ifo
WORKING
PAPERS

357
2021
June 2021

Unraveling the Black Box of Power Market Models

Mathias Mier, Valeriya Azarova

ifo INSTITUTE
Leibniz Institute for Economic Research
at the University of Munich

Imprint:

ifo Working Papers
Publisher and distributor: ifo Institute – Leibniz Institute for Economic Research at the University of Munich
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49(0)89 9224 0, Telefax +49(0)89 985369, email ifo@ifo.de
www.ifo.de

An electronic version of the paper may be downloaded from the ifo website:
www.ifo.de

ifo Working Paper No. 357

Unraveling the Black Box of Power Market Models

Abstract

Detailed numerical models of power markets with millions of variables and equations are often perceived as black boxes, because differences in results cannot be traced back to single equations or assumptions, respectively. We unravel parts of those black box by determining the impact of different investment cost specifications including the role of varying discount and interest rates. We further expand our analysis to the impact of simplifications strategies (foresight, spatial resolution, temporal resolution) that are applied to contain numerical feasibility of such models. The choice of investment cost modeling (and related discount and interest rates) has the highest impact on results. Increasing or decreasing, respectively, complexity in turn, has only minor impacts. Our findings questions the current focus of the literature on complexity of power market models neglecting the most relevant factor, which is the choice of handling investment costs.

JEL code: C61, C68, Q40, Q41
Keywords: Energy system modeling, power market modeling, investment behavior, firm behavior, spatial resolution, temporal resolution, decarbonization

Mathias Mier*
ifo Institute – Leibniz Institute for Economic Research
at the University of Munich
Poschingerstr. 5
81679 Munich, Germany
mier@ifo.de

Valeriya Azarova
ifo Institute – Leibniz Institute for Economic Research
at the University of Munich
Poschingerstr. 5
81679 Munich, Germany
azarova@ifo.de

* Corresponding author.

Unraveling the Black Box of Power Market Models

Mathias Mier[a], Valeriya Azarova[a]

[a]*ifo Institute for Economic Research at the University of Munich*

Abstract

Detailed numerical models of power markets with millions of variables and equations are often perceived as black boxes, because differences in results cannot be traced back to single equations or assumptions, respectively. We unravel parts of those black box by determining the impact of different investment cost specifications including the role of varying discount and interest rates. We further expand our analysis to the impact of simplifications strategies (foresight, spatial resolution, temporal resolution) that are applied to contain numerical feasibility of such models. The choice of investment cost modeling (and related discount and interest rates) has the highest impact on results. Increasing or decreasing, respectively, complexity in turn, has only minor impacts. Our findings questions the current focus of the literature on complexity of power market models neglecting the most relevant factor, which is the choice of handling investment costs.

Keywords: Energy system modeling, Power market modeling, Investment behavior, Firm behavior, Spatial resolution, Temporal resolution, Decarbonization

JEL Classification: C61, C68, Q40, Q41

1. Introduction

Power market models are a widely decision and analysis support tools used to provide policy recommendations and elaborate scenarios of future energy systems and support to national and regional energy planning and policy-making (Cao et al., 2016). To cite some examples, several TIMES-based optimization models are being used by international organizations and governmental institutions to provide insights on how to reach national and international climate goals (Merkel et al., 2014). DICE

*Corresponding author: mier@ifo.de

model is used to elaborate an efficient strategy for coping with the threat of global warming. PRIMES, simulation-based model, has been used by European Commission to analyze of taxation policy and emissions trading and includes a detailed representation of energy technologies in the EU (Connolly et al., 2010).

Improved computational capabilities on one hand, and increased demand in power market modeling with a growing number of climate change abatement polices and initiatives being recently discussed or adopted (the European Green Deal, in Europe, Clean Energy and Pollution Reduction Act in USA, California, Climate Change Response Amendment Act, New Zealand), triggered a boost in the number and diversity of models. According to Amerighi et al. (2010), various European research institutions had developed more than 68 models only for the European power market.

However, despite the research community's overall tendency towards more transparency and open access publishing, the majority of available models are still so-called "black boxes". In other words, closed for outside researchers for replication or comparison, usually providing very little to no descriptions or source codes. This closeness of models is often mentioned as one of the key reasons for conflicting results reported for similar policy analyses produced by different models (Müller et al., 2018, Morrison, 2018). Furthermore, lack of transparency and formal guidance on model design and missing evidence on the impact of key configurations on models' results, leads to the absence of a well-established best practice approach for inclusion of necessary parameters and constraints for a specific analysis and research question (DeCarolis et al., 2017).

Hence, a high variety of model designs and assumptions and a lack of publicly available codes or in some cases even detailed descriptions, impede clear comparison of the models. This complicates the choice of a suitable model configuration for a specific research question, as well as policy recommendation provision based on the model, and impedes synergy between models.

In this paper, we use the example of EUREGEN, the European power market model, to demonstrate how changing a model's design namely (1) handling of investment cost, (2) foresight, (3) spatial resolution, and (4) temporal resolution, impacts four key outcomes of the model: i. technology mix, ii. storage, iii. transmission, iv. emissions and cost. We start with comparing three different investment cost specifications, including the role of discounting, and varying interest rates. We then analyze the impact of different levels of foresight under four different scenarios of CO2 price evolution. Thirdly, we compare outcomes of six spatial resolutions under four different scenarios of transmission boundaries. Finally, we analyze eight temporal resolutions and evaluate the role of storage under six different scenarios of storage cost.

Additionally, increasing complexity of energy systems triggers models and research questions to gain in complexity (Bale et al., 2015). Consequently, the models need to carefully balance between complexity and accuracy of the results (Ridha et al., 2020, Priesmann et al., 2019). Thus, a trade-off between high resolution in time, space, techno-economic detail and sector coupling becomes one of the pressing key challenges of future models (Prina et al., 2020). At the same time, existing (and future) power market models require a high degree of detail in representing the underlying system to serve their main aim of policy recommendation provision. Owing to the underlying complexity of the power market, this requires many variables, parameters, restrictions, and assumptions (Babrowski et al., 2014).In this paper, we complement available literature by demonstrating strategies of model simplification and a reduction of computational times, as well as their impact on model's outcomes. In this regard, intra-model comparison conducted in this paper can be used as a guideline making it possible to design the model specific to the analyzed research question.

In this way, our analysis contributes to greater transparency in power market modeling and serves as a milestone in establishing best practice in power market models. Thus, elaborating guidelines for appropriate model features, which will help to conduct and refine analysis and improve models' comparability.

We find that configuration of the investment behavior has the highest impact on the model outcomes in terms of both installed capacity and system cost and CO2 emissions. Hence, it is advised to carefully select an investment behavior representation strategy, allowing the most accurate reflection of the market context. The other tested model specifications including model horizon, and spatial and temporal resolution are also evaluated and discussed in the respective sections, followed by a conclusion in section 6.

2. Investment Cost

The specification of investment cost is one of the key features in power market models, despite only receiving scant coverage in the literature. For instance, dynELMOD (Gerbaulet and Lorenz, 2017) is configured to include investment on an annuity basis. When investments occur, the entire cost is not accounted for in the year of investment, however, the to-be-paid annuities are tracked over the investment's economic lifetime. This also takes account of the remaining model periods. Another possibility would be to use weighted average cost of capital (WACC) as applied by (Bachner et al., 2019) in their analysis. In this section, we show how various strategies of handling the investment costs including annuity (with impact of varying interest rates), WACC and applied in the base version of a EUREGEN model overall investment in the respective period (total investment cost, see detailed formulation below), impact the model's outcomes. The essential objective of this analysis is not to show which investment cost implementation approach within the power market models is superior, but rather to demonstrate the applicability of each of those for specific research questions. This enables a reflection on investment behavior for better targeted policy implications and policy instruments in this context.

2.1. Investment Cost Specifications

Consider technologies j (e.g., wind onshore), regions r (e.g., Germany), time periods $t = 2015, 2020, ..., 2050$, and the period of installation $v = 1960, 1965, ..., 2050$. We use subscript j, r to denote variables and parameters and parentheses for periods v, t, e.g., $Q_{jr}(v)$ is the capacity installed in period v and $C_{jr}(v)$ the constant unit cost. The discount factor δ follows from the discount rate ν and reflects that each period t accounts for $t_{step} = 5$ years, e.g.,[1]

$$\delta(t) \quad = \quad \frac{(1+\nu)^{t_{step}} - 1}{\nu(1+\nu)^{t-t_{base}}}, \tag{1}$$

where $t_{base} = 2015$ serves as focal point.

Normal specification.. The *normal*specification considers all investment cost in the period of investment. The objective is thus given by:

$$\min_{\mathbf{Q},...} \quad \textstyle\sum_t \delta(t) \sum_r \quad \left[\frac{1}{t_{step}} \sum_j \sum_{v=t} Q_{jr}(v)\, C_{jr}(v) \times \Gamma_{jr}(v,t) + ... \right], \tag{2}$$

where $\mathbf{Q}$ is the vector of investment decisions. $Q_{jr}(v)\, C_{jr}(v)$ are direct cost of investing into a technology and Γ is the end-effect, which reflects that the depreciation time of an investment might expand beyond the model horizon, e.g.,

$$\Gamma_{jr}(v,t) \quad = \quad \frac{\sum_t \delta(t)\, \Lambda_{ir}(v,t)}{\sum_{t_{long}} \delta(t_{long})\, \Lambda_{ir}(v, t_{long})}, \tag{3}$$

[1] 2020 reflects the time period 2016 to 2020, 2025 reflects 2021 to 2025, ...

where t_{long} reflects an unconstrained time horizon. Λ is a binary parameter that takes the value 1 when the investment is still under depreciation and 0 else, e.g.,

$$\Lambda_{jr}(v,t) = \begin{cases} 1 & \text{if } t \leq v + t_{jr,depr}(v) \\ 0 & \text{if } t > v + t_{jr,depr}(v) \end{cases}, \tag{4}$$

where $t_{jr,depr}(v)$ is the depreciation time of an investment.[2]

Annuity specification.. The *annuity* specification assumes that an investment is financed by loan capital only. The annuity reflects interests (i is the interest rate) and repayment, e.g.,

$$A_{jr}(v) = \frac{i(1+i)^{t_{jr,depr}(v)}}{(1+i)} - 1. \tag{5}$$

Investments cause a stream of cost over the entire deprecation time of the respective investment. The underlying objective becomes:

$$\min_{\mathbf{Q},\ldots} \quad \textstyle\sum_t \delta(t) \sum_r \quad \left[\displaystyle\sum_j \sum_{v \leq t} Q_{jr}^{new}(v) C_{jr}(v) \times \mathit{\Lambda}_{jr}(v,t) A_{jr}(v) + \ldots \right]. \tag{6}$$

Capital cost specification.. The *capital cost* specification assumes that a capital stock is subject to capital cost, best reflected by the weighted average cost of capital $WACC$. The difference to the annuity approach is that the depreciation time of an investment does not matter for the height of the cost, so that the objective is:

$$\min_{\mathbf{Q},\ldots} \quad \textstyle\sum_t \delta(t) \sum_r \quad \left[\displaystyle\sum_j \sum_{v \leq t} Q_{ir}(v) C_{jr}(v) \times \mathit{\Lambda}_{jr}(v,t) WACC + \ldots \right]. \tag{7}$$

Illustrative example.. Consider a wind turbine investment (100% are the installation cost) with a depreciation time of 25 years. Installing a wind turbine in 2040 translates into using it in three periods (15 years) of the model horizon (until 2050). This

[2]The installation period v reflects potential technological progress with respect to lifetime and also depreciation time. It might also reflect changing investor behavior.

translates into an end-effect of 78.16% given a discount rate of 7% and 60% when neglecting discounting. Note that a wind turbine investment in 2020 leads to end-effects of 100% for both with and without discounting. The annuity of the 2040 turbine is 8.58%. This annuity needs to be paid every year and thus the investor pays $5 \times 8.58 = 42.9\%$ in 2040, 2045, and 2050, accumulating in far more cost ((95.29% with and 128.7% without discounting) than in the normal specification (78.16% or 60%, respectively). The WACC specification just takes 7% instead of the above mentioned 8.58%, leading to an overall cost of the 2040 wind turbine of 77.75% (with discounting) or 105% (without discounting), respectively. Without discounting, the 2020 wind turbine investment results in costs of $5 \times 42.9 = 214.5\%$ in the annuity specification and of 175% in the WACC specification. Moreover, the WACC specification deviates less from the annuity when looking at investment with a longer depreciation time, such as nuclear power (40 years). The annuity then reduces to 7.5% and the annuity specification is then quite close to the WACC one in terms of an investment's profitability.

2.2. Impact of Investment Cost Specifications

We now present results on how the three investment cost specifications perform with and without discounting cashflows. Figure 1 demonstrates the evolution of installed capacity by technology types in the upmost part; stored energy by technology types in the second part; net trade capacities (NTC, light blue bars with scale on left axis) and transfers (blue triangles with scale on right axis) in the third part; as well as CO2 emissions (grey bars with scale on left axis) and system cost (orange diamonds with scale on right axis) for the three different specifications of investment costs with and without discounting from 2015 (base year) to 2050 (end of the model horizon). 2015 serves as benchmark for comparing specifications because there are no endogenous investments in this period. From 2020 onwards, values differ fundamentally between specifications.

We start with installed capacities and the specifications that consider discounting (normal, WACC, annuity). The *annuity* specification serves as benchmark for all relative differences. The annuity specification results in the lowest aggregate capacity levels, with the normal specification being the highest in the medium-term (+15.1% in 2020, +9.1% in 2035, +7.2% in 2050), and the WACC specification the highest in the long-term (+9.6% in 2020, +7.2% in 2035, +9.3% in 2050). With respect to the specific technologies enhanced by each of the specifications, we find that the exploitation of wind onshore in the normal (WACC) specification is 63.5% higher (40.4%) in 2020, 35.2% (26.9%) higher in 2035, and still 7.2% (8.4%) higher in 2050. There are no differences for solar power in 2020. Differences remain negligible in

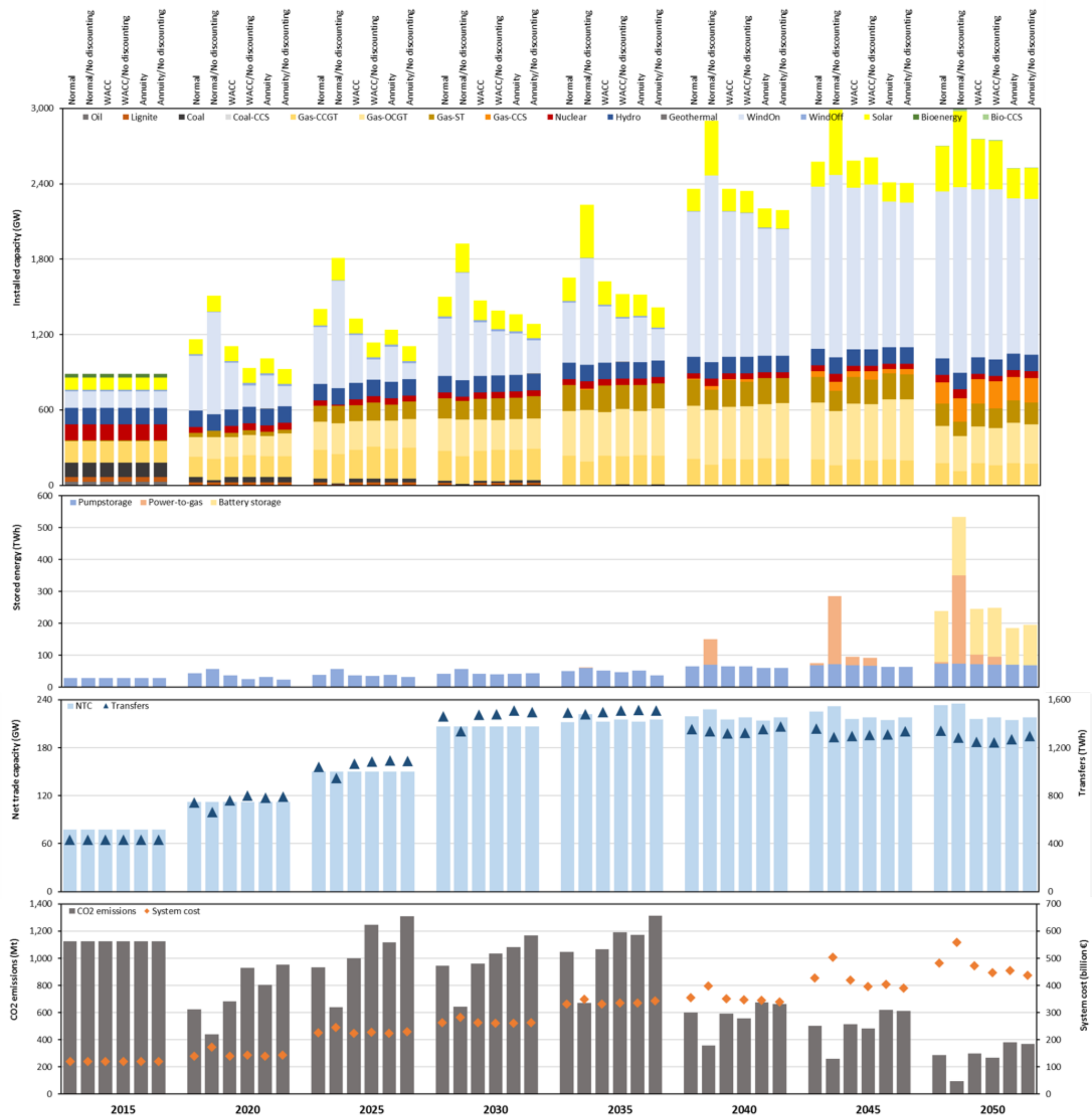
2035 (12.7% or 12.6%, respectively), but become fundamental in 2050 (+52.9% or +67.6%, respectively). Differences for gas power technologies are insignificant until 2045 (below 2% absolutely), but the normal (WACC) specification leads to 5.1% (2.1%) lower gas capacity in 2050.[3]

Now consider the specifications that neglect discounting. Now, the respective specification with discounting serves as a benchmark for relative values. Neglecting discounting leads to 15.8% (8.3%) lower capacity levels in 2020 and 6.3% (6.7%) lower ones in 2035 for the WACC (annuity) specification. From 2040 onwards, difference are below 1% in absolute and thus negligible. We observe the opposite for the normal specifications. No discounting results in 29.8% higher capacity levels in 2020. This value persists until 2035 (35%) and decreases to 10.7% in 2050. Interestingly, the normal specification differs regarding the direction of the capacity level. Under WACC and annuity specifications, capacity levels are lower, whereas capacity levels are higher under the normal specification when neglecting discounting. Considering specific technologies, we find that gas power tends to be higher in the short-term (11.9% for normal, 5.5% for WACC, 4.1% for annuity) and the mid-term for all three specification when neglecting discounting. This changes in 2035 (normal) or 2040 (WACC, annuity), so that gas power is 15.4% (normal), 1.7% (WACC), or 1.1% (annuity), respectively, lower in 2050. Wind onshore patterns are reversed for WACC and annuity (-54.3% or -38.9% in 2020 and 0.9% or 0.4% in 2050, respectively). In turn, for the normal specification, we already observe 87.3% higher wind onshore capacity in 2020. This prevails until 2035 (77.5%) and reduces to 11.2% higher capacity in 2050. Whereas differences in solar power are negligible for WACC and annuity specifications, solar power is fundamentally higher in the normal specification without discounting. Differences reach 158.8% higher capacity in 2045 but decrease slightly in 2050 (67.4%).

The behavior of the normal specification without discounting is also reflected in stored energy (second part of Figure 1), transmission (third part), and CO2 emissions and system cost (fourth and lowest parts). The fundamentally higher wind and solar capacity leads to a more effective use of existing pump-storage capacities, the application of power-to-gas from 2040 onwards, along with relatively high battery usage in 2050. The higher amount of stored energy is accompanied by lower transfers (blue triangle with scale on the right) from 2020 to 2030, whereas net trade capacities (NTC, blue bars with scale on the left) are comparable. Remaining differences for

[3]We refrain from presenting detailed outcomes for oil, lignite, coal, coal-CCS, nuclear, hydro, wind offshore, geothermal, bioenergy, and bio-CCS technologies because their shares do not drive the overall amount of capacity.

Figure 1: Results for different investment cost specifications



Evolution of installed capacity, storage, NTC and CO2 emissions and system cost for model with 12 regions are shown for respective period and configuration of investment cost

stored energy, transfers, and NTC are negligible.

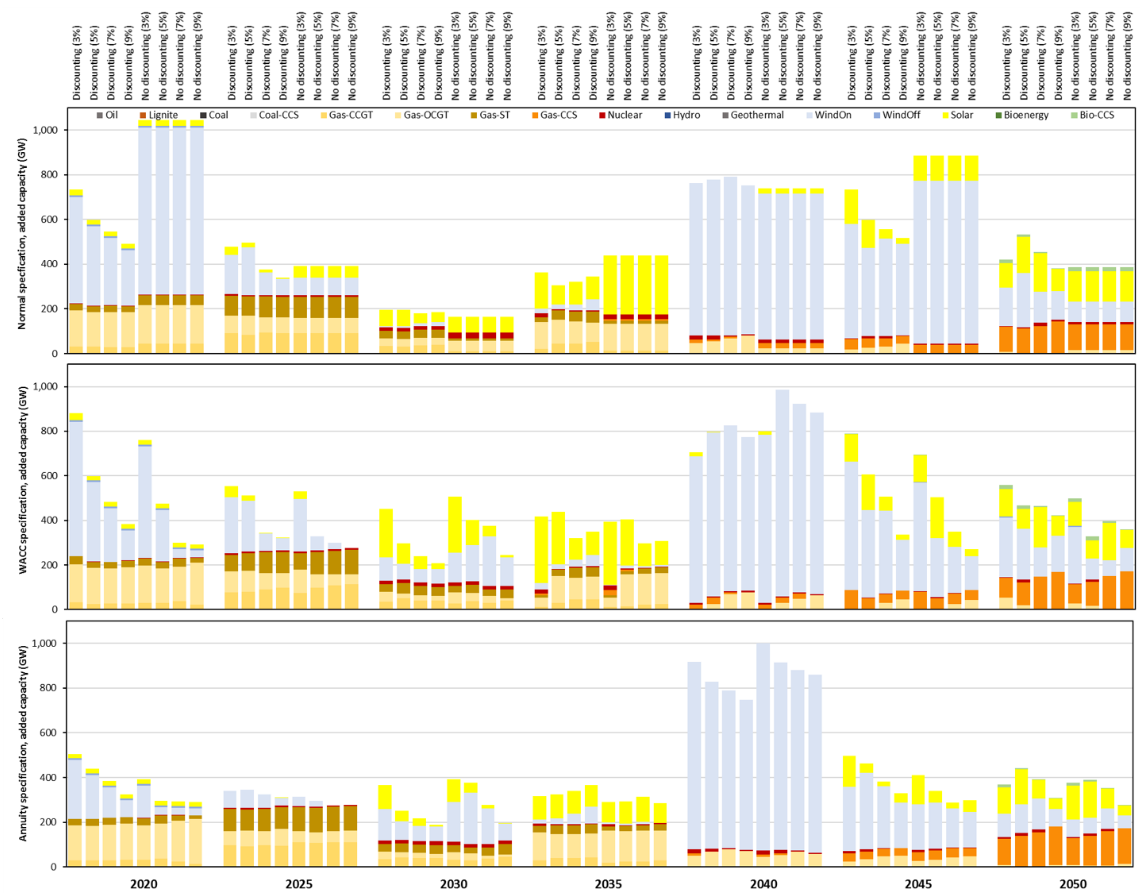
The lowest part of Figure 1 depicts CO2 emissions as gray bars with scale on the left axis and system cost as orange diamonds with scale on the right axis. Starting with system cost, which are calculated based on the annuity specification neglecting discounting, which results in the lowest cost for this specification. Observe that system costs for normal investment specification without discounting are fundamentally higher than in all other specifications. The spire from 2020 continues over time and even increases from 2040 until 2050. Yet, when looking at the corresponding CO2 emissions for this specification we observe also that this specification has the lowest CO2 emissions. The latter has an additional value, especially in the context of the European Green Deal goals to reach climate-neutrality by 2050. The aggressive expansion of wind onshore technology in 2020 and fundamentally higher solar power leads to 66.4% lower emissions (97 Mt vs. 289 Mt) in contrast to the specification with discounting. This results in a reduction of CO2 emissions by 91.3% (to 2015 values). The other specifications, irrespective of discounting, seem to have higher CO2 emissions, both in the medium- and long-term. Conversely, system cost are the lowest for the annuity specification that neglects discounting, which is consistent with the metric of calculating system cost here: we use annuities and neglect discounting. However, note that system costs and investment cost, in particular, often have low explanatory power because the aim of power market models is to reflect firms' investment behavior (to make good predictions and evaluate the impact of certain policies) rather than attempting to perfectly optimize the system.

2.3. Varying Interest and Discount Rates

Note from Subsection 2.1 that interest and discount rates are treated the same. The normal specification calculates end-effects that reflect discounting but does not discount a future stream of costs emanating from a specific investment. In turn, the WACC and annuity specification calculate annual payments following the investment decision. Those annual payments depend on the interest rate, and discounting determines the valuation of those streams of cost over time. The specifications without discounting are thus not fully comparable. The end-effects in the normal specification seem to be highly sensitive when neglecting discounting, whereas the WACC and annuity are independent of discounting because only the stream is subject to discounting. We thus vary the underlying discount and interest rate from seven per cent to three, five, and nine per cent, respectively. Figure 2 shows added capacity from all three specifications with and without discounting.

Our analysis demonstrates that lower interest and discount rates lead to higher investments over time for all three specifications. Looking at normal investment cost specification (upper part of the Figure 2), we observe significant differences with and without discounting. This is especially the case in 2020 and 2045, where we observe significantly higher investments in wind onshore without discounting, or 2025 and 2035 where the same effects are observed for solar. However, we also see that varying

Figure 2: Added capacities for different investment cost specifications



Aggregated values for model with 12 regions are shown for respective period.

interest rate has almost no effect on total investment in this specification without discounting, while with discounting the variation is higher especially in the first periods, the difference dissipates in mid-term and is once again more pronounced during the last two periods. A similar pattern holds for the annuity with slightly higher variation without discounting, but lower variation between similar interest rates with and without discounting. WACC specification is more sensitive to interest rate size variation without discounting. Nevertheless, the differences between the specifications of normal, annuity and WACC without discounting, are far less pronounced.

2.4. Assessment of Investment Cost Specifications

The normal specification is most sensitive to discounting and tends to foster the highest capacity investments. The annuity specification leads to lowest capacity investment because investments are most expensive under such configuration. For example, the annuity of wind onshore investments (interest rate of seven per cent, depreciation time of 25 years) is 8.58 per cent and that of coal power (depreciation time of 40 years), is 7.5 per cent. The corresponding WACC invariably used (no matter of depreciation time) seven per cent, which explains the higher investments here. In turn, the normal approach does not reflect weighted cost or cost for borrowed capital, but rather the absolute sum. Not accounting for risk premia in that sum, that is, using the same investment cost as with the other two approaches, leads to the highest investments. Discounting impedes these effects to some extent, pushing the outcome of the normal approach closer to that of the WACC specification. Reductions in interest and discount rates reduce differences between the specifications, particularly when neglecting discounting. However, the importance of discounting becomes clear. In the long-run, differences become negligible, but when neglecting discounting investors might postpone investments into later stages, which does not reflect reality. Hence, we consider the annuity approach (1) as the most widespread and (2) as the least sensitive one and decide to adhere to that specification (and use discounting) in the following analyses.

3. Foresight

Reducing the foresight (or planning horizon) is one way to increase the precision of the power market's decision by holding computation time reasonable and retaining numerical feasibility. In other words, switching from a perfect foresight version to a myopic one (Priesmann et al., 2019, Babrowski et al., 2014). In this section, we analyze how different planning horizons affect results.

3.1. Foresight Specifications

Intertemporal models optimize the entire planning horizon, that is, they optimize over all t (6). Myopic models optimize only the current time period and use the results from that one as inputs for optimizing the succeeding period. Typically, compared to perfect foresight, myopic models lead to delayed or canceled investments and overall higher system costs (Nerini et al., 2017). Yet, such an approach presents advantages when increasing numerical complexity of the optimization problem with regards to temporal, spatial, or technological resolution, and when investors indeed act myopic (Li, 2017, Heuberger et al., 2018). For instance, as far as the authors are aware there are no intertemporal models enabling the modeling of all the 8760 hours of the year but rather two myopic models that are able to use 8760 hours for calculating of the European power system (Siala et al., 2020). However, Siala et al. (2020) only focus on the comparison of myopic and intertemporal models, and neglect the possibilities inbetween. We now account for a rolling foresight, where a subset of periods is optimized at all times. T is the sample of all periods. Denote by $T' := \{t', ..., t'+x\}$ a subsample of all periods with x describing the foresight of the respective specification (as number of periods). For example, $x = 3$ is a specification that invariably optimizes the current period including three periods ahead. The optimization problem, applying the annuity specification, becomes

$$\min_{\mathbf{Q},...} \quad \textstyle\sum_{t \in T'} \delta\left(t\right) \sum_r \quad \left[\displaystyle\sum_j \sum_{v \leq t} Q_{jr}\left(v\right) C_{jr}\left(v\right) \times \varLambda_{jr}\left(v,t\right) A_{jr}\left(v\right) + ... \right]. \tag{8}$$

3.2. Impact of Rolling Foresight

Figure 3 presents the evolution (from 2015 to 2050) of installed capacity by technology type (first diagram), stored energy by technology type (second diagram), NTC (light blue bars with the scale on the left axis), transfers (blue triangles with scale on the right axis), and CO2 emissions (gray bars with scale on the left axis) as well as system cost (orange diamonds with scale on the right axis) for eight different model specifications. The *myopic* specification only optimizes the respective period, whereas the other myopic specifications look *one* to *six periods* ahead. For example, the myopic specification only optimizes 2015, then 2020, then 2025, ... The three periods specification optimizes 2015 to 2030, 2020 to 2035, ..., and finally 2035 to 2050. The intertemporal specification finally optimizes all periods from 2015 to 2050 simultaneously.

Looking five or six periods ahead, matches (almost perfectly) the intertemporal results. Looking one to four periods ahead leads to considerable differences compared with the intertemporal optimum in the mid-term (2020 to 2035). In the long-term (2040 to 2050), differences are negligible. The only specification that structurally lags behind the intertemporal optimum is the myopic specification. Looking at installed capacities, gas, nuclear, and hydro capacities show almost no difference, whereas wind onshore (and solar) capacity is +49 per cent (+zero per cent) in 2015, +86.1 per cent (+8.7 per cent) in 2035, and +2.4 (10.6 per cent) in 2050. Differences for wind onshore become smaller in size, but are persistent, leading to six per cent higher system cost and 3.6 per cent more CO2 emissions in the myopic version. Those differences impact stored energy, too. Pump storage usage also tends to be higher in the myopic version, while battery storage usage is smaller (in 2050) compared to more forward-looking models. Similarly, NTC and transfers are considerably smaller as well.

Interestingly, although the capacity mixes (and CO2 emissions) already differ considerably from 2020 onwards, system costs remain similar until 2035. From 2040 onwards, capacity mixes and CO2 emissions become similar, but there is an increase in system cost that reinforces over time. Here, it is clear that intertemporal specifications are superior to myopic ones.

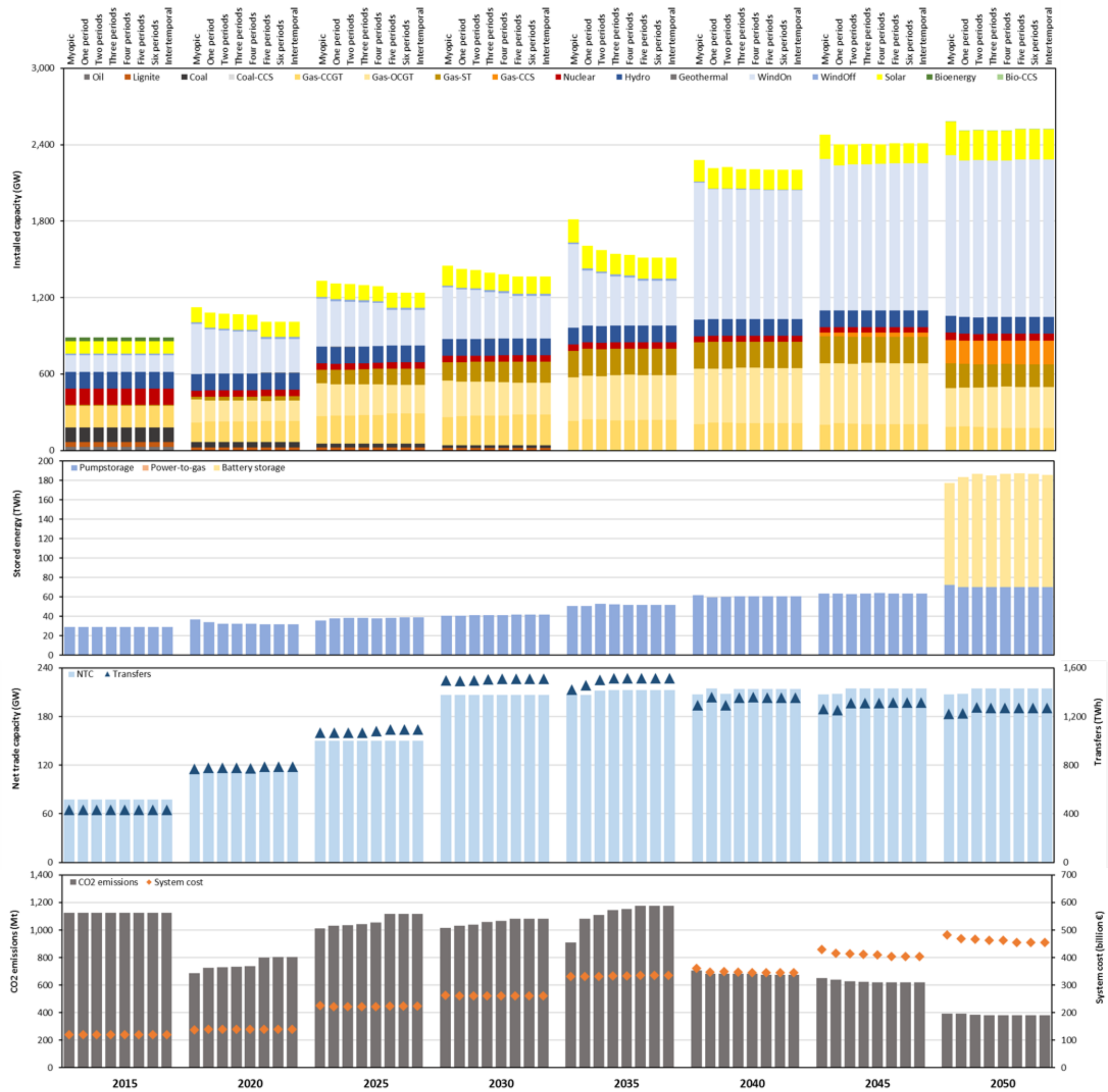
What happens in 2040? From that period onwards, the model can install higher wind turbines (120 meters high) that have better resource profiles for the same cost. Here, it becomes apparent that the far forward-looking models save the best resource spots for this new technology, whereas less forward- looking models do not due to a failure to foresee that technology boost. This also explains why myopic models can install more capacity at similar system cost until 2035—this is accomplished by using better resource sites. However, the lack of foresight becomes crucial over the long run.

3.3. Varying Carbon Prices

Myopic models can perform quite well when there is no fundamental change in market conditions (cost, technological improvements, policy change) (Siala et al., 2020, Heuberger et al., 2018). Having said that technology boosts, as described in the prior subsection, seem to considerably change results. We now analyze another change in market conditions, namely different CO2 price trajectories (see Table 1) that could fundamentally impact results from less forward-looking models. As extremes, we now analyze the myopic and intertemporal specifications in a changing market environment.

The *normal CO2 price* calibration assumed CO2 prices increase from 7.75 EUR/t in 2015, to 15 EUR/t in 2020, and to 132 EUR/t in 2050. The scenario *flat CO2 price* assumes no further price increase from 2025 onwards. *Double* and *triple CO2 price*

Figure 3: Results for different foresight specifications



Aggregated installed capacity for model with 12 regions are shown for respective periods.

assume doubled or tripled CO2 prices, respectively, from 2025 onwards. The flat scenario reflects a demolition of the EU ETS, the double CO2 price scenario is closer to developments reflecting latest changes of the EU ETS (market stability reserve with canceling mechanisms), and the triple scenario seems to be a good reflection of future changes regarding the goals underpinning the European Green Deal (Azarova and Mier, 2020).

Table 1: Carbon price variations

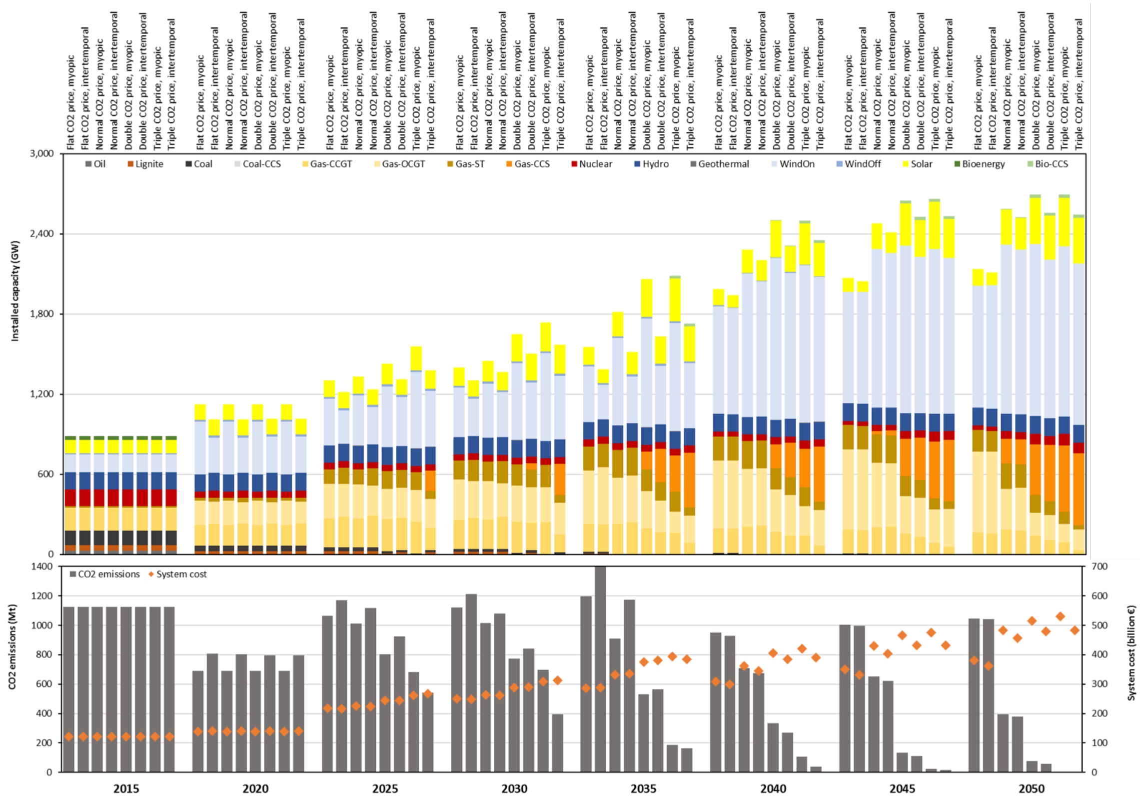
	Normal CO2 price	Double CO2 price	Flat CO2 price	Triple CO2 price
2015	7.75	7.75	7.75	7.75
2020	15	15	15	15
2025	22	44	15	66
2030	27	54	15	81
2035	56	112	15	168
2040	68	136	15	204
2045	102	204	15	306
2050	132	264	15	396

Figure 4 shows the evolution of installed capacity (upper diagram), and CO2 emissions (gray bars with scale on the left axis) as well as system cost (orange diamonds with scale on the right axis) for the myopic and the intertemporal specification for the four different market environments. Start considering installed capacities. Already in 2020, the myopic version fundamentally differs from the intertemporal one. These differences reduce over time in the flat scenario (+11.4% in 2020, +12.1% in 2035, and +1.1% in 2050). Differences are considerably higher for double and triple CO2 prices (+5.3% or +5.9%, respectively, in 2050).

We now focus on CO2 emissions and cost, observe that system costs are similar between myopic and intertemporal specifications until 2035. From 2040 onwards, this fundamentally changes due to the aforementioned technology boost (higher wind turbines are available from 2040 onwards). The use of better sites from 2020 to 2035 in the myopic versions keeps the cost low. The intertemporal models save the spots to use them from 2040 onwards, leading to structural cost differences from 2040 onwards. Moreover, differences in costs are smallest in the flat price scenario and highest in the one with tripled CO2 prices. Interestingly, the use of better wind spots in the myopic specifications in 2020 already does not change system costs, but leads to fundamentally lower CO2 emissions (-14.3%). This pattern vanishes in from 2040 onwards. The intertemporal models then tend to deliver lower emissions, whereas the differences are small.

Moreover, it is not only the aggregate capacity that changes, but also the decomposition. Intertemporal models rely less on classic gas power technologies (gas-CCGT, gas-OCGT, gas-ST) but more on gas-CCS, resulting in lower wind onshore

Figure 4: Installed capacities, CO2 emissions, and system costs for different CO2 price trajectories



and solar capacity. In the long-term this translates into lower CO2 emissions and lower system costs which is especially apparent for double and triple CO2 price scenarios.

3.4. *Assessment of Foresight Specifications*

Our analysis illustrates the pitfalls of relying entirely on myopic models that neglect any foresight (of firms). Such models rely on future costs and prices being subject to uncertainties that increase with the length of the considered time-frame (Keppo and Strubegger, 2010). Our findings contradict the literature suggesting that myopic models perform better when reflecting the decision-making framework relevant for a more realistic representation of energy system and actual investment behavior as argued by Martinsen et al. (2006, 2007), Poncelet et al. (2016), Hedenus et al. (2006). The claims that myopic models are superior when such aspects as reactions to sudden changes (such as a sudden increase of oil and gas prices), flexibility of technical scenarios (such as the use of long lived and high investment technologies by varying energy policy) need to be examined. However, fully myopic specifications, not having any foresight seem to create costly path dependencies and lead to fundamentally different results when market conditions change dynamically. Therefore, we suggest considering a foresight specification looking at least a few periods ahead when the target is to reflect investment behaviour. At the same time, opting for a myopic horizon instead of a perfect foresight allows expanding analyses with higher spatial, temporal, or technological resolutions with manageable computational times (Poncelet et al., 2016, Babrowski et al., 2014). However, it can also result in an oversized and underutilised power system when a disruptive technological change is introduced in the model Heuberger et al. (2018). We, therefore, suggest running sensitivities with different foresight horizons (eventually at reduced temporal and spatial resolutions), to gain an understanding of the bias in results related to time horizon configuration of the model.

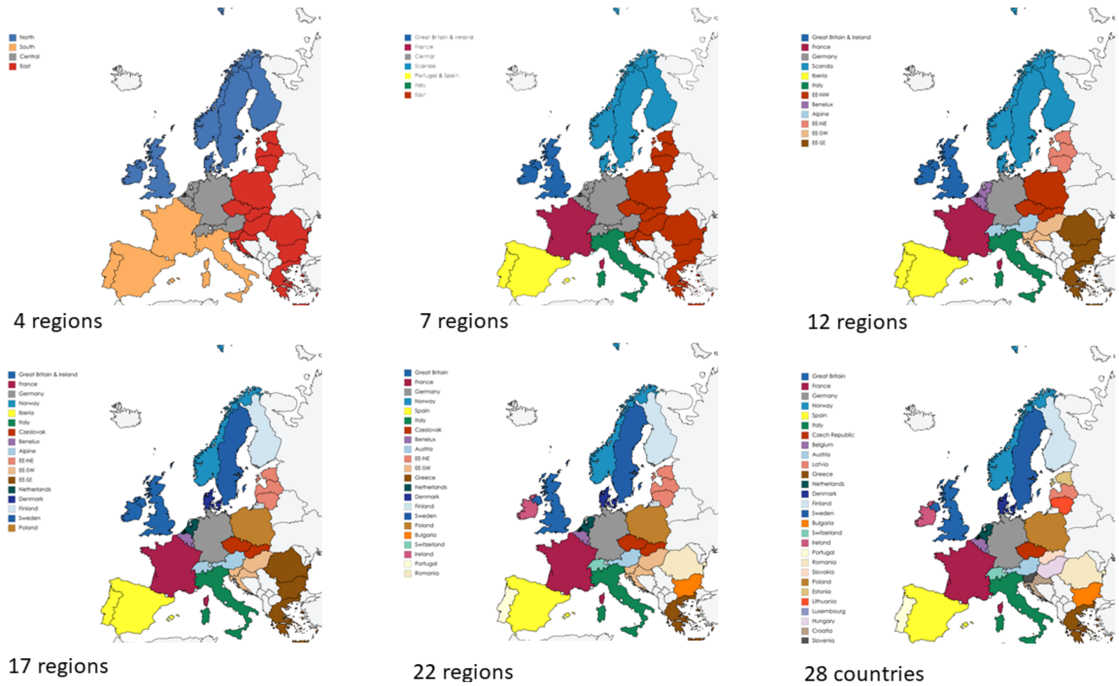
4. Spatial Resolution

The selected spatial resolution is a crucial determinant for power market models (Frew and Jacobson, 2016, Siala et al., 2019, 2020). Rising complexity due to an increasing share of renewable resources and the decentralization of the electricity provision, require a high spatial resolution to correctly depict resource potentials and resulting electricity flows correctly. At the same time, the degree of spatial resolution can have a significant effect on computation time and the numerical feasibility of the models. We thus analyze various degrees of spatial resolutions. Figure 5 shows the country composition of six different spatial resolutions that are analyzed in this section.

4.1. Aggregation Routine

We consider 28 countries presenting the European power market. Those countries can be aggregated differently to regions to reflect the reality of markets (e.g., Scandinavian countries are organized in the Nordic markets) or to reduce numerical complexity. Reducing spatial resolution is always a matter of calibration (Siala et al., 2020), because (sub-)country-level values need to be aggregated. This is not problematic when analyzing stock values such as installed capacities but becomes critical as soon it is about different cost, efficiencies, or other region-specific technological parameters such as resource potential and full-load hours of intermittent renewables, pump hydro availability, and the depiction of transmission between countries. Transmission is particularly important because grouped countries assume infinite transmission capabilities (at no cost or losses, respectively) within grouped countries.

Figure 5: Spatial resolutions considered



Such an aggregation enables focus on transmission bottlenecks between regions, and neglect those between countries grouped to one region. Another important point is the different resource potential for intermittent renewables.

The used calibration has unified technology cost (capacity, fixed and variable operation and maintenance) across European countries. Commodity prices for oil, natural gas, coal, and uranium differ only marginally due to upfront cost chains. Transmission (line) cost and losses are subject to specific country-pairs, depending on the overall length of a line and the transmission technology used (AC lines or DC cables, respectively). We calibrate the model with country-level values and use different aggregation routines to obtain regional counterparts. We now describe the most important aggregation routines.

Transmission.. Denote by ρ a stock parameter (e.g., installed capacity) and by ϱ a specific or relative parameter (e.g., capacity cost, line losses). $r - r'$ is a neighbouring region pair. $cty - cty'$ is the corresponding country pair, where $cty \subseteq r$ and $cty' \subseteq r'$. We can then describe the mapping between neighbouring regions and containing countries so that stock, specific, and relative parameters can be aggregated to regional values, e.g.,

$$\rho_{r-r'} = \sum_{cty \subseteq r} \left(\sum_{cty' \subseteq r'} \rho_{cty-cty'} \right), \tag{9}$$

$$\varrho_{r-r'} = \max_{cty \subseteq r} \left(\max_{cty' \subseteq r'} \varrho_{cty-cty'} \right). \tag{10}$$

We chose the maximum for specific or relative parameters because of neglecting bottlenecks between countries. Averaging, for example, would simplify transfers to an unreasonable amount.

Load timeseries.. Denote by $D_r(t), D_{cty}(t)$ regional or country-level demand, respectively, and by $d_r(h), d_{cty}(h) \in (0,1]$ the 2015 load profiles (1 reflects the peak hour) that are used to represent future periods as well. Regional demand follows from simple aggregation, i.e., $D_r(t) = \sum_{cty \subseteq r} D_{cty}(t)$, and the corresponding load profile follows from

$$d_r(h) = \frac{\sum_{cty \subseteq r} d_{cty}(h) D_{cty}(2015)}{\sum_{cty \subseteq r} D_{cty}(2015)}. \tag{11}$$

Intermittent renewables timeseries.. We consider high, medium, and low resource classes for solar and wind power (onshore and offshore). By assumption, the medium class is always three times the size of high and low classes (Siala et al., 2019). $Q^{max}_{j,class}$ is the resource potential of an intermittent technology by class, i.e., $Q^{max}_{j,class,r} = \sum_{cty \subseteq r} Q^{max}_{j,class,cty}$ describes the aggregation of country-level to regional potentials. Denote by $\Psi_{j,class}$ full-load hours of an intermittent technology (wind, solar, hydro) and by $\psi_{j,class}(h) \in (0,1]$ the corresponding profile (one reflects maximum and zero minimum in-feed). We obtain regional profiles by weighting with the corresponding resource potential, i.e.,

$$\psi_{j,class,r}(h) = \frac{\sum_{cty \subseteq r} \psi_{j,class,cty}(h) Q^{max}_{j,class,cty}}{\sum_{cty \subseteq r} Q^{max}_{j,class,cty}}. \tag{12}$$

4.2. Impact of Spatial Resolution

Figure 6 shows the evolution (2015 to 2050) of electricity generation by technology in the first panel, stored energy by technology in the second panel, NTC (blue bars with scale on left axis). While also presenting transfers (blue triangles with scale on the right axis) in the third panel, and CO2 emissions (gray bars with scale on left axis) as well as system cost (orange diamonds with scale on right axis) in the fourth panel for six different spatial resolutions.

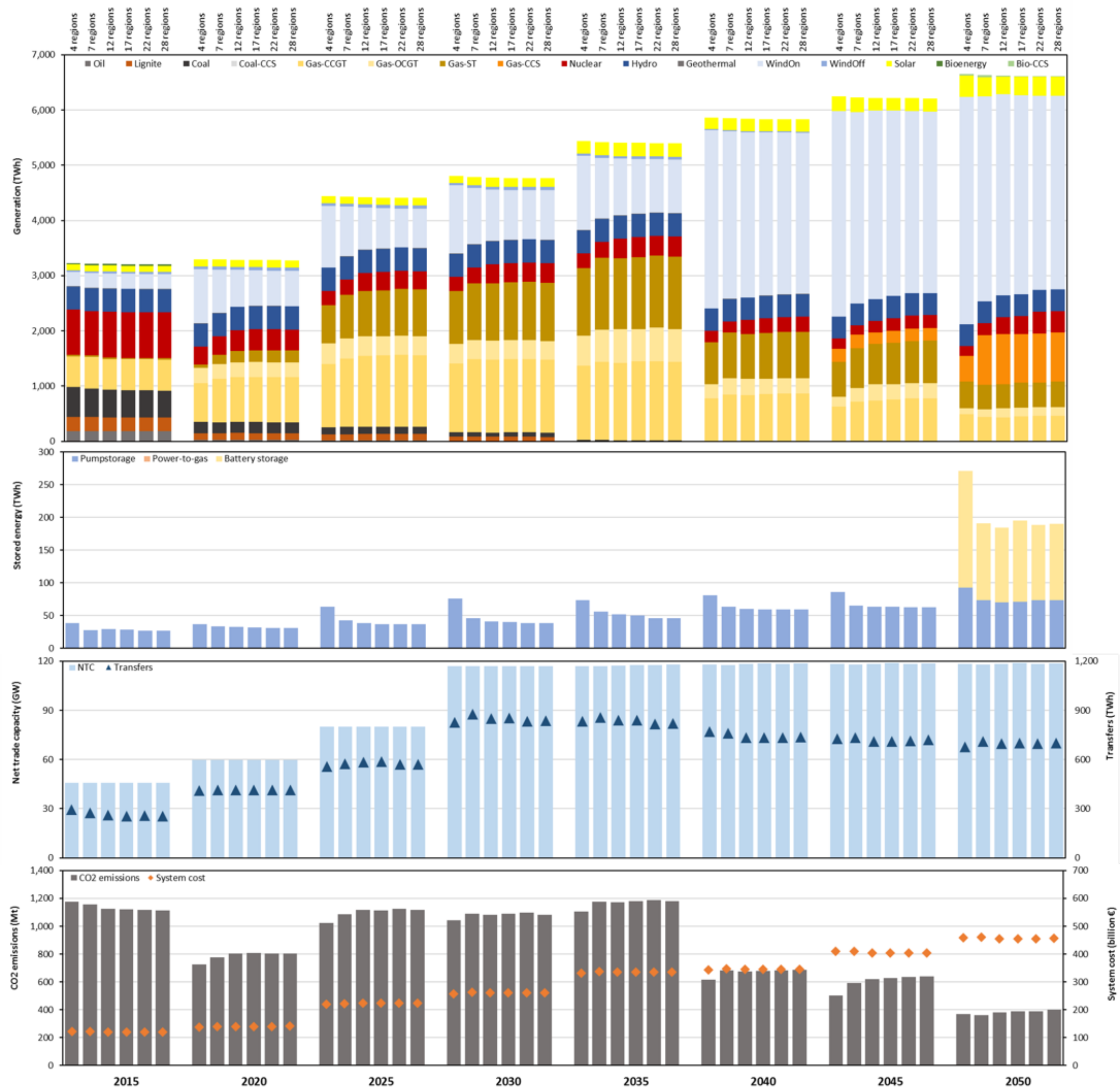
The first panel demonstrates how lower spatial resolutions foster the exploitation of wind power and solar power, albeit to a lesser extent. More wind power then substitutes gas and nuclear power. In particular, gas-CCS is fundamentally lower in configuration with four regions when compared to all the others (-47.5 per cent or -412 TWh, wind onshore generation is 616 TWh or 17.6 per cent higher). This also enhances the use of pump and battery storage (see second panel), especially in the long-term. A similar trend is observed for a seven regions configuration. However, already an aggregation to twelve and 17 regions shows results that are relatively comparable with the full country resolution (28 countries).

The expansion of wind power in more aggregated configurations might be fostered either by (1) no transmission boundaries within a region, or (2) better possibilities to transmit to other regions as well. The first point can only by addressed by changing resource potential and profiles in response to the aggregation routine. In turn, the second is a matter of calibration. We can test our transmission calibration by controlling NTC and transfers in the third panel of Figure 6. Observe that NTC values are (almost) constant, since we only aggregate NTC and transfers values between the regions of the lowest spatial resolution. NTC and transfers within North, South, Central, and East are thus neglected in all specifications. Furthermore, the transfer volume is quite stable across the different specifications. It seems that transfers are a bit higher in the mid-term for lower resolutions, with this effect leveling out over the long-term.

Observe that total electricity generation is almost the same between resolutions. Differences can be traced back to lower (absolute) line losses for lower resolutions and, eventually, different storage patterns. In other words, higher storage use comes at higher storage losses. However, storage patterns do not have much impact on our results. In turn, results are driven by the transmission boundaries that are partly neglected when aggregating 28 countries to, for example, four regions—North, South, Central, and East.

Finally, observe that that CO2 emissions and system cost for twelve, 17, and 22 regions differ only slightly from the configuration with 28 countries. Until 2045, CO2 emissions are fundamentally lower for the lowest resolutions (four and seven regions) due to high wind expansion. This effect is still observable in 2050, but substitution for gas-CCS and nuclear almost entirely compensates that wind boost in the higher resolutions. Interestingly, there is almost no difference in system cost until 2040.

Figure 6: Results for different spatial resolutions



From 2045 onwards, lower resolutions experience higher cost because the effect from averaging resource potential now has a strong impact on our results.

4.3. Varying Upper Bounds for Transmission Capacity Expansion

We now choose the two extreme configurations, four regions and 28 countries, and control for the impact of spatial resolution by varying the upper bounds for NTC expansion. *No transmission* sets NTC values to zero and thus prevents all transmission. All remaining differences can then be attributed to aggregation flaws when aggregating timeseries and potential of intermittent renewables. Besides the *simple NTC* scenario, we also model *double* and *triple NTC* upper bounds from 2035 onwards.[4] Figures 7 shows results.

Start with the upper panel that presents the evolution of electricity generation. Observe that higher NTC bounds foster wind onshore expansion and reduce gas and gas-CCS production. This effect is stronger for the configuration with 28 countries, underlying that the averaging of full-load hours and resource potential of intermittent technologies underestimates renewables diffusion.

Now turn to the lower panel that presents NTC (blue bars with scale on the left) and transfers (blue triangles with scale on the right). Until 2030, NTC and transfers are the same for all variations and both configurations, except those neglecting transmission. Allowing higher NTC expansion leads to fundamental differences in the double and triple NTC variation. Interestingly, for the simple NTC variation resulting NTC and transfers are very similar.

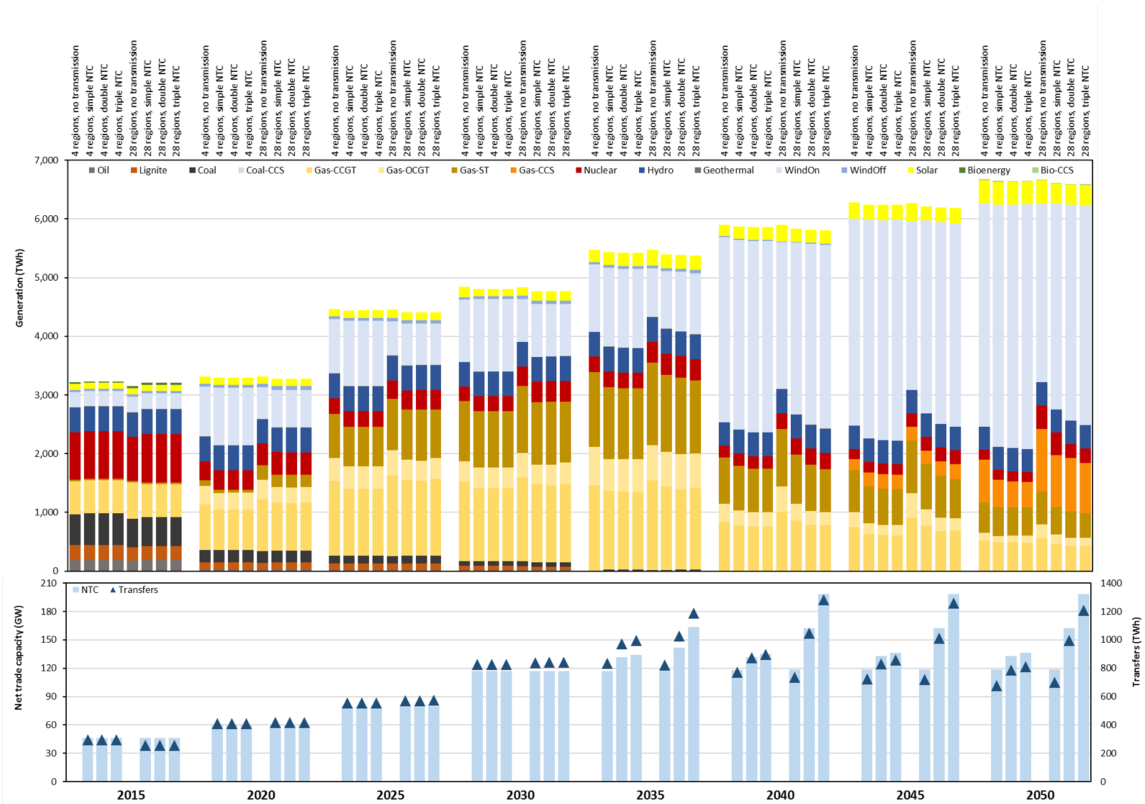
4.4. Assessment of Spatial Resolution

On the whole, our aggregation routine regarding transmission modeling demonstrates fine performance and comparable results. However, wind onshore expansion is still fostered by grouping countries given that bottlenecks are neglected. A possible option for resolving this issue would be to scale resource potential or profiles. Additionally, the usage of more resource classes (by country) could be useful for avoiding the loss of too much information when aggregating. Note that fostering wind power also affects emissions and corresponding system costs.

Moreover, changing market environments such as higher NTC boundaries that allow for more transmission capacity expansions, increase differences between spatial resolutions because the effect of neglecting bottlenecks and averaging resource potential becomes more severe

[4]NTC expansion until 2030 is fixed by the ten-year-development plan of transmission system operators.

Figure 7: Generation, NTC, and transfers for different transmission expansion scenarios



5. Temporal Resolution

Ideally, a power market model should run serially for all hours of a year to capture the time-linked behavior of storage and to avoid omission of any solution-constraining hours Frew and Jacobson (2016). However, such a solution is computationally unreasonable in most cases Ridha et al. (2020). Selecting the temporal resolution of a power market model hence becomes a philosophic question (a matter of taste and preferences), as underlined by the variance in approaches. In general, there are two types of philosophies here: (1) One tries to model the resolution of year (which consists of 8760 hours) in as much detail as possible. When this is not possible, the number of hours presented is reduced until the model is (numerically feasible) solvable again. The weighting of the reduced time series is often simple, so that each hour represents the same number of hours. (2) One tries to reduce temporal complexity by choosing a subset of hours and weight them, both done either via algorithms or based on some experience. We apply both approaches and evaluate how much simplification is possible. For the algorithm results, as well as the simple selection outcome, we apply scaling of load, wind, and solar timeseries to meet annual demand by region, and technology- and class-specific full-load hours of intermittent technologies (hydro, solar, wind); otherwise, it would not be possible to compare results of timescales.

5.1. Selection and Weighting of Hours

Selection and weighting algorithm.. EUREGEN inhibits an algorithm to determine representative hours and its weighting to reflect the extremes of wind and solar power, as well as load by region (Blanford and Weissbart, 2019). We thus calculate a representative hourly timeseries for wind, solar, and load for each region; where one represents the maximum availability of wind and solar in-feed or maximum demand, respectively. We then create one- and three-dimensional vertices of these extremes and search for the hours best reflecting corners of those vertices. A bubble tolerance around those closest hours creates a set of eligible hours, which can be chosen to reflect the extremes of wind and solar. The chosen bubble tolerance also determines the allowed error in the final selection. Thus, the number of eligible hours is high for big bubbles but becomes small for high tolerances because just a few hours are appropriate for depicting the entire set. Bubble tolerances of ten per cent lead to 34 segments, tolerances of five per cent to 60, two per cent to 92, and a tolerance of 0.33 per cent to 117 segments (and the smallest final deviations for all checked bubble tolerances). The weighting then reduces the error to the hourly timeseries.

Simple selection and weighting.. The simple selection and weighting decide to take the first four days of each season. We do so to retain the computation time of one run within days (and not weeks). The four days results in 384 hours in total. Each hour is weighted accordingly by 22.8125. We then reduce the selected days to the first two of each season (192 hours), the first day of each season (92 hours), and finally just select the first summer and first winter day (48 hours).

Mapping between hours and representative hours.. Denote by $x\left(h,h'\right)$ the mapping of hours h to representative hours h' and by $w\left(h'\right)$ the weight of the representative hour. Using superscript red to indicate reduced timeseries, we have $d_r^{red}\left(h'\right) = \sum_{x(h,h')} d_r\left(h\right)$ and $\psi_{j,class,r}^{red}\left(h'\right) = \sum_{x(h,h')} \psi_{j,class,r}\left(h\right)$. Annual demand ($D_r^{peak}$ denotes peak demand by region) or full-load hours, respectively, calculate according to

$$
\begin{aligned}
D_r(t) &= \sum_h d_r(h)\, D_r^{peak}, && (13)\\
D_r^{red}(t) &= \sum_h d_r^{red}(h')\, w(h')\, D_r^{peak}, && (14)\\
\Psi_{j,class,r} &= \sum_h \psi_{j,class,r}(h), && (15)\\
\Psi_{j,class,r}^{red} &= \sum_h \psi_{j,class,r}^{red}(h')\, w(h'). && (16)
\end{aligned}
$$

Scaling of timeseries.. We scale load and intermittent renewables timeseries so that annual demand or full-load hours match the values from a 8760-hours-timeseries. This scaling is necessary to guarantee comparability of different temporal resolutions. The scaling factor calculates according to $D_r(t)/D_r^{red}(t)$—which is constant for all t—or $\Psi_{j,class,r}/\Psi_{j,class,r}^{red}$, respectively. Those scaling factors are applied on the reduced timeseries to obtain the scaled ones (superscript *scale* indicates scaled timeseries),e.g.,

$$
\begin{aligned}
d_r^{scale}(h') &= d_r^{red}(h')\,\frac{D_r(t)}{D_r^{red}(t)}, && (17)\\
\psi_{j,class}^{scale}(h') &= \psi_{j,class}^{red}(h')\,\frac{\Psi_{j,class,r}}{\Psi_{j,class,r}^{red}}. && (18)
\end{aligned}
$$

5.2. Impact of Temporal Resolution and Choice of Representative Hours

We expect temporal resolution to have the highest impact on storage behavior. We therefore not only evaluate electricity generation, but also stored energy. Figure 8 demonstrates the outcomes of our analysis. We begin by looking at the upmost panel, representing the outcomes of various temporal resolution configuration on the generation mix. The first four bars from the left correspond to selection algorithm, the last four to simple selection strategy. Both strategies going from left to right increases the number of representative hours. In the short-term, both simple selection and algorithm configurations seem to provide relatively similar results, with an enhanced onshore wind (for instance 17.2 per cent more wind is already observed in 2020 for 48 hours vs 34 segments models) and solar (starting from 2035 where 13.7 per cent more solar is observed for the same model pair) in the case of simple selection. Additionally, looking at the results of simple selection we observe that lower temporal resolution leads to overestimation of wind generation, as already observed in 2020. Hence, we find similar results as suggested by (Haydt et al., 2011) or (Pina et al., 2011) indicating that lower temporal resolutions lead to overestimation of RES generation. The indicated differences in generation mix become more pronounced in the mid-term and are especially evident in the long-term where significantly more wind and solar is observed for sample selection configurations in 2045 and 2050. However, if we compare the best algorithm selection (117 segments) and 384 hours, differences between the two approaches' generation mix are negligible, while computational cost savings are not.

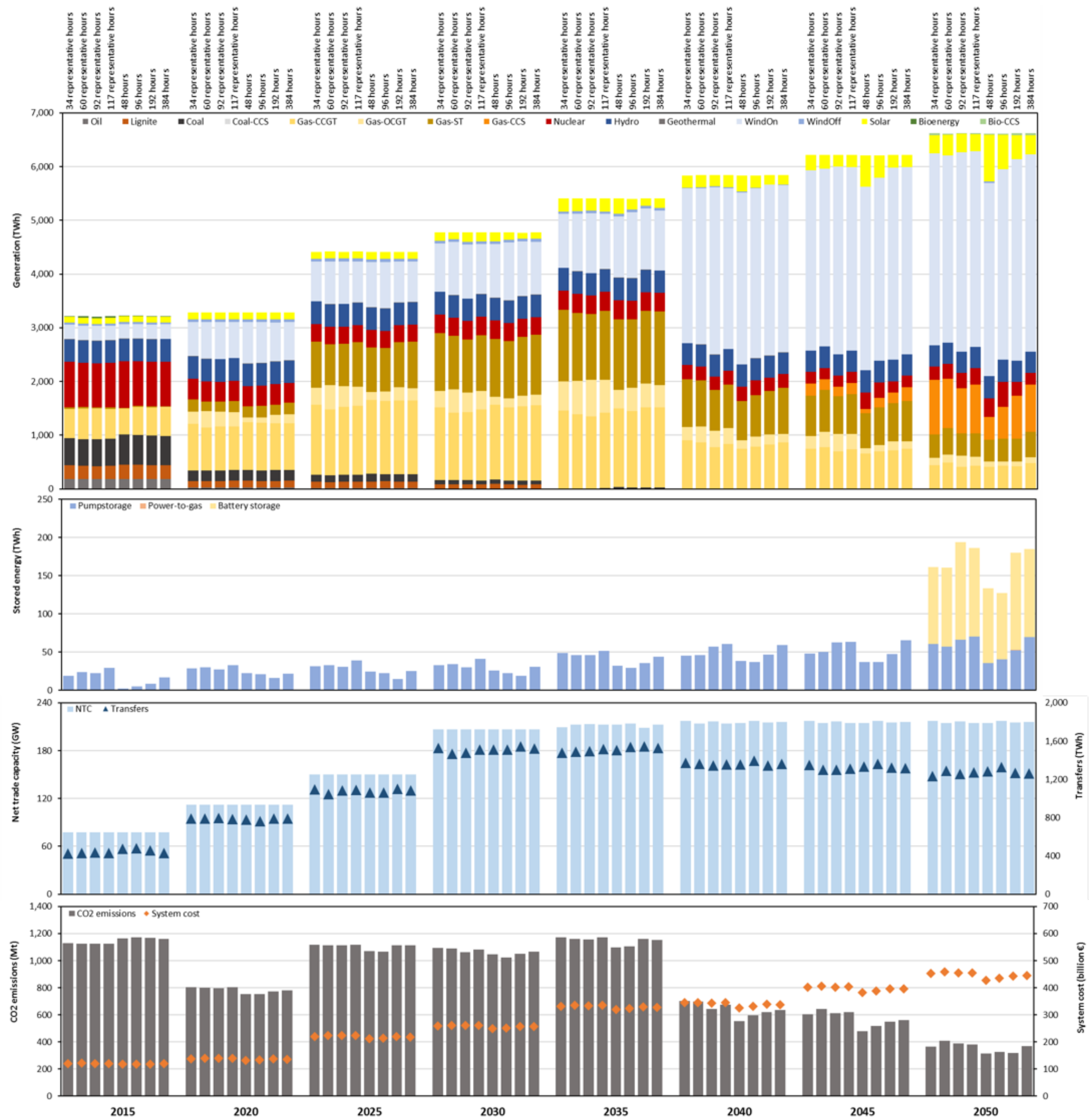
Looking at the storage outcomes, as expected we observe more variation across the temporal configurations. Specifically, we find that simple selection configurations require far less storage in the short term, for instance, -42 per cent for the highest resolution pair of 117 segments versus 384 hours model is required in 2020. Interestingly, these differences dissipate in the long-term, with the difference between the same pair of models being only three per cent towards 2050. Overall, we can conclude that in terms of storage, similar dynamics are observed within the two approaches: increased temporal resolution for both simple selection and algorithm increases storage demand. Looking at the outcomes of temporal resolution on the system cost and CO2 emissions, our analysis suggests that simple selection results in lower system cost and CO2 emission levels. These effects can potentially be attributed to specific weather conditions (high wind and sunshine hours) of the selected days. These could be evaluated by selecting random days within a season as opposed to the first two days of the season.

5.3. Varying Storage Cost

In this section we demonstrate sensitivity of selection algorithm with respect to storage. We do so by varying the level of capital cost of storage charge and discharge: simple cost - is the usual cost used in all the previous analyses. We also include one third, half, double, and even triple the cost as well as specification without storage possibility for comparison. We use the selection algorithm with the 117 segments.

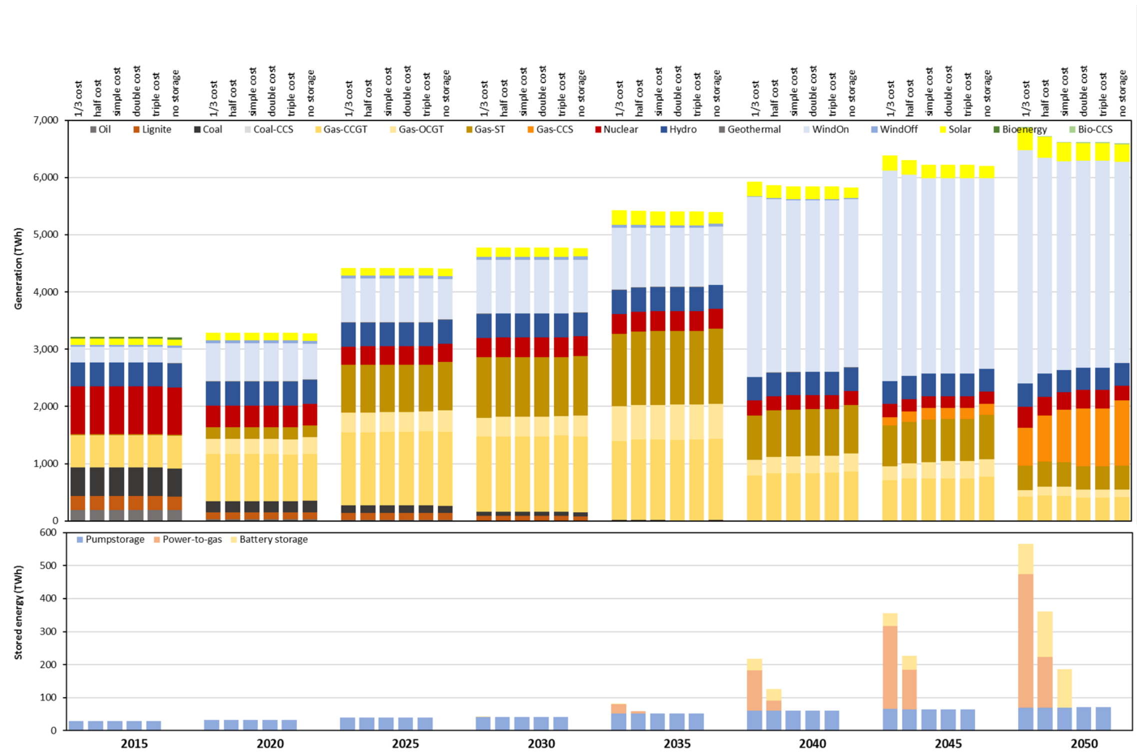
The results presented in Figure 9 show that the price of storage (both reducing and increasing) has no significant impact on generation up until 2040. From 2040 (after introducing a new type of wind turbine and an increase in CO2 prices) we observe that lower storage cost enhances wind and solar production. For instance, the difference in 2050 between 1/3 cost and triple the cost translates into 13 per cent higher wind and 24 per cent higher solar generation. Additionally, higher storage costs promote a higher share of Gas-CCS technology with 54 per cent difference between the two extreme configurations. Observe that the configuration without

Figure 8: Results for different temporal resolutions



storage at all seems to be quite similar to usual configuration in terms of the generation mix outcomes. Hence, while in the short term the impact of storage price on the generation capacity is negligible, long-term changes are more pronounced. Yet, in general considering overall comparability of base configuration and the one without storage, neglecting storage offers an additional possibility to improve computational

Figure 9: Generation and storage mix for varying storage cost



times without significant impact on generation mix.

5.4. Assessment of Temporal Resolution

Looking at the results of simple selection of hours and algorithm application, we conclude that both approaches provide quite comparable results. Particularly the versions with the highest resolution (384 hours vs 177 segments), demonstrate only negligible differences in the outcomes, while the simple selection configuration facilitates considerable savings in computational times. However, one should consider that simple selection and lower temporal resolution for both approaches might enhance wind and solar generation. This could potentially overestimate their share in the generation mix, thereby reducing demand for storage and underestimating system costs and CO2 emissions.

6. Conclusion

Following the approach of Ridha et al. (2020), we demonstrate the impact of key dimensions of the power market model on its outcomes. Among the analyzed features of the power market models including investment configuration, foresight, spatial aggregation, and temporal resolution, we find that investment configuration has the highest impact on the model's outcomes in terms of installed capacity, storage, NTC and system costs and CO2 emissions.

Each set of investment behavior has a set of advantages and pitfalls. The normal specification as applied in EUREGEN might be the closest to the economic perspective of a central planner that optimizes the power system. The annuity specification in turn has high-risk premiums, depending on the investment's depreciation (or amortization) time. The WACC specification neglects the depreciation time of an investment by applying a constant WACC for each unit of bounded capital. One cannot evaluate which approach better fits the respective market environment without knowing it. The European power market provides arguments for all three approaches. For example, a state-owned monopolist in France owns most of the system. Such a constellation favors the normal specification. However, other countries in Europe experience new firms that challenge incumbent ones. For new firms, the annuity approach (higher share of loan capital) might fit best, whereas incumbent ones could stay with the WACC specification because their owned capital is higher. The "fit" might also change over time when looking at the energy transition and large-scale exploitation of wind power. Hence, investment configuration, while significantly affecting the outcomes, should be adjusted to specific contexts and research perspectives. At the same time, one should be bear in mind that the normal specification is the most sensitive to discounting and tends to foster highest capacity investments. While the annuity specification leads to lowest capacity investment since investments are most expensive under such configuration.

Considering the foresight, the two conflicting approaches, myopic and perfect foresight as well as an in-between approach of rolling myopic horizon (allowing a certain limited foresight) are shown, and although both approaches have strong adepts, supporting one or the other, we suggest that fully myopic models must be used with caution in case a significant change in environment (technological boost, change in legislation) is expected in future. Additionally, we find that intertemporal models perform better regarding cost (-6.04 per cent in contrast to full myopic) but decarbonizes only slightly better. The cost impact increases under scenarios assuming more fundamental market changes (e.g., double, or triple CO2 prices)

For the spatial resolution, we conclude that twelve or 17 regions model produces quite similar results to the 28 countries model. Yet, a higher level of aggregation leads to enhanced onshore wind technology expansion, which should be considered if a specific research question or policy analysis requires higher aggregation levels. Finally, testing the impact of temporal resolution, we find that even simple selection of hours is relatively comparable with selection algorithm results. Hence if computational times are to be improved for a specific analysis, we recommend considering the simplification of hours selection. Having said that, both higher aggregation and lower temporal resolution overestimate the share of wind and solar power in the generation mix, which should be accounted for in resulting policy recommendations.

Acknowledgments. We gratefully acknowledge the financial support by the German Federal Ministry for Economic Affairs and Energy (grant number 020E100374092).

Declaration of interests. The authors declare that they have no known competing conflict of interest or financial interests that could have appeared to influence the work reported in this paper.

CRediT author statement. **Mathias Mier:** Conceptualization, Methodology, Investigation, Visualization, Writing - Original Draft. **Valeriya Azarova:** Conceptualization, Investigation, Writing - Original Draft.

References

Amerighi, O., U. Ciorba, G. Giannakidis, and M. Tommasino (2010). Initial Categorization of Tools Used for the Long Term Energy System Analysis-First Results from ATEsT Project.

Azarova, V. and M. Mier (2020). Market Stability Reserve under exogenous shock: The case of COVID-19 pandemic. *Applied Energy*, 116351.

Babrowski, S., T. Heffels, P. Jochem, and W. Fichtner (2014). Reducing computing time of energy system models by a myopic approach. *Energy systems 5*(1), 65–83.

Bachner, G., J. Mayer, and K. W. Steininger (2019). Costs or benefits? assessing the economy-wide effects of the electricity sector's low carbon transition–the role of capital costs, divergent risk perceptions and premiums. *Energy Strategy Reviews 26*, 100373.

Bale, C. S., L. Varga, and T. J. Foxon (2015). Energy and complexity: New ways forward. *Applied Energy 138*, 150–159.

Blanford, G. J. and C. Weissbart (2019). A framework for modeling the dynamics of power markets – the eu-regen model. *ifo Institute, Munich*.

Cao, K.-K., F. Cebulla, J. J. G. Vilchez, B. Mousavi, and S. Prehofer (2016). Raising awareness in model-based energy scenario studies—a transparency checklist. *Energy, Sustainability and Society 6*(1), 28.

Connolly, D., H. Lund, B. V. Mathiesen, and M. Leahy (2010). A review of computer tools for analysing the integration of renewable energy into various energy systems. *Applied energy 87*(4), 1059–1082.

DeCarolis, J., H. Daly, P. Dodds, I. Keppo, F. Li, W. McDowall, S. Pye, N. Strachan, E. Trutnevyte, W. Usher, et al. (2017). Formalizing best practice for energy system optimization modelling. *Applied energy 194*, 184–198.

Frew, B. A. and M. Z. Jacobson (2016). Temporal and spatial tradeoffs in power system modeling with assumptions about storage: An application of the power model. *Energy 117*, 198–213.

Gerbaulet, C. and C. Lorenz (2017). dynelmod: A dynamic investment and dispatch model for the future european electricity market. Technical report, DIW Data Documentation.

Haydt, G., V. Leal, A. Pina, and C. A. Silva (2011). The relevance of the energy resource dynamics in the mid/long-term energy planning models. *Renewable energy 36*(11), 3068–3074.

Hedenus, F., C. Azar, and K. Lindgren (2006). Induced technological change in a limited foresight optimization model. *The Energy Journal* (Special Issue# 1).

Heuberger, C. F., I. Staffell, N. Shah, and N. Mac Dowell (2018). Impact of myopic decision-making and disruptive events in power systems planning. *Nature Energy 3*(8), 634–640.

Keppo, I. and M. Strubegger (2010). Short term decisions for long term problems-The effect of foresight on model based energy systems analysis. *Energy 35*(5), 2033–2042.

Li, F. G. (2017). Actors behaving badly: Exploring the modelling of non-optimal behaviour in energy transitions. *Energy Strategy Reviews 15*, 57–71.

Martinsen, D., V. Krey, and P. Markewitz (2007). Implications of high energy prices for energy system and emissions—the response from an energy model for germany. *Energy policy 35*(9), 4504–4515.

Martinsen, D., V. Krey, P. Markewitz, and S. Vogele (2006). A time step energy process model for Germany-model structure and results. *Energy Studies Review 14*(1), 35.

Merkel, E., D. Fehrenbach, R. McKenna, and W. Fichtner (2014). Modelling decentralised heat supply: An application and methodological extension in TIMES. *Energy 73*, 592–605.

Morrison, R. (2018). Energy system modeling: Public transparency, scientific reproducibility, and open development. *Energy Strategy Reviews 20*, 49 – 63.

Müller, B., F. Gardumi, and L. Hülk (2018). Comprehensive representation of models for energy system analyses: Insights from the Energy Modelling Platform for Europe (EMP-E) 2017. *Energy strategy reviews 21*, 82–87.

Nerini, F. F., I. Keppo, and N. Strachan (2017). Myopic decision making in energy system decarbonisation pathways. A UK case study. *Energy strategy reviews 17*, 19–26.

Pina, A., C. Silva, and P. Ferrão (2011). Modeling hourly electricity dynamics for policy making in long-term scenarios. *Energy Policy 39*(9), 4692–4702.

Poncelet, K., E. Delarue, D. Six, and W. D'haeseleer (2016). Myopic optimization models for simulation of investment decisions in the electric power sector. In *2016 13th International Conference on the European Energy Market (EEM)*, pp. 1–9. IEEE.

Priesmann, J., L. Nolting, and A. Praktiknjo (2019). Are complex energy system models more accurate? An intra-model comparison of power system optimization models. *Applied Energy 255*, 113783.

Prina, M. G., G. Manzolini, D. Moser, B. Nastasi, and W. Sparber (2020). Classification and challenges of bottom-up energy system models - a review. *Renewable and Sustainable Energy Reviews 129*, 109917.

Ridha, E., L. Nolting, and A. Praktiknjo (2020). Complexity profiles: A large-scale review of energy system models in terms of complexity. *Energy Strategy Reviews 30*, 100515.

Siala, K., C. de la Rúa, Y. Lechón, and T. Hamacher (2019). Towards a sustainable european energy system: Linking optimization models with multi-regional input-output analysis. *Energy Strategy Reviews 26*, 100391.

Siala, K., M. Mier, L. Schmidt, L. Torralba-Díaz, S. Sheykkha, and G. Savvidis (2020). Which model features matter? An experimental approach to evaluate power market modeling choices. *arXiv preprint arXiv:2010.16142*.