraday patterns in the past are the most similar to that of the forecasting period.

The resulting calibration dataset that is shown in Figure 4 contains two sets of information: the recent dynamics such as agent strategic behaviours provided by the structural breaks method and the patterns that are driven by market fundamentals yielded by the hourly clustering technique. All in all, this combined dataset discards the information pertaining to dissimilar market regimes according to recent price behaviours and forecasted market regime indicators in an automated fashion.

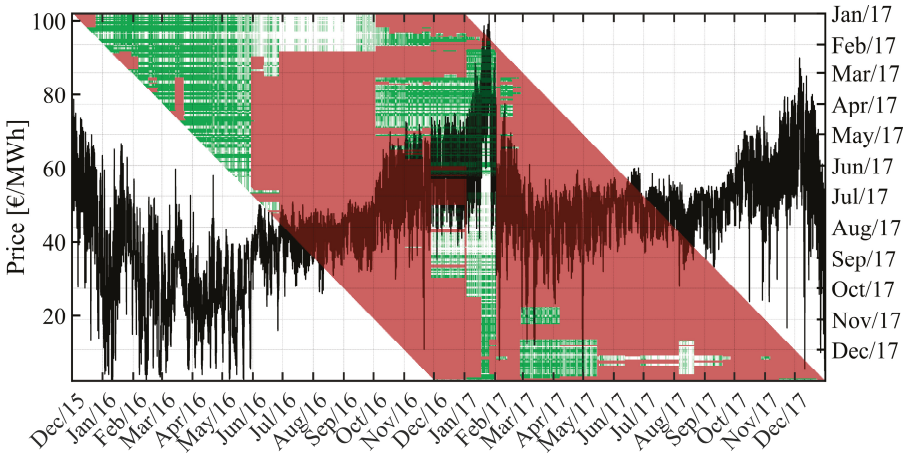


Figure 4. Addition of an hourly clustering method to the period selection methodology.

2.2.3. Neural Network Validation Set

Considering the length of the filtered dataset, a validation set is obtained following the similar days method performed in [19], which selects the top 20% of days in the most recent segment (i.e., between the most recent structural break and immediately prior to the forecasting period) as per their similarity with respect to the forecasting period in terms of daily patterns regarding exogenous variables such as expected demand.

2.3. Artificial Neural Network Model

As displayed in Figure 1, four outputs of the fundamental model are combined with common predictors to form the set of input variables for the NN model. This set consists of the following factors:

- Expected values of demand, wind and solar generation
- Expected mean temperature in the Iberian Peninsula
- Two dummy variables corresponding to working days or a Sunday/holiday, thus Saturdays would correspond to both dummy variables being false
- Actual electricity market prices with the following lags: one day, two days, one week and two weeks
- Commodity related month-ahead forward prices: API2 coal, NBP natural gas and European CO_2 emission allowances
- Day-ahead Iberian electricity market futures
- Fundamental model output variables: market-clearing prices; and coal, CCGT and hydro production levels.

This set of variables has been obtained and validated by means of a variable selection procedure based on mutual information and partial mutual information in order to analyse their dependency with respect to electricity prices and their redundancy with respect to the other explanatory variables when used to predict electricity prices. However, it is worth noting that this work's contributions are centred on the period selection methods.

Once the proposed period filtering methods have been carried out, the remaining data are used as training inputs to a NN forecasting method. The literature suggests that the most suitable NN configuration is a single hidden and output layer architecture, as stated in [30]. Another well-established choice in the literature is the Levenberg-Marquardt training algorithm. The hyperbolic tangent sigmoid activation function was utilised for the hidden layer's neurons and a pure linear activation function

has been resorted to for the output layer. However, due to the lack of consensus in the literature with regards to the number of neurons to be set in the hidden layer, several variations were tested (a range from 10 to 60 with a step of 5 neurons). The validation set mean-square error MSE of the neural networks was used in order to choose the optimal number of the hidden layer's neurons. Moreover, in order to consider the random initialisation of the weights of the NN training algorithm, a high number of replications of the NN forecasting procedure were carried out. This is also done in order to improve the likelihood of the NN training algorithm going from local to global minima.

2.4. Model Performance Metrics and Evaluation Criteria

Consistent with most forecasting research, the models are evaluated using three error metrics: mean absolute percentage error (MAPE), mean absolute error (MAE) and root-mean-square error (RMSE). Note that electricity prices can sometimes approach zero and in such cases, MAPE may approach infinite values. Nevertheless, the case study utilised in this work does not include any actual zero-price occurrences. In order to assess performance comparisons in a statistically significant manner, the Diebold-Mariano (DM) test has been carried out with a 5% significance level [31].

3. Case Studies, Results and Discussion

The case study for this work is the Iberian electricity system throughout the entire year of 2017. Early 2017 was characterised by very cold weather, low renewable energy generation, high natural gas prices and external disruptions originated by the decommissioning of nuclear power plants in France. This caused the price surge that is seen in Figure 3, which represents the 2017 maximum price of 101.99 €/MWh (up from 2016's peak of 75 €/MWh). Furthermore, a steady increase in the European CO_2 emission allowance prices began in the late summer of 2017. More specifically, these prices rose by approximately 25% throughout the year of 2017. Therefore, this case study poses a highly challenging task with disruptions and evolving changes and is, therefore, a suitable test of the methodology proposed in this paper.

First of all and as per Figure 1, the cost production model is run so as to obtain the necessary information to carry out the remaining stages of the proposed methodology. This provides the fundamental model output variables: market clearing prices as well as CCGT, hydro and coal unit generation output levels. Furthermore, given that the aim of this work is to provide forecasts for the entire year of 2017 and that the NN model's initial training dataset is of 13 months, all input variables must be made available from December 2015 up to December 2017. Once these 13 months are filtered according to the methodologies presented in the previous section, the NN model is run to provide rolling forecasts for every day of the year in 2017. Therefore, the actual forecasting horizon is of one day.

In order to specifically assess the ingredients of this methodology, the proposed hybrid forecasting model is split into stages, where each stage adds one of the techniques detailed in the previous subsection as follows:

- Stage 0: A base hybrid fundamental-econometric model without filtering any periods and variables and using 120 days of calibration data, although a limited filtering procedure in winter 2017 reduced this data length by roughly 70%. This coincides with the Proposed Model 2 that was presented in [19].
- Stage 1: 13 months of calibration data are used and these are filtered via the structural breaks technique.
- Stage 2: The K-means hourly clustering procedure is added to the calibration period selection method.

These models will be referred to as PMS_i, i.e., the Proposed Model at its Stage *i*. As in other works, for instance [11], the performance of these models has been analysed for every season of the year and compared with that of six other electricity price forecasting models, some of which correspond to well-established methodologies in the literature. The first chosen benchmark model (Benchmark 1 or

BM_1) consists of the proposed simple average of [19] between the forecasts of a pure NN model and the base hybrid fundamental-econometric model (PMS_0). Benchmark two (BM_2) only involves this pure NN model that utilises the same input variables as BM_1/PMS_0 (except those pertaining to the fundamental model) and the same calibration window. This 120-day window includes four months within the 13-month window established in this work, more specifically, the 13th, 12th, 2nd and 1st month prior to the forecasting day [19].

The third benchmark model (BM_3) is related to a linear regression model with several autoregressive terms and exogenous components. This ARX model, introduced in [32] and recently utilised in [15], includes a logarithmic transform that was modified so as to account for the lower price cap of zero in the Iberian electricity market:

$$p_{d,h} = \beta_{h,1}p_{d-1,h} + \beta_{h,2}p_{d-2,h} + \beta_{h,3}p_{d-7,h} + \beta_{h,4}p_{d-1}^{min} + \beta_{h,5}z_{d,h} + \beta_{h,6}D_{Sat} + \beta_{h,7}D_{Sun} + \beta_{h,8}D_{Mon} + \varepsilon_{d,h} \quad (1)$$

$$n_{d,h} = \frac{\left(P_{d,h} - \mu_T\right)}{\sigma_T} \quad (2)$$

$$p_{d,h} = sgn\left(n_{d,h}\right)\left[log\left(n_{d,h} + \frac{1}{c}\right) + log(c)\right] \quad (3)$$

According to Equation (1), the log-price at day d and hour h, $p_{d,h}$, depends upon: lagged prices, such as $p_{d-1,h}$; the minimum log-price during the 24 h of day d minus one (i.e., p_{d-1}^{min}); the load forecast ($z_{d,h}$); and three dummy variables that specify if day d is a Saturday, Sunday or Monday. The logarithmic transform of Equation (3) is the mirror-log transform, where prices, $P_{d,h}$, are first normalised in Equation (2) across the training period T, and the parameter c was set to 1/3 according to the application presented in [33]. The transformations that have been applied to the explanatory variables regarding past electricity prices stabilise their variance and ensure stationarity, as observed in [33]. Three months prior to the forecasting day were used as calibration data. The next benchmark (BM_4) is the extension of BM_3 as per the work presented in [15], which performs a weighted average of forecasts from the ARX model of Equation (1) across the following calibration windows (in terms of days prior to the forecast day): 56, 84, 112, 714, 721 and 728 days. The weights of these six forecasts are computed by means of an inverse MAE weighting procedure when testing the ARX models on the day prior to the forecast day.

Benchmark five (BM_5) is related to a SARIMAX model, whose SARIMA noise presents the following notation: $SARIMA(1,0,0)_{168}(1,0,2)_{24}(1,0,0)_1$. A daily and weekly seasonality was considered, as well as the expected demand as an exogenous variable. This model was created following [34,35], with the Box-Jenkins methodology. Furthermore, the Box-Cox transformation was used to stabilise the price variance [36]. The final benchmark (BM_6) is related to a simple naïve approach that sets the forecast to the actual electricity price value corresponding to the previous week.

The proposed model, in all of its stages as well as the six benchmark models, have been tested for every day of the year in 2017 and their error measures across the four seasons of 2017 are shown in Table 1. Furthermore, the average calibration dataset windows for each of the models involving a NN forecasting technique are displayed in Table 2.

Table 1. Forecast error measures for the proposed and benchmark models: MAPE (%), MAE & RMSE (€/MWh).

Model	Winter			Spring			Summer			Autumn			Entire 2017		
	MAPE	**MAE**	**RMSE**	**MAPE**	**MAE**	**RMSE**	**MAPE**	**MAE**	**RMSE**	**MAPE**	**MAE**	**RMSE**	**MAPE**	**MAE**	**RMSE**
PMS_0—Base model [19]	11.68	4.756	5.479	8.106	2.882	3.407	**4.450**	2.070	2.517	6.812	3.453	4.129	7.744	3.282	3.874
PMS_1—Structural breaks	11.02	4.266	4.917	7.706	2.487	2.960	4.501	2.063	2.509	**6.237**	3.150	3.820	7.348	2.984	3.543
PMS_2—Hourly clustering	**10.05**	**4.133**	**4.785**	**7.303**	**2.433**	**2.892**	4.467	**2.045**	**2.504**	6.284	**3.139**	**3.818**	**7.012**	**2.930**	**3.492**
BM_1—$\frac{1}{2}(PMS_0 + BM_2)$ [19]	11.30	4.610	5.334	7.902	2.832	3.349	4.477	2.079	2.525	6.756	3.409	4.069	7.591	3.224	3.810
BM_2—Pure NN [19]	11.12	4.562	5.308	7.804	2.826	3.342	4.605	2.136	2.588	6.834	3.440	4.089	7.575	3.233	3.823
BM_3—ARX [32]	16.79	6.838	7.809	13.58	4.765	5.552	7.153	3.262	3.885	10.51	5.066	6.055	11.99	4.972	5.814
BM_4—W. ARX [15]	16.27	6.390	7.314	13.21	4.500	5.268	7.015	3.211	3.857	10.14	4.880	5.874	11.64	4.736	5.568
BM_5—SARIMAX	15.06	8.113	10.84	9.293	4.150	5.585	5.097	2.473	4.531	7.654	4.454	4.959	9.248	4.780	6.460
BM_6—Naïve approach	25.93	10.53	11.48	17.55	6.225	7.092	9.343	4.266	5.030	12.82	6.387	7.567	16.37	6.828	7.773

Table 2. Average calibration length window of the NN models (days).

Model	Winter	Spring	Summer	Autumn	Overall
PMS_0, BM_1 & BM_2 [19]	36.67	120.0	120.0	120.0	99.17
PMS_1 (Figure 3)	152.9	237.0	324.7	300.5	254.2
PMS_2 (Figure 4)	288.8	324.5	344.7	348.1	326.7

Compared with the base model of PMS_0, the implementation of the structural breaks technique increased the NN training set by well beyond the predefined number of 120 days that was established in [19]. The reason behind the reduced dataset during the 2017 winter is due to its high instability, and it was observed in [19] that a reduction of the 120-day dataset provided useful results. This agrees with the rationale that consists of increasing adaptability on unstable periods by reducing the calibration window in order to remove structural breaks from the input dataset. However, in this work, an average dataset of 152.9 days yields lower forecasting errors. Furthermore, PMS_1 discards most of the previous winter, which is considerably different from the 2017 winter as depicted in Figure 3. This also seems to be the case for spring, as the 2016 spring yielded approximately twice as much hydro generation as the 2017 spring. In general, the structural breaks algorithm provides a generally lower error throughout 2017. However, summer 2017 appears to be the exception, where prices are relatively stable and thus, it lacks room for improvement, as proven by the generally low errors yielded by most models.

Furthermore, the lengthening of PMS_1's calibration dataset with the hourly clustering technique of PMS_2 further reduces the overall forecasting error. This is more notable during winter, where the average calibration dataset is greatly increased to 288.8 days. As for the other seasons, a calibration dataset of approximately one year proves to be beneficial for electricity price forecasting with NN models even with the hourly arrangement and does not seem to cause any overfitting issues. Although PMS_2 yields a lower error overall, the statistical significance of these error measures must be verified in order to confirm its superiority against its competitors, especially the highest-ranked models according to Table 1. Therefore, a DM test was carried out for PMS_2 against every other model. The DM test statistic is evaluated with a 5% significance level, such that a DM statistic < -1.96 implies significant outperformance. The results of the DM test statistic are shown in Table 3.

Table 3. DM statistic values of PMS2 against every other model.

Model Comparison	Winter	Spring	Summer	Autumn	Entire 2017
PMS_2 vs. PMS_0	−8.834	−12.31	**−0.975**	−6.787	−14.75
PMS_2 vs. PMS_1	−2.903	−3.199	**−1.833**	**−0.436**	−3.917
PMS_2 vs. BM_1	−6.528	−8.042	−2.883	−5.726	−11.29
PMS_2 vs. BM_2	−6.316	−11.36	−3.262	−6.877	−13.18
PMS_2 vs. BM_3	−21.54	−28.70	−22.94	−21.69	−44.76
PMS_2 vs. BM_4	−18.78	−26.37	−21.18	−19.59	−40.67
PMS_2 vs. BM_5	−21.21	−17.96	−4.454	−16.85	−29.70
PMS_2 vs. BM_6	−34.44	−33.17	−27.75	−28.66	−58.82

The three values in bold indicate the three occasions that PMS_2 was unable to significantly outperform. The comparison with PMS_0 suggests that the increase in calibration data window lengths does not significantly contribute to summer forecasts, albeit not detrimental to the accuracy. This may also imply that a robust calibration period selection is not highly crucial in such a stable market regime. Therefore, the same conclusion can be drawn from the summer comparison with PMS_1. Furthermore, the DM statistic value in autumn when compared with PMS_1 may indicate that the information provided by the hourly clustering method is not significantly different than that provided by the structural breaks technique. However, these values indicate that PMS_2 is significantly outperforming all other models throughout the year in 2017.

4. Conclusions

This research presents a novel short-term hybrid electricity price forecasting methodology which is comprised of three main elements: a cost-production optimisation model, a sophisticated period filtering approach and a neural network (NN) model. These three elements were utilised sequentially with the calibration selection procedure as the main focus of this work. Given a forecasting day, the structural patterns in actual prices corresponding to the 13 months prior to that day are analysed and those deemed unimportant were discarded. A K-means clustering method was also applied to relate the moments in the prior 13 months to the forecasting day in terms of the estimated fundamental market-clearing prices, expected thermal gap and expected mean temperatures in the Iberian Peninsula. The key innovation of this approach is to move beyond the conventional forecasting principles which suggest that adaptive methods should place most emphasis upon recent data and that regime-switching should likewise model transitions from the latest regime. The approach recognises that the most relevant dataset in the episodic, recurrent nature of electricity dynamics may not be the most recent. Another unique feature of this methodology is the definition of a calibration period that is not driven by heuristic assumptions or any other specific predefinitions.

The results and analyses indicate the following. The combination of structural break analysis and hourly clustering provides a dynamic calibration period appropriate for the forecasting model estimation. In validation, this sophisticated training window selection for the NN model yields appealing results in every market circumstance present in the relatively challenging case study of the Iberian electricity market of 2017. The period selection technique is more selective in volatile market conditions, such as early 2017, albeit providing a considerably longer training window length than other works which claim that employing much shorter calibration windows is most suitable in these situations. In addition, the proposed methodology proves most useful during volatile periods, whilst the accuracy is marginally increased in stable market regimes, such as summer 2017.

Overall, this short-term fundamental-econometric electricity price forecasting model, which features a unique hybridisation approach, has yielded appropriate results when applied to a real-size electricity system with complex price dynamics, such as the Iberian power exchange of 2017. Furthermore, the performance of this proposal is superior to that of other benchmark models. Although only one market has been chosen as the case study, the results may be generalised for other markets due to the high number of special circumstances that the Iberian power system experienced throughout the year 2017. However, there seems to be room for improvement regarding the utilised structural breaks period selection algorithm, as it is highly challenging to ascertain a convenient compromise between accuracy and computational burden. Transient spikes for example cannot all be considered structural breaks, yet it may be beneficial if these are more adequately considered in a computationally feasible manner. Furthermore, more complex neural network topologies may be tested in conjunction with this calibration period selection methodology, such as convoluted or LSTM neural networks.

Author Contributions: Conceptualization, R.A.d.M., D.W.B., A.B. and J.R.; Data curation, R.A.d.M.; Formal analysis, R.A.d.M.; Investigation, R.A.d.M.; Methodology, R.A.d.M., D.W.B., A.B. and J.R.; Software, R.A.d.M.; Supervision, D.W.B., A.B. and J.R.; Validation, R.A.d.M. and A.B.; Visualization, R.A.d.M.; Writing—original draft, R.A.d.M., D.W.B. and A.B.; Writing—review & editing, R.A.d.M., D.W.B., A.B. and J.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bello, A.; Reneses, J.; Muñoz, A.; Delgadillo, A. Probabilistic forecasting of hourly electricity prices in the medium-term using spatial interpolation techniques. *Int. J. Forecast.* **2016**, *32*, 966–980. [CrossRef]
2. Contreras, J.; Espínola, R.; Nogales, F.J.; Conejo, A.J. ARIMA models to predict next-day electricity prices. *IEEE Trans. Power Syst.* **2003**, *18*, 1014–1020. [CrossRef]

3. Cipriano, A.; Lira, F.; Núñez, F.; Muñoz, C. Short-term forecasting of electricity prices in the Colombian electricity market. *IET Gener. Transm. Distrib.* **2009**, *3*, 980–986. [CrossRef]
4. Sánchez De La Nieta, A.A.; González, V.; Contreras, J. Portfolio decision of short-term electricity forecasted prices through stochastic programming. *Energies* **2016**, *9*, 1069. [CrossRef]
5. Catalão, J.P.S.; Mariano, S.J.P.S.; Mendes, V.M.F.; Ferreira, L.A.F.M. Short-term electricity prices forecasting in a competitive market: A neural network approach. *Electr. Power Syst. Res.* **2007**, *77*, 1297–1304. [CrossRef]
6. Amjady, N.; Daraeepour, A.; Keynia, F. Day-ahead electricity price forecasting by modified relief algorithm and hybrid neural network. *IET Gener. Transm. Distrib.* **2010**, *4*, 432. [CrossRef]
7. Yan, X.; Chowdhury, N.A. Mid-term electricity market clearing price forecasting using multiple support vector machine. *IET Gener. Transm. Distrib.* **2014**, *8*, 1572–1582. [CrossRef]
8. Monteiro, C.; Ramirez-Rosado, I.J.; Fernandez-Jimenez, L.A.; Conde, P. Short-term price forecasting models based on artificial neural networks for intraday sessions in the Iberian electricity market. *Energies* **2016**, *9*, 721. [CrossRef]
9. Hong, T.; Pinson, P.; Fan, S.; Zareipour, H.; Troccoli, A.; Hyndman, R.J. Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond. *Int. J. Forecast.* **2016**, *32*, 896–913. [CrossRef]
10. Tahmasebifar, R.; Sheikh-El-Eslami, M.K.; Kheirollahi, R. Point and interval forecasting of real-time and day-ahead electricity prices by a novel hybrid approach. *IET Gener. Transm. Distrib.* **2017**, *11*, 2173–2183. [CrossRef]
11. Bento, P.M.R.; Pombo, J.A.N.; Calado, M.R.A.; Mariano, S.J.P.S. A bat optimized neural network and wavelet transform approach for short-term price forecasting. *Appl. Energy* **2018**, *210*, 88–97. [CrossRef]
12. Qiao, W.; Yang, Z. Forecast the electricity price of U.S. using a wavelet transform-based hybrid model. *Energy* **2020**, *193*, 116704. [CrossRef]
13. Chang, Z.; Zhang, Y.; Chen, W. Electricity price prediction based on hybrid model of adam optimized LSTM neural network and wavelet transform. *Energy* **2019**, *187*, 115804. [CrossRef]
14. Pesaran, M.H.; Timmermann, A. Selection of estimation window in the presence of breaks. *J. Econom.* **2007**, *137*, 134–161. [CrossRef]
15. Marcjasz, G.; Serafin, T.; Weron, R. Selection of calibration windows for day-ahead electricity price forecasting. *Energies* **2018**, *11*, 2364. [CrossRef]
16. Avci, E.; Ketter, W.; van Heck, E. Managing electricity price modeling risk via ensemble forecasting: The case of Turkey. *Energy Policy* **2018**, *123*, 390–403. [CrossRef]
17. Chen, D.; Bunn, D.W. Analysis of the nonlinear response of electricity prices to fundamental and strategic factors. *IEEE Trans. Power Syst.* **2010**, *25*, 595–606. [CrossRef]
18. Paraschiv, F.; Fleten, S.E.; Schürle, M. A spot-forward model for electricity prices with regime shifts. *Energy Econ.* **2015**, *47*, 142–153. [CrossRef]
19. De Marcos, R.A.; Bello, A.; Reneses, J. Short-term electricity price forecasting with a composite fundamental-econometric hybrid methodology. *Energies* **2019**, *12*, 1067. [CrossRef]
20. Mandal, P.; Senjyu, T.; Funabashi, T. Neural networks approach to forecast several hour ahead electricity prices and loads in deregulated market. *Energy Convers. Manag.* **2006**, *47*, 2128–2142. [CrossRef]
21. Karakatsani, N.V.; Bunn, D.W. Forecasting electricity prices: The impact of fundamentals and time-varying coefficients. *Int. J. Forecast.* **2008**, *24*, 764–785. [CrossRef]
22. Bello, A.; Bunn, D.W.; Reneses, J.; Muñoz, A. Parametric density recalibration of a fundamental market model to forecast electricity prices. *Energies* **2016**, *9*, 959. [CrossRef]
23. Bello, A.; Bunn, D.W.; Reneses, J.; Muñoz, A. Medium-term probabilistic forecasting of electricity prices: A hybrid approach. *IEEE Trans. Power Syst.* **2017**, *32*, 334–343. [CrossRef]
24. González, V.; Contreras, J.; Bunn, D.W. Forecasting power prices using a hybrid fundamental-econometric model. *IEEE Trans. Power Syst.* **2012**, *27*, 363–372. [CrossRef]
25. De Marcos, R.A.; Bello, A.; Reneses, J. Electricity price forecasting in the short term hybridising fundamental and econometric modelling. *Electr. Power Syst. Res.* **2019**, *167*, 240–251. [CrossRef]
26. Transparency Platform of the Spanish System Operator. Available online: https://www.esios.ree.es/en (accessed on 13 November 2018).
27. Transparency Platform of the ENTSO-E. Available online: https://transparency.entsoe.eu/ (accessed on 11 September 2018).

28. Zeileis, A.; Kleiber, C.; Walter, K.; Hornik, K. Testing and dating of structural changes in practice. *Comput. Stat. Data Anal.* **2003**, *44*, 109–123. [CrossRef]
29. Cheikh, M.; Jarboui, B.; Loukil, T.; Siarry, P. A method for selecting Pareto optimal solutions in multiobjective optimization. *J. Inform. Math. Sci.* **2010**, *2*, 51–62.
30. Bello, A.; Reneses, J.; Muñoz, A. Medium-term probabilistic forecasting of extremely low prices in electricity markets: Application to the Spanish case. *Energies* **2016**, *9*, 193. [CrossRef]
31. Diebold, F.X.; Mariano, R.S. Comparing predictive accuracy. *J. Bus. Econ. Stat.* **1995**, *13*, 253–263. [CrossRef]
32. Weron, R.; Misiorek, A. Short-term electricity price forecasting with time series models: A review and evaluation. In *Complex Electricity Markets*; IEPŁ & SEP: Łódz, Poland, 2006; pp. 231–254.
33. Uniejewski, B.; Weron, R.; Ziel, F. Variance stabilizing transformations for electricity spot price forecasting. *IEEE Trans. Power Syst.* **2017**, *33*, 2219–2229. [CrossRef]
34. Box, G.E.P.; Jenkins, G.; Reinsel, G.C. *Time Series Analysis—Forecasting and Control, Fourth Edition*; John Wiley & Sons: Hoboken, NJ, USA, 2008. [CrossRef]
35. Pankratz, A. Building dynamic regression models: Model identification. In *Forecasting with Dynamic Regression Models*; John Wiley & Sons: Hoboken, NJ, USA, 2012; Volume 935, pp. 167–201. [CrossRef]
36. Box, G.E.P.; Cox, D.R. An analysis of transformations. *J. Stat. Soc. Ser. B.* **1964**, *26*, 211–252. [CrossRef]

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

Energies Editorial Office
E-mail: energies@mdpi.com
www.mdpi.com/journal/energies

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland

Tel: +41 61 683 77 34
Fax: +41 61 302 89 18

www.mdpi.com

ISBN 978-3-03943-822-8