Congestion management in power systems - Long-term modeling framework and large-scale application

AUTHORS

Joachim Bertsch

Simeon Hagspiel

Lisa Just

EWI Working Paper, No 15/03

Original Version: June 2015

Updated Version: December 2016

The final publication is available at Springer via
http://dx.doi.org/10.1007/s11149-016-9310-x

Institute of Energy Economics at the University of Cologne (EWI)
www.ewi.uni-koeln.de

Institute of Energy Economics
at the University of Cologne (EWI)

Alte Wagenfabrik
Vogelsanger Straße 321a
50827 Köln
Germany

Tel.: +49 (0)221 277 29-100
Fax: +49 (0)221 277 29-400
www.ewi.uni-koeln.de

CORRESPONDING AUTHOR

Simeon Hagspiel
Department of Economics
University of Cologne
simeon.hagspiel@uni-koeln.de

ISSN: 1862-3808

Congestion management in power systems – Long-term modeling framework and large-scale application

Joachim Bertsch, Simeon Hagspiel and Lisa Just

Abstract

In liberalized power systems, generation and transmission services are unbundled, but remain tightly interlinked. Congestion management in the transmission network is of crucial importance for the efficiency of these inter-linkages. Different regulatory designs have been suggested, analyzed and followed, such as uniform zonal pricing with redispatch or nodal pricing. However, the literature has either focused on the short-term efficiency of congestion management or specific issues of timing investments. In contrast, this paper presents a generalized and flexible economic modeling framework based on a decomposed inter-temporal equilibrium model including generation, transmission, as well as their inter-linkages. The model covers short-run operation and long-run investments and hence, allows to analyze short and long-term efficiency of different congestion management designs that vary with respect to the definition of market areas, the regulation and organization of TSOs, the way of managing congestion besides grid expansion, and the type of cross-border capacity allocation. We are able to identify and isolate implicit frictions and sources of inefficiencies in the different regulatory designs, and to provide a comparative analysis including a benchmark against a first-best welfare-optimal result. To demonstrate the applicability of our framework, we calibrate and numerically solve our model for a detailed representation of the Central Western European (CWE) region, consisting of 70 nodes and 174 power lines. Analyzing six different congestion management designs until 2030, we show that compared to the first-best benchmark, i.e., nodal pricing, inefficiencies of up to 4.6% arise. Inefficiencies are mainly driven by the approach of determining cross-border capacities as well as the coordination of transmission system operators' activities.

JEL classification: C61, C63, D47, E61, L50, Q40

Keywords: Power system economics, unbundling, congestion management, transmission pricing, inter-temporal equilibrium model

1. Introduction

The liberalization of power systems entails an unbundling of generation and grid services to reap efficiency gains stemming from a separate and different organization. While there is competition between generating

We thank Felix Höffler for helpful comments, as well as Tom Brown and Energynautics for their cooperation to realize the large-scale application. The project was funded by the German Federal Ministry for Economic Affairs and Energy based on ruling of the Deutsche Bundestag. The financial support through grant 03ESP239 is gratefully acknowledged. Furthermore, J. Bertsch and S. Hagspiel acknowledge funding of the German research society DFG through grant HO 5108/2-1. The responsibility for the content of this publication lies solely with the authors.

firms, transmission grids are considered a natural monopoly and are operated by regulated transmission system operators (TSOs). However, strong inter-linkages remain between these two parts of the power system: From a transmission perspective, TSOs are responsible for non-discriminatory access of generating units to transmission services while maintaining a secure grid operation. They are thus strongly influenced by the level and locality of generation and load. Furthermore, due to Kirchhoff's laws, operation and investment decisions of one TSO may affect electricity flows in the area of another TSO. From a generation firms' perspective, activities are impacted by restrictions on exchange capacities between markets or operational interventions by the TSOs to sustain a reliable network.

An efficient regulatory design of those inter-linkages between generation and grid will positively affect the overall efficiency of the system, for instance by providing locational signals for efficient investments into new generation or transmission assets. To ensure an efficient coordination of short (i.e., operational) and long-term (i.e., investment) activities in the generation and grid sectors, congestion management has been identified to be of utmost importance (e.g., Chao et al. (2000)). Different regulatory designs and options are available to manage congestion, including the definition of price zones as well as various operational and investment measures. Because it is able to deliver undistorted and hence efficient price signals, nodal pricing is a powerful market design to bring along efficiency. This was shown in the seminal work of Schweppe et al. (1988) and Hogan (1992). Nevertheless, many markets deviate and pursue alternative approaches, e.g., due to historical or political reasons. For instance, most European countries deploy national zonal market areas with uniform electricity prices. Implicitly, several challenges are thus imposed upon the system: First, in zonal markets, intra-zonal network congestion remains unconsidered by dispatch decisions. However, if a dispatch induces intra-zonal congestion (which is typically often the case), it might be necessary to reconfigure the dispatch, known as re-dispatch. Alternatively, the dispatch can be impacted by charging grid costs directly to generators in order to avoid congestion in the market clearing process (a so-called generator- or g-component, also known as grid connection charge). Such charges reflect the locational scarcity of the grid, and are thus conceptually similar to nodal prices, depending on the calculation method applied (see Brunekreeft et al. (2005) for a comprehensive discussion). Second, cross-border capacity needs to be managed. Whereas historically, cross-border capacities have often been auctioned explicitly, many market areas are now turning to implicit market coupling based on different allocation routines, such as net-transfer capacities (NTC) or flow-based algorithms (Brunekreeft et al. (2005), Oggioni and Smeers (2012), Oggioni and Smeers (2013)).[1]

The literature has investigated various regulatory designs to manage congestion in power systems from different perspectives. Static short-term efficiency of nodal pricing – as shown by Schweppe et al. (1988) – was confirmed, e.g, by van der Weijde and Hobbs (2011) who compare nodal pricing and NTC based market coupling in a stylized modeling environment. Furthermore, several papers have quantified the increase in social welfare through a switch from zonal to nodal pricing for static real world case studies (see for example: Green (2007), Leuthold et al. (2008), Burstedde (2012), Neuhoff et al. (2013)). Similarly, Daxhelet and Smeers (2007) show that generator and load components reflecting their respective impact on congestion have a positive effect on static social welfare (as well as its distribution), while Oggioni and Smeers (2012) investigate different congestion management designs in a six node model and find that a single TSO or

[1]Under implicit market coupling, cross-border capacities and prices are implicitly taken into account during the joint clearing process of coupled markets.

multi-lateral arrangements for counter-trading between several TSOs may improve efficiency. Oggioni et al. (2012) and Oggioni and Smeers (2013) show that in a zonal pricing system, the configuration of zones as well as the choice of counter-trading designs have a significant impact on efficiency.

A second line of literature deals with the dynamic long-term effects of congestion management, i.e., the investment perspective. On the one hand, issues of timing (e.g., due to uncertainty or commitment) in settings consisting of multiple players (such as generation and transmission) have been addressed. Höffler and Wambach (2013) find that long-term commitment of a benevolent TSO may lead to inefficient investment decisions due to the locational decisions of investments in generation. In contrast, Sauma and Oren (2006) and Rious et al. (2009) formulate the coordination problem between a generation and a transmission agent as a decomposed problem, and find that a prospective coordinated planning approach as well as transparent price signals entail efficiency gains, though some inefficiencies remain and the first best is not realized. On the other hand, imperfect simultaneous coordination (e.g., due to strategic behavior or hidden information) has been investigated by Huppmann and Egerer (2014) for the case of multiple TSOs being active in an interconnected system. They find that a frictionless coordinated approach outperforms the system outcome with strategic TSOs maximizing social welfare within their own jurisdiction.

With this paper, we contribute to the above literature with a generalized and flexible economic modeling framework for analyzing the short as well as long-term effects of different congestion management designs in a decomposed inter-temporal equilibrium model including generation, transmission, as well as their inter-linkages. Specifically, with our framework we are able to represent, analyze and compare different TSO organizations, market areas (i.e., nodal or zonal pricing), grid expansion, redispatch or g-components, as well as calculation methods for cross-border capacity allocation (i.e., NTC and flow-based). A major advantage of our analytical and numerical implementation is its flexibility to represent different congestion management designs in one consistent framework. We are hence able to identify and isolate frictions and sources of inefficiencies by comparing these different regulatory designs. Moreover, we are able to benchmark the different designs against a frictionless welfare-optimal result, i.e., the "first best". In order to exclusively focus on the frictions and inefficiencies induced by the congestion management designs, we do not address issues of timing, such as uncertainty or sequential moving. Instead, we assume perfect competition, perfect information, no transaction costs, utility-maximizing agents, continuous functions, inelastic demand and an environment where generation and grid problems are solved simultaneously. As an additional contribution, we calibrate and numerically solve our model for a large-scale problem. Specifically, we investigate a detailed representation of the Central Western European (CWE) region.[2] To tackle the complex nature of the optimization problem, we develop a numerical solution algorithm based on decomposition, while a detailed analysis of the convergence behavior suggests that the results obtained may be globally optimal. Thereby, we offer a sound indication on how different congestion management designs perform in practice, and provide empirical evidence that nodal pricing is the efficient benchmark while alternative designs imply inefficiencies of up to 4.6% until 2030.

The paper proceeds as follows: In Section 2, we analytically develop our modeling framework. In Section 3, a numerical solution method to solve this framework is proposed. In Section 4, we apply the methodology to a detailed representation of the CWE region in scenarios up to the year 2030. Section 5 concludes and

[2]The CWE region is one of seven regional initiatives to bring forward European market integration. The countries within this area are Belgium, France, Germany, Luxemburg and the Netherlands.

provides an outlook on future research.

2. Economic framework

In order to develop a consistent analytical modeling framework for different congestion management designs, we start with the well-known model for an integrated optimization problem for planning and operating a power system.[3] By design, this model does not contain any frictions and inefficiencies. Hence, the results obtained are necessarily first best and may serve as the efficient benchmark for alternative settings. Moreover, it corresponds to the concept of nodal pricing as introduced by Schweppe et al. (1988).[4]

To depict various congestion management designs, we make use of the possibility to separate an integrated optimization problem into multiple levels (or, in other words, subproblems). Even though the model structure is then different, it can be shown that both formulations of the problem yield the same results. However, in the economic interpretation we can take advantage of the separated model structure representing unbundled generation and transmission sectors. On the generation stage, competitive firms decide about investments in and dispatch of power plants, whereas the transmission stage consists of one or multiple TSOs that efficiently expand and operate transmission grid capacities.[5] Lastly, with generation and transmission separated, we are able to introduce six practically relevant congestion management designs through the manipulation of the exchange of information between and among the two levels, and show how they deviate from the first best.

Even though the modeling framework would allow to study an extensive range of congestion management designs, we restrict our attention to four settings (and two additional variations) that are both, relevant in practical applications and sufficiently different from each other. Specifically, our settings vary in the definition of market areas (nodal or coupled zonal markets), the regulation and organization of TSOs (one single TSO for all zones or several zonal TSOs), the way of managing congestion besides grid expansion (redispatch and g-component) and different alternatives for cross-border capacity allocation (NTC vs. flow-based market coupling). We consider Net Transfer Capacity (NTC) and flow-based market coupling as cross-border capacity allocation algorithms because they have been used extensively in the European context (see, e.g., Glachant (2010)). NTCs are a rather simplified version of cross-border trade restrictions, widely neglecting the physical properties of the grid as well as its time-varying characteristics. Under flow-based market coupling, cross-border transmission capacities are calculated taking into account the impact of (cross-border) line flows on every line in the system (e.g., Oggioni and Smeers (2013)), hence providing a much better consideration of the physical grid properties which is crucially important in case of meshed networks. As a consequence, more capacity can generally be offered for trading between markets, and a better usage of existing infrastructures is achieved. The analyzed settings are summarized in the following Table 1.

Noticeably, despite the separated generation and transmission levels, agents are in all settings assumed to act rationally and simultaneously while taking into account the activities of the other stage.[6] Furthermore, we assume perfect competition on the generation stage and perfect regulation of the TSOs in the sense

[3] Such a model is typically applied to represent the optimization problem of a social planner or an integrated firm optimizing the entire electricity system, including generation and transmission.

[4] One main difference in our model is the assumption of an inelastic demand which was necessary to formulate and solve the model as a linear program. We will elaborate on this issue in Section 2.1.

[5] Efficient in this context means that the TSO(s) are perfectly regulated to expand and operate the grid at minimal costs.

[6] I.e., sequential moving and issues of timing are not considered.

	Market area and coupling	TSO scope	TSO measures
I	Nodal markets	One TSO	Grid expansion
II - NTC	Zonal markets, NTC-based coupling	One TSO	Grid expansion, zonal redispatch
II - FB	Zonal markets, Flow-based coupling	One TSO	Grid expansion, zonal redispatch
III - NTC	Zonal markets, NTC-based coupling	Zonal TSOs	Grid expansion, zonal redispatch
III - FB	Zonal markets, Flow-based coupling	Zonal TSOs	Grid expansion, zonal redispatch
IV	Zonal markets	Zonal TSOs	Grid expansion, zonal g-component

Table 1: Analyzed congestion management designs

that TSO activities are aligned with social objectives. TSOs as well as generators are price taking, with an independent institution (e.g., the power exchange) being responsible for coordinating the activities of the different participating agents and for market clearing.[7] Importantly, while in the first best design all information is available to all agents, alternative congestion management designs may induce an adverse (e.g., aggregated) availability of information. The solution of the problem is an intertemporal equilibrium which would be unique if the optimization was strictly convex. Unfortunately, we were not able to rigorously verify (strict) convexity in our case. However, we will thoroughly discuss the issue in the context of the numerical implementation in Section 3. Noticeably, with the above assumptions, our general modeling approach can be thought of as a way to compare today's and future performances of different congestion management designs based on today's state of the system, today's information horizon, as well as rational expectations about future developments and resulting investment decisions.[8]

For the sake of readability (and in contrast to the large-scale application presented in Section 4), we make some simplifications in the theoretical framework: dispatch decisions are realized in several points of time, but invest decisions are undertaken only once. Furthermore, we neglect different types of generation technologies that may be available at a node. This simplification does not change any of the conclusions drawn from the theoretical formulation.[9]

For developing the economic modeling framework in the following subsections, we will deploy parameters, variables and sets as depicted in Table A.2 in Appendix Appendix A.

2.1. Setting I – First Best: Nodal pricing with one TSO

By design, nodal pricing avoids any inefficiency by covering and exchanging all information present within the problem – leading to a welfare optimal electricity system. It hence represents the first best setting in our analysis of different congestion management designs. With the assumption of a social planner or perfect competition and regulation, nodal prices can be derived from locational marginal costs (of generation and capacity) in a market clearing that implicitly considers the physical properties of the electricity network (specifically, loop flows). Abstracting from economies of scale and lumpiness of investment, it can be shown

[7]By assuming perfect competition and an inelastic demand, we are able to treat the general problem as a cost minimization problem. This assumption is commonly applied for formulation of electricity markets in the literature. An alternative formulation with a welfare maximization approach would be possible, but wouldn't impact the general conclusions.

[8]In our numerical application, this approach is supplemented with discounted future cash flows. See Section 4 for further details.

[9]To include multiple instances in time for investments, the formulation could easily be adapted by adding an index to all parameters, variables and equations related to installed capacities (generation and transmission). In the same vein, an additional index could be inserted to account for different types of generation technologies.

that an efficient and unique equilibrium exists under nodal prices (Caramanis (1982), Joskow and Tirole (2005), Rious et al. (2009)). In line with these findings, we assume constant marginal grid costs as well as continuous generation and transmission expansion.[10] Another assumption in our formulation is an inelastic (yet time-varying) demand. The reason for assuming an inelastic demand is mainly triggered by the excessive computational burden that would be induced by an elastic demand in the numerical solution approach (an inelastic demand allows us to formulate and solve the model as a linear instead of a non-linear program). As a drawback, the assumption of an inelastic demand differs from the formulation in Schweppe et al. (1988) and leads to the artifact that demand can never set the price. However, scarcity rents to cover capacity costs are still possible under perfect information and competition (including entry and exit of generators). For instance, consider the bid of a peak load plant during a single peak load hour when it is dispatched and pivotal. The bid will consist of the variable costs plus the long-term marginal costs of the capacity. If the bid was lower, the peak load plant would leave the market due to an overall loss. If the bid was higher, another peak load plant would enter the market due to the possibility of making a profit. This forces the peak load plant to bid its true variable plus marginal capacity costs. Once accepted, this bid can be interpreted as the resulting market prices under capacity scarcity. Lastly, note that off-peak hours can also have capacity components in prices if there is a diversified mix of generation technologies, characterized by different cost structures.

The following optimization problem *P1* is similar to the formulation of an integrated problem for operating generation and transmission as in Schweppe et al. (1988), except for the major change of demand being inelastic. In this formulation, a social planner or an integrated firm minimizes total system costs of the operation and investment of generation and transmission.

P1 *Integrated Problem*

$$\min_{\overline{G}_i, G_{i,t}, T_{i,j,t}, \overline{P}_{i,j}} \quad X = \sum_i \delta_i \overline{G}_i + \sum_{i,t} \gamma_{i,t} G_{i,t} + \sum_{i,j} \mu_{i,j} \overline{P}_{i,j} \tag{1a}$$

$$\text{s.t.} \quad G_{i,t} - \sum_j T_{i,j,t} = d_{i,t} \qquad \forall i,t \qquad | \, \lambda_{i,t} \tag{1b}$$

$$G_{i,t} \leq \overline{G}_i \qquad \forall i,t \tag{1c}$$

$$|T_{i,j,t}| = |P_{i,j,t}(\overline{P}_{k,l}, G_{k,t}, d_{k,t})| \leq \overline{P}_{i,j} \qquad \forall i,j,t \qquad | \, \kappa_{i,j,t} \tag{1d}$$

$$T_{i,j,t} = -T_{j,i,t} \qquad \forall i,j,t \tag{1e}$$

Indices i, j, k, l represent nodes in the system. Generation $G_{i,t}$, generation capacity $\overline{G}_i$, trade $T_{i,j,t}$ and transmission capacity $\overline{P}_{i,j}$ are optimization variables. Additional capacities can be installed at the costs of δ_i for generation and $\mu_{i,j}$ for transmission. Nodal prices are derived from the dual variables $\lambda_{i,t}$ of the equilibrium constraint which states that the demand level $d_{i,t}$ at node i can be either satisfied by generation at the same node or trade between nodes (Equation (1b)). Equations (1c) and (1d) mirror that generation is restricted by installed generation capacities, and physical flows by installed transmission capacities. Furthermore, trades from node i to node j are necessarily equal to negative trades from node j to node i (Equation (1e)). As the market clearing fully accounts for the transmission network in the

[10] This assumption is certainly more critical for transmission investments which require a certain magnitude to be realized. Generation investment might also be lumpy, but smaller plant sizes are possible.

nodal pricing regime, trade between adjacent nodes is equal to physical flows on the respective line, i.e., $T_{i,j,t} = P_{i,j,t}$ (Equation (1d)).

Load flows on transmission lines are based on Kirchhoff's laws, which we represent based on a linearized load flow approach.[11] Thereby, flows are impacted by generation ($G_{k,t}$) and demand ($d_{k,t}$), i.e., power balances of all nodes in the system, as well as by the physical properties of the transmission system, represented by installed transmission capacities $\overline{P}_{k,l}$. Thus, there is a functional dependency of flows and trades on generation, demand, and line capacities throughout the system, i.e., $T_{i,j,t} = T_{i,j,t}(\overline{P}_{k,l}, G_{k,t}, d_{k,t})$.

As has been shown, e.g., by Conejo et al. (2006), an integrated optimization problem can be decomposed into subproblems which are solved simultaneously, while still representing the same overall situation and corresponding optimal solution. In our application, we take advantage of this possibility to represent separated generation and transmission levels in problem *P1'*. The generation stage *P1'a* states the market clearing of supply and demand while respecting generation capacity constraints. As in *P1*, the same nodal prices are obtained by the dual variable $\lambda_{i,t}$ of the equilibrium constraint (2b). Instead of including the explicit grid expansion costs in the cost minimization, the objective function of the generation stage now contains transmission costs which assign transmission prices $\kappa_{i,j,t}$ to trade flows between two nodes i and j. These prices are derived from the dual variable of the equilibrium constraint on the transmission stage (Equation (2g)). We assume that the TSO is perfectly regulated to minimize costs of grid extensions accounting for the physical feasibility of the market clearing as determined on the generation stage while considering all grid flows and related costs (problem *P1'b*). As trade is a function of $\overline{P}$, which in turn is the decision variable in the transmission problem, the market clearing conditions need to reoccur in the transmission problem.

P1'a *Generation*

$$\min_{\overline{G}_i, G_{i,t}, T_{i,j,t}} \quad X = \sum_i \delta_i \overline{G}_i + \sum_{i,t} \gamma_{i,t} G_{i,t} + \sum_{i,j,t} \kappa_{i,j,t} T_{i,j,t} \tag{2a}$$

$$\text{s.t.} \qquad G_{i,t} - \sum_j T_{i,j,t} = d_{i,t} \qquad \forall i,t \qquad | \, \lambda_{i,t} \tag{2b}$$

$$G_{i,t} \leq \overline{G}_i \qquad \forall i,t \tag{2c}$$

$$T_{i,j,t} = -T_{j,i,t} \qquad \forall i,j,t \tag{2d}$$

P1'b *Transmission*

$$\min_{\overline{P}_{i,j}} \quad Y = \sum_{i,j} \mu_{i,j} \overline{P}_{i,j} \tag{2e}$$

$$\text{s.t.} \qquad G_{i,t} - \sum_j T_{i,j,t} = d_{i,t} \qquad \forall i,t \tag{2f}$$

$$|T_{i,j,t}| = |P_{i,j,t}(\overline{P}_{k,l}, G_{k,t}, d_{k,t})| \leq \overline{P}_{i,j} \qquad \forall i,j,t \qquad | \, \kappa_{i,j,t} \tag{2g}$$

$$T_{i,j,t} = -T_{j,i,t} \qquad \forall i,j,t \tag{2h}$$

As can be seen, all terms of *P1* reappear in *P1'*, however, allocated to two separated levels. Mathematically, the equivalence of *P1* and *P1'* is shown in Appendix Appendix C, where the first order conditions

[11] We will use the PTDF approach shown in Appendix Appendix B in our numerical implementation in Section 3, as this enables a linearization of the generally non-linear load flow problem, given a fixed transmission network (cf. Hagspiel et al. (2014)).

of both formulations are compared.

2.2. Setting II: coupled zonal markets with one TSO and zonal redispatch

In zonal markets, a number of nodes are aggregated to a market with a uniform price. In contrast to nodal pricing, coupled zonal markets only consider aggregated cross-border capacities between market zones during market clearing (instead of all individual grid elements). Thus, the obtained prices for generation do not reflect the true total costs of the entire grid infrastructure. This is due to the fact that zonal prices only reflect those cross-border capacities that limit activities between zonal markets. Cross-border capacities can be allocated in different ways. We consider Net Transfer Capacity (NTC) and the more sophisticated flow-based market coupling as cross-border capacity allocation algorithms (see Oggioni and Smeers (2013)). Under the latter regime, more capacity can generally be offered for trading between markets, and a better usage of existing infrastructures is achieved.

Because intra-zonal congestion is neglected in the zonal market-clearing, it needs to be resolved in a subsequent step by the TSO. Besides the expansion of grid capacities, in *Setting II* we provide the TSO with the opportunity of zonal redispatch. The TSO may instruct generators located behind the bottleneck to increase production (positive redispatch), and another generator before the bottleneck to reduce production (negative redispatch).[12] We assume here a perfectly discriminating redispatch: the TSO pays generators that have to increase their production their variable costs, and in turn receives the avoided variable costs of generators that reduce their supply.[13] As the generator with positive redispatch was not part of the original dispatch, it necessarily has higher variable costs than the generator that reduces supply. Thus, the TSO has to bear additional costs that are caused by the redispatch which amount to the difference between the variable costs of the redispatched entities. Assuming further that the TSO has perfect information about the variable costs of the generating firms, redispatch measures of the TSO have no impact on investment decisions of generating firms as the originally dispatched generation capacity is still able to cover capital costs from the spot market result. As the TSO is assumed to have perfect information, no strategic behavior (e.g., through overscheduling of transactions) is possible that would create artifical congestion and necessitate redispatch. However, additional costs for the economy are induced by inefficient investment decisions of those generators that are not aligned with the overall system optimum due to missing locational price signals.

In the formulation of problem *P2a* zonal pricing is represented by the zonal market indices n, m, each containing one or several nodes i. Market clearing, depicted by the equilibrium Equation (3f), now takes place on zonal instead of nodal markets. The corresponding dual variable $\lambda_{m,t}$ represents zonal prices, which do not include any grid costs except for cross-border capacities. This is indicated by the term $\sum_{m,n,t} \kappa_{m,n,t} T_{m,n,t}$ instead of the nodal formulation (with $\kappa_{i,j,t}$) above. Transmission prices are determined on the transmission stage (Equation (3j)). However, contrary to nodal pricing, these prices are calculated based on some regulatory rule (e.g., NTC or FB) and are thus inherently incomplete since they do not

[12]Redispatch is always feasible due to the fact that the TSO can foresee congestion and hence, counteract by expanding line capacities. Note that in practice, this might be a critical prerequisite that can not always be easily fulfilled, especially when line expansions are impossible or delayed. In fact, this was a key consideration in Texas for moving to a nodal design (e.g., see Baldick and Niu (2005)). For a European context, the aspect is studied thoroughly in Ehrenmann and Smeers (2005) and Bertsch et al. (2015).

[13]It is noteworthy that this assumption refers to European electricity market design, while it would not hold for the zonal designs in California and ERCOT (before they changed to nodal prices). In the latter markets, redispatched generators were settled at uniform clearing prices set by the most expansive unit for increasing and cheapest unit for decreasing generators.

represent real grid scarcities.[14] In addition to grid expansion, the TSO may relieve intra-zonal congestion and optimize the situation by means of redispatch measures $R_{i,t}$ at costs of $\gamma_{i,t}R_{i,t}$.

P2a *Generation*

$$\min_{\overline{G}_i, G_{i,t}, T_{m,n,t}} \quad X = \sum_i \delta_i \overline{G}_i + \sum_{i,t} \gamma_{i,t} G_{i,t} + \sum_{m,n,t} \kappa_{m,n,t} T_{m,n,t} \tag{3a}$$

$$\text{s.t.} \quad \sum_{i \in \mathbf{I_m}} G_{i,t} - \sum_n T_{m,n,t} = \sum_{i \in \mathbf{I_m}} d_{i,t} \quad \forall m,t \quad | \, \lambda_{m,t} \tag{3b}$$

$$G_{i,t} \leq \overline{G}_i \quad \forall i,t \tag{3c}$$

$$T_{m,n,t} = -T_{n,m,t} \quad \forall m,n,t \tag{3d}$$

P2b *Transmission*

$$\min_{\overline{P}_{i,j}, R_{i,t}} \quad Y = \sum_{i,j} \mu_{i,j} \overline{P}_{i,j} + \sum_{i,t} \gamma_{i,t} R_{i,t} \tag{3e}$$

$$\text{s.t.} \sum_{i \in \mathbf{I_m}} G_{i,t} - \sum_n T_{m,n,t} = \sum_{i \in \mathbf{I_m}} d_{i,t} \quad \forall m,t \tag{3f}$$

$$|T_{i,j,t}| = |P_{i,j,t}(\overline{P}_{k,l}, R_{k,t}, G_{k,t}, d_{k,t})| \leq \overline{P}_{i,j} \quad \forall i,j,t \quad | \, \kappa_{i,j,t} \tag{3g}$$

$$\sum_{i \in \mathbf{I_m}} R_{i,t} = 0 \quad \forall m,t \tag{3h}$$

$$0 \leq G_{i,t} + R_{i,t} \leq \overline{G}_i \quad \forall i,t \tag{3i}$$

$$\kappa_{m,n,t} = g(\kappa_{i,j,t}) \tag{3j}$$

$$T_{m,n,t} = -T_{n,m,t} \quad \forall m,n,t \tag{3k}$$

The following two examples illustrate the fundamental differences between *Setting I* and *II*.

Example for 2 nodes and 2 markets: If the electricity system consists of 2 nodes and 2 markets (Figure 1, left hand side), *Setting I* and *II* are identical: There is only one element $i \in \mathbf{I_m}$, such that Equation (3h) fixes variable R_i to zero. Equation (3i) is then no longer relevant, and the cost term of redispatch in the objective function ($\sum_{i,t} \gamma_{i,t} R_{i,t}$) becomes zero. The only difference remaining between *1'b* *P2b* is then Equation 3j. However, due to $I = M$, it follows that $\kappa_{m,n,t} = \kappa_{i,j,t}$, which, inserted on the generation level, yields equivalence of problems *P1'* and problem *P2* for the chosen example.

Example for 3 nodes and 2 markets: Figure 1, right hand side, shows an electricity system consisting of two markets m and n, where m includes one node (1) and n two nodes $(2,3)$ at a point in time t. Function g for calculating the transmission price $\kappa_{m,n,t}$ (Equation (3j)) between the markets has to be defined, e.g., by averaging the single line prices $\kappa_{m,n,t} = (\kappa_{1,2,t} + \kappa_{1,3,t})/2$. Still, the TSO cannot supply the locational fully differentiated prices $\kappa_{1,2,t}, \kappa_{1,3,t}$ and $\kappa_{2,3,t}$ to the market, and hence, efficient allocation of investments is (partly) achieved *between* the markets, but not *within* the markets. Redispatch does not fully solve this problem, because it is revenue-neutral and does not affect the investment decision.

[14] Note that the duality of the problem would also allow for an alternative formulation of the cross-border transmission constraint by means of quantity constraints instead of prices. Hence, the cost of transmission in the objective function of the generation stage ($\sum_{m,n,t} \kappa_{m,n,t} T_{m,n,t}$) would disappear and an additional constraint for trading would be implemented ($|T_{m,n,t}| \leq C_{m,n}, \forall m,n,t$). The restriction of trading volumes $C_{m,n,t}$ would be calculated on the transmission stage *P2b* via a constraint $C_{m,n} = h(\overline{P}_{i,j})$ instead of the prices $\kappa_{m,n,t}$. These prices would then be the dual variable of the volume constraint on the generation stage, and necessarily coincide with $\kappa_{m,n,t}$.

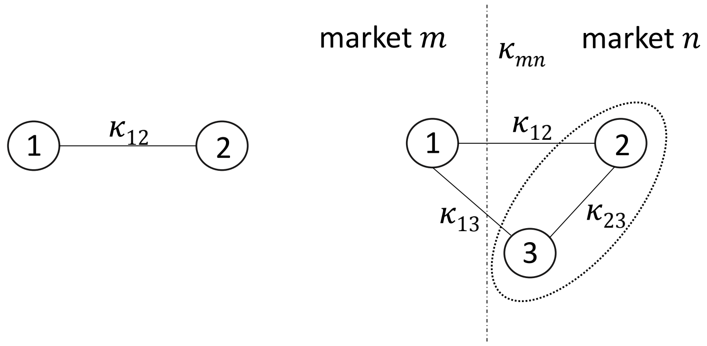

Figure 1: Two simple examples. Left: 2 nodes, 2 markets. Right: 3 nodes, 2 markets

Overall, *Settings I* and *II* differ in the way grid costs are reflected on the generation stage. Specifically, *Setting II* lacks locational differentiated prices, thus impeding efficient price signals $\kappa_{i,j,t}$ for the generation stage. Of course, the level of inefficiency depends substantially on the regulatory rule determining the calculation of prices based on a specification of function $g(\kappa_{i,j,t})$. In general, it is clear that the closer the specification of g reflects real-time conditions and the more it enables the full usage of existing grid infrastructures, the more efficiently the general problem will be solved. While we limit our analysis in this section to this general finding, we will discuss two possible specifications often implemented in practice (NTC and flow-based market coupling) in the empirical example in Section 4. Given the inefficiency induced by the specification of function g, the question remains whether and how redispatch measures may help to relieve the problem. We find that the resulting inefficiency cannot be fully resolved by redispatch because the latter remains a zonal measure (Equation (3h)). Hence, the TSO cannot induce an efficient usage of generation and transmission across zonal borders. Furthermore, investments into generation capacities are not influenced by redispatch and only zonal prices as well as their costs are considered.[15] Hence, the setting lacks locational signals for efficient generation investments within zonal markets.

2.3. Setting III: coupled zonal markets with zonal TSOs and zonal redispatch

In this setting, we consider zonal markets with zonal TSOs being responsible for grid expansion as well as a zonal redispatch. Thus, the problem on the generation stage remains exactly the same as in the previous setting (i.e., *P3a* = *P2a*). However, the transmission problem changes, such that now multiple zonal TSOs are considered. Each TSO solves its own optimization problem according to the national regulatory regime (in our case corresponding to a cost-minimization within the zones). Formally, problem *P3b*, now consists of multiple separate optimization problems for each zonal TSO, with the objective to minimize costs from zonal grid as well as from zonal redispatch measures. However, cross-border line capacities are also taken into account. As these are by definition located within the jurisdiction of two adjacent market areas, the

[15]For obtaining a unique equilibrium we assume that costs differ over all nodes, such that decisions for generation and investments are unambiguously ordered.

two corresponding TSOs have to negotiate about the extension of these cross-border capacities. In fact, cross-border capacities built by two different TSOs may be seen as a Leontief production function, due to the fact that the line capacities built on each side are perfect complements. Corresponding costs from inter-zonal grid extensions are assumed to be shared among the TSOs. Due to the fact that situations may arise where an agreement on specific cross-border lines between neighboring TSOs cannot be reached (which would imply that an equilibrium solution cannot be found), we assume the implementation of a regulatory rule that ensures the acceptance of a unique price for each cross-border line by both of the neighboring TSOs. For instance, the regulatory rule may be specified such that both TSOs are obliged to accept the higher price offer, or, equivalently, the lower of the two capacities offered for the specific cross-border line.

As a consequence, grid capacities, especially cross-border capacities, are extended inefficiently as they do not result from an optimization of the entire grid infrastructure. In addition – just as in the previous setting – inefficient investment incentives for generation and grid capacities are caused by the lack of locational differentiated prices. Hence, overall, system outcomes in *Setting III* must be inferior or at most equal to those of *Setting II*.[16]

The mathematical program as well as further technical details of *Setting III* can be found in the Appendix Appendix D.

2.4. Setting IV: coupled zonal markets with zonal TSOs and g-component

In this last setting, we again consider coupled zonal markets with zonal TSOs. However, instead of having the possibility to perform a zonal redispatch (as in *Setting III*), zonal TSOs may now determine local, time-varying prices for generators, i.e., a g-component, at each node belonging to its zone to cope with intra-zonal congestion. A g-component charges grid costs directly to generators in order to avoid congestion in the market clearing process reflecting the impact of generators on the grid at each node and each instant of time. Thus, grid costs are being transferred to the generating firms which consider them in their investment and dispatch decision. In other words, TSOs are able to provide locationally differentiated prices (and hence, generation and investment incentives) for generators within their zone. Noticeably, we do not consider an international g-component here as this would yield the same results as a nodal pricing regime due to generators considering the full set of information concerning grid costs. However, two frictions that may cause an inefficient outcome of this setting remain. When determining nodal g-components, zonal TSOs only consider grid infrastructures within their zone, and not within the entire system. Furthermore, as in *Setting III*, the desired expansion of cross-border lines, which is here assumed to be solved by some regulatory rule ensuring successful negotiation, may deviate between/across neighboring TSOs.

The mathematical program as well as further technical details of *Setting IV* can be found in the Appendix Appendix E.

3. Numerical solution approach

Our approach to numerically solve the problem depicted in the previous section builds on the concept of decomposition. In fact, it follows the approach already applied in the context of *Setting I* (Section

[16] The only mathematical difference of problem *P3b* compared to *P2b* is that the transmission level is partitioned into several optimization problems that are solved separately from each other. Hence, compared to problem *P2b* where the transmission level is solved comprehensively, this represents a more restrictive problem that must be inferior (or at most equal) to the one of *P3b*.

2.1), where we decomposed the integrated problem into two separate levels that are solved simultaneously and showed that they can – in economic terms – be interpreted as generation and transmission levels. Algorithmically, according to Benders (1962), decomposition techniques can be applied to optimization problems with a decomposable structure that can be advantageously exploited. The idea of decomposition generally consists of splitting the optimization problem into a master and one or several subproblems that are solved iteratively. For the problem we are dealing with, namely the simultaneous optimization of generation and grid infrastructures under different congestion management designs and a varying number of TSOs, decomposing the overall problem entails two major advantages: First, the decomposition allows to easily implement variations of the generation and transmission levels including the underlying congestion management design. Hence, the model can be flexibly adjusted to represent the various settings described in the previous section. Second, the iterative nature of the solution process resulting from the decomposition allows to readily update PTDF matrices every time changes in the grid infrastructure have been made, according to Equation (B.11) and the PTDF calculation procedure presented in Section Appendix B. This iterative update of the grid properties, as applied in Hagspiel et al. (2014) and Ozdemir et al. (2015), successively linearizes the non-linear optimization problem to ensure a consistent representation of generally non-linear grid properties, and allows for solving a corresponding linear problem.[17] In turn, linear problems can be solved effectively for global optima using standard techniques, such as the Simplex algorithm (e.g., Murty (1983)).

Even though the PTDF update ties in nicely with the iterative solution of the decomposed problem, it also imposes a particular challenge stemming from the non-linearity in the PTDF calculation (see Appendix Appendix B). Specifically, despite the successive linearization and iterative solution, the non-linearity of the transmission expansion problem remains. Hence, neither the existence and uniqueness of a global optimum of the problem, nor the convergence of the solution algorithm can generally be guaranteed (e.g., Bazaraa et al. (2006)).[18] However, we can build on numerical experience that has been gained by two papers that are closely related to ours in terms of the algorithmic approach: The analysis in Hagspiel et al. (2014) is closest to our application as they deploy the same successive PTDF update to co-optimize generation and transmission assets (including operation and investment). Based on numerical examples, they show that the algorithm converges in a large number of configurations, including small analytically tractable test systems as well as large-scale applications. Furthermore, they do not detect issues of multiple equilibria in their analysis. In a very similar vein, Ozdemir et al. (2015) develop a methodology based on successive linear programming and Gauss-Seidel iteration to jointly optimize transmission and generation capacities. They report that even though they cannot guarantee convergence or global optimality either, their approach shows good performance. In the course of preparing the results presented in this paper, we were able to confirm the above findings in several model runs where we varied starting values over a broad range and did not find evidence neither against convergence nor against uniqueness of our optimum. Hence, even though not guaranteed, empirical evidence indicates that we are facing a numerical problem that we are able to

[17] Accordingly, in our model $PTDF$ is depicted as a parameter that is updated in each iteration instead of a variable.

[18] Noticeably, this would change if the problem was strictly convex. Then, there would be a unique equilibrium, corresponding to a global optimum. Furthermore, deploying a Benders-type decomposition, the algorithm would preserve convexity and guarantee that the iterative solution converges towards this global optimum (Benders (1962) and, e.g., Conejo et al. (2006) for a general overview). Unfortunately, however, to the best of our knowledge a thorough rigorous analytical analysis of the properties and solvability of the transmission expansion problem and particular instances therein is still missing. Meanwhile, it would be beyond the scope of this paper to approach this challenging problem.

reliably solve with our algorithm while converging towards an optimal solution. In our application, the optimal solution represents an intertemporal equilibrium without uncertainty. However, note that due to the converging nature of the algorithm, we can only *approach optimality while retaining an optimality gap (i.e., a deviation from the optimal problem solution). As the gap is expected to decrease, this results in a tradeoff between accuracy and solution time. Numerically, the tradeoff is solved by setting a convergence threshold.* Interestingly, in economic terms, the iterative algorithm to solve the decomposed problem can be readily interpreted as a price adjustment by a Walrasian auctioneer, also know as tatonnement procedure (e.g., Boyd et al. (2008)).

With some minor modifications, we can directly follow the (economically intuitive) formalization developed in the previous section and implement separate optimization problems representing the different tasks of generation and grid as well as the various settings (*I-IV*). We follow the Benders decomposition approach described in Conejo et al. (2006), while considering the transmission capacities as complicating variables. We define the generation stage as the master problem, whereas the subproblem covers the transmission stage.[19] The principle idea of the solution algorithm is to solve the simultaneous generation and transmission stage problem iteratively, i.e., in a loop that runs as long as some convergence criterion is reached. In this process, optimized variables and marginal values are exchanged between the separated generation and grid levels reflecting the configuration of congestion management and TSO organization. For the settings described in the previous section, prices, which are iterated and thus adjusted, differ with respect to the information they contain and hence determine to which degree efficiency can be reached. Compared to nodal pricing (*Setting I*), the other settings provide prices or products that describe the underlying problem only incompletely – and hence, entail an inefficient outcome.

The numerical algorithm to solve the nodal pricing model is sketched below. Parameters that save levels of optimal variables for usage in the respective other stage are indicated by $^{(\cdot)}$. It should be noticed that for the sake of comprehensibility, we still represent a simplified version of a more complete power system model that would need to account for multiple instances in time for investments, multiple generation technologies, etc. However, the extension is straightforward and does not change the principle approach depicted here.

Information passed from the transmission to the generation stage is captured by α, for which a Benders cut (lower bound constraint) is added in each iteration u up to the current iteration v (Equation (4e)). This Benders cut consists of total grid costs $Y^{(u)}$ as well as the marginal costs each unit of trade $T_{i,j,t}$ is causing in the grid per node, denoted by $\kappa^{(u)}_{i,j,t}$. Both pieces of information are provided in the highest possible temporal and spatial resolution. As these components occur in the objective function of the generation stage (via α), the optimization will try to avoid the additional costs it is causing on the transmission stage, e.g., by moving power plant investments to alternative locations. The variable α is needed to correctly account for the impact of the transmission on the generation stage. On the transmission stage, the TSO is coping with the exchange (i.e., trade) of power stemming from the dispatch situation delivered by the master problem, thereby determining the marginal costs the trade is causing on the transmission stage, i.e., κ. Power flows are calculated by linearized load-flow equations represented by PTDF matrices mapping. The TSO then expands the grid such that it supports the emerging line flows at minimal costs.

[19]Noticeably, the model could be inverted such that the master problem represents the grid sector which would, however, not change any of the results obtained.

$v = 1$; convergence=false

While(convergence=false) {

Master problem: generation

$$\min_{\overline{G}_i, G_{i,t}, T_{i,j,t}, \alpha} \quad X = \sum_i \delta_i \overline{G}_i + \sum_{i,t} \gamma_{i,t} G_{i,t} + \alpha \tag{4a}$$

$$\text{s.t.} \qquad G_{i,t} - \sum_j T_{i,j,t} = d_{i,t} \qquad \forall i, t \tag{4b}$$

$$G_{i,t} \le \overline{G}_i \qquad \forall i, t \tag{4c}$$

$$T_{i,j,t} = -T_{j,i,t} \qquad \forall i, j, t \tag{4d}$$

$$Y^{(u)} + \sum_{i,j} \kappa^{(u)}_{i,j,t} \cdot (T_{i,j,t} - T^{(u)}_{i,j,t}) \le \alpha \qquad \forall u = 1, ..., v-1 | v > 1 \tag{4e}$$

$$G^{(v)}_{i,t} = \text{Optimal value of } G_{i,t} \qquad \forall i, t \tag{4f}$$

Sub-problem: transmission

$$\min_{\overline{P}_{i,j}, T_{i,j,t}} \quad Y = \sum_{i,j} \mu_{i,j} \overline{P}_{i,j} \tag{4g}$$

$$\text{s.t.} \qquad |P_{i,j,t}| = \left| \sum_k PTDF^{(v)}_{k,i,j} \cdot (G^{(v)}_{k,t} - d_{k,t}) \right| \le \overline{P}_{i,j} \quad \forall i, j, t \quad | \, \kappa^{(v)}_{i,j,t} \tag{4h}$$

$$Y^{(v)} = \text{Optimal value of } Y \tag{4i}$$

$$PTDF^{(v)} = \text{PTDF matrix calculated based on } \overline{P}_{i,j} \tag{4j}$$

if(convergence criterion < threshold; convergence=true)

$v = v + 1$

};

As regards the representation of settings *II-IV*, only very few modifications are needed compared to the nodal pricing regime (*Setting I*). The numerical algorithmic implementation of the various settings and modifications directly follows the procedure discussed in Section 2 and is thus not discussed again in detail here.[20]

4. Large-scale application

In this section, we apply the previously developed methodology to a detailed representation of the power sector in the Central Western European (CWE) region up to the year 2030. The application demonstrates the suitability of the modeling framework for large-scale problems and allows to assess and quantify the welfare losses in the considered region caused by different congestion management designs.

Given its historical, current and foreseen future development, the CWE region appears to be a particularly timely and relevant case study for different congestion management designs. In order to increase the market

[20] Nevertheless, for the sake of completeness and reproducibility, we have included one more complete model formulation illustrating the main differences of the other settings in Appendix Appendix F.

integration of European electricity markets towards an internal energy market, the European Union (EU) has declared the coupling of European electricity markets, which are organized in uniform price zones, an important stepping stone (see, e.g., Glachant (2010)). As for the cross-border capacity allocation, after a phase of NTC (Net Transfer Capacities) based market coupling, the CWE region is currently implementing a flow-based market coupling which is expected to increase the efficiency of the utilization of transmission capacities as well as overall social welfare (Capacity Allocating Service Company (2014)). Even though nodal pricing regimes have often been discussed for the European power sector (see, e.g., Ehrenmann and Smeers (2005) or Oggioni and Smeers (2012)), it can be expected that uniform price zones that correspond to national borders will remain. In fact, zonal markets coupled via a flow-based algorithm have been declared the target model for the European power sector (ACER (2014)).

In each zonal market, the respective zonal (i.e., national) TSO is responsible for the transmission network. Thereby, TSOs are organized and regulated on a national level, such that they can be assumed to care mainly about grid operation and expansion planning within their own jurisdiction. Although there are an umbrella organization (ENTSO-E) and coordinated actions, such as the (non-binding) European Ten-Year-Network-Development-Plan (TYNDP), the incentives of the national regulatory regime to intensify cross-border action might fall short of effectiveness. At the same time, Europe is heavily engaged in the large-scale deployment of renewable energies, hence causing fundamental changes in the supply structure. Generation is now often built with respect to the availability of primary renewable resources, i.e., wind and solar irradiation, and not necessarily close to load. This implies that the current grid infrastructure is partly no longer suitable and needs to be substantially redesigned, rendering an efficient congestion management even more important than before.

4.1. Model configuration and assumptions

The applied model for the generation stage belongs to the class of partial equilibrium models that aim at determining the cost-optimal electricity supply to customers by means of dispatch and investments decisions based on a large number of technological options for generation. As power systems are typically large and complex, these models are commonly set up as a linear optimization problem which can efficiently be solved. Our model is an extended version of the linear long-term investment and dispatch model for conventional, renewable, storage and transmission technologies as presented in Richter (2011) and applied in, e.g., Jägemann et al. (2013) or Hagspiel et al. (2014). In contrast to previous versions, the CWE region, i.e., Belgium, France, Germany, Luxembourg and Netherlands, is considered with a high spatial (i.e., nodal) resolution. In order to account for exchanges with neighboring countries, additional regions are defined, but at an aggregated level: Southern Europe (Austria, Italy and Switzerland), South-West Europe (Portugal and Spain), North-West Europe (Ireland and UK), Northern Europe (Denmark, Finland, Norway and Sweden), and Eastern Europe (Czech Republic, Hungary, Poland, Slovakia and Slovenia). Figure 2 depicts the regional coverage and aggregation as they are represented in the model. In total, the model represents 70 nodes (or markets) and 174 power lines (AC and DC).

The model determines a possible path of how installed capacities will develop and how they are operated in the future assuming that electricity markets will achieve the cost-minimizing mix of different technologies which is obtained under perfect competition and the absence of market failures and distortions. Among a number of techno-economic constraints, e.g., supply coverage or investment decisions, the model also includes a number of politically implied constraints: nuclear power is phased-out where decided so, and then only allowed in countries already using it; a CO_2-Quota is implemented corresponding to currently discussed targets for the European energy sector, i.e., 20% reduction with respect to 1990 levels in 2020, and 40% in 2030 (European Commission (2013, 2014)); nation-specific 2020 targets for renewable energy sources are assumed to be reached until 2020 whereas from 2020 onwards there are no further specific renewable energy targets. At the same time, endogenous investments into renewable energy technologies are always possible.

The utilized model for the transmission stage is based on PTDF matrices which are calculated using a detailed European power flow model developed by Energynautics (see Ackermann et al. (2013) for a detailed model description). The number of nodes (70) corresponds to the nodal markets implemented in the generation market model and represents generation and load centers within Europe at an aggregated level. Those nodes are connected by 174 high voltage alternating current (AC) lines (220 and 380kV) as well as high voltage direct current (HVDC) lines. Even though the model is generally built for AC load flow calculations, it is here used to determine PTDF matrices for different grid expansion levels. Details on how the PTDF matrices are calculated can be found in Appendix Appendix B.

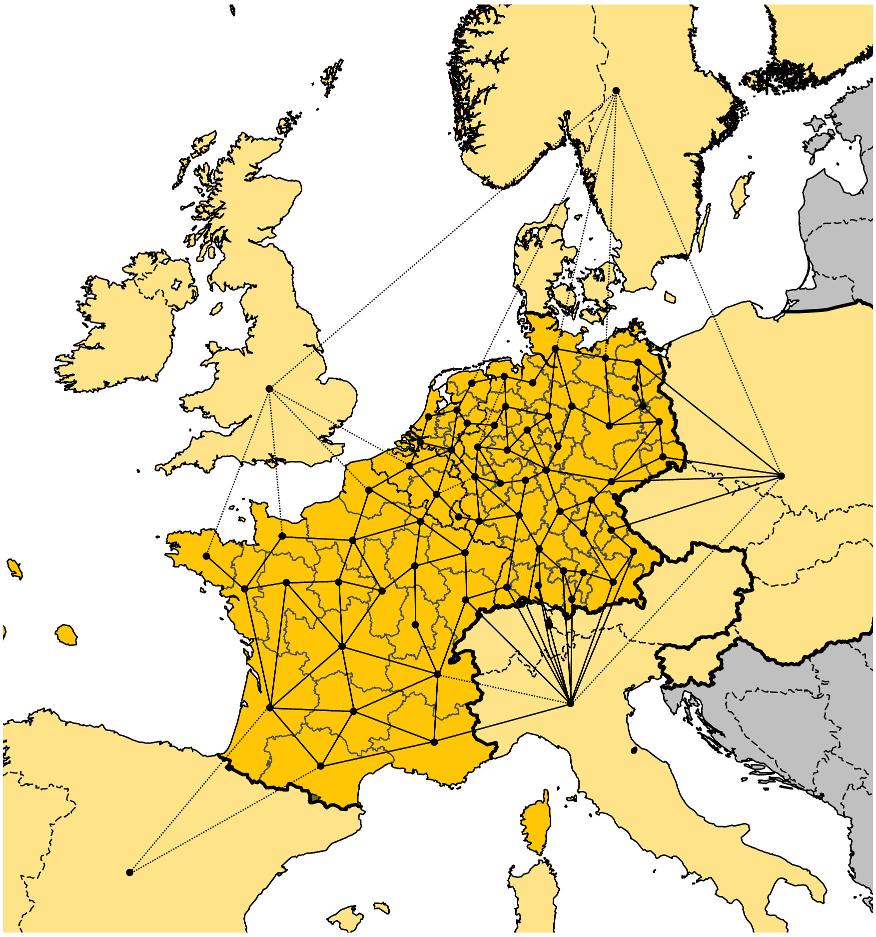

Figure 2: Representation of the CWE and neighboring regions in the model

As a starting point, the optimization takes the situation of the year 2011, based on a detailed database developed at the Institute of Energy Economics at the University of Cologne which in turn is largely based on the Platts WEPP Database (Platts (2009)). From these starting conditions, the development for the years 2020 and 2030 is optimized.[21] As for the temporal resolution, we represent the operational phase by nine typical days representing weekdays and weekend as well as variations in and interdependencies between demand and power from solar and wind. One of the typical days represents an extreme day during the week with peak demand and low supply from wind and solar. Specific numerical assumptions for the generation and transmission model can be found in the Appendix Appendix G.

As in *Settings II-IV* zonal markets are being considered, assumptions about the cross-border price function $g(\kappa_{i,j,t})$ are necessary. For the NTC-based coupling of market zones, we define function $g(\kappa_{i,j,t}) = 1.43 \cdot \frac{\kappa_{i,j,t}\overline{P}_{i,j}}{\sum_{i,j}\overline{P}_{i,j}}$ $\forall i,j \in \mathbf{I_{m,cb}}$ for each market border. The function consists of the weighted average of cross-border line marginals multiplied by a security margin. The security margin is the inverse of the ratio of NTC capacity to technical line capacity and has been derived heuristically by comparing currently installed cross-border grid capacities with NTC values reported by ENTSO-E for the CWE region. For flow-based market coupling, we set this security margin to one, in order to account for enhanced cross-border capacities provided to the power market.[22] In the case of zonal TSOs, we have made the following two assumptions: Differing interest of TSOs regarding cross-border line extensions are aligned by taking the smaller one of the two expansion levels.[23] The costs of cross-border lines are shared half-half by the two TSOs concerned, i.e., $\sigma_{i,j} = 0.5$.

4.2. Results and discussion

As usual in a Benders decomposition, we trace convergence based on the difference between an upper (i.e., the objective value of the integrated problem with solution values of the current iteration) and a lower bound (i.e., the objective of the master problem with the same solution values). We found that all settings reach a convergence threshold of 2.5% or less within 20 to 60 iterations (corresponding to a solution time of 2 to 7 days).[24] For practical reasons, we let all settings solve for one week and – after having double-checked that the convergence threshold of 2.5 % is met – take the last iteration to obtain our final results. The convergence threshold is chosen to keep the solution process computationally tractable, but is also based on empirical observations as well as expected convergence behavior. In fact, a lower convergence criterion increases computational time significantly, while further improvements on the objective value and optimized capacities are hardly observable. For a more thorough discussion of the convergent behavior of our problem, the reader is referred to Appendix Appendix H.

Costs are reported as accumulated discounted system costs.[25] In the generation sector, costs occur due to investments, operation and maintenance, production as well as ramping, whereas in the grid sector,

[21] Technically, we implement the optimization routine up to 2050, but only report results until 2030. This is necessary to avoid problematic results at the end of the optimization timeframe.

[22] Of course, this is just a simple representation of the cross-border capacity allocation. However, a more detailed representation is rather complex and would go beyond the scope of this paper. For more sophisticated models of flow-based capacity allocation, the reader is referred to Kurzidem (2010).

[23] Equation (F.1m) in Appendix Appendix F. Note that this assumption may influence the equilibrium solution of the coordination between the TSOs. Due to the fact that the minimum of the line capacities is chosen, the solutions for the TSOs are no longer continuous. Hence, some equilibria might be omitted during the iterative solution of the problem. We accept this shortfall in our numerical approach for the sake of the large-scale application. The general approach, however, remains valid, and a process for determining all equilibria could be implemented in the numerical solution method (e.g., through randomized starting values).

[24] All models were coded in GAMS 24.2.2 and solved with CPLEX 12.6 on a High Performance Computer with two processors (1600 and 2700Mhz) and physical/virtual memory of 98/150GB.

[25] The discount rate is assumed to be 10% throughout all calculations.

investment as well as operation and maintenance costs are considered. Overall costs of electricity supply can be considered as a measure of efficiency and are reported in the following Figure 3 for the different settings. Besides the absolute costs, which are subdivided into generation and grid costs, the relative cost increase with respect to the overall costs of the nodal pricing setting is also depicted.

Considering the optimality error in the obtained solution, it should be stressed that the exact differences reported here do not necessarily persist after full convergence. However, based on the above discussion about convergence, the general conclusions and order of magnitude are expected to remain valid.

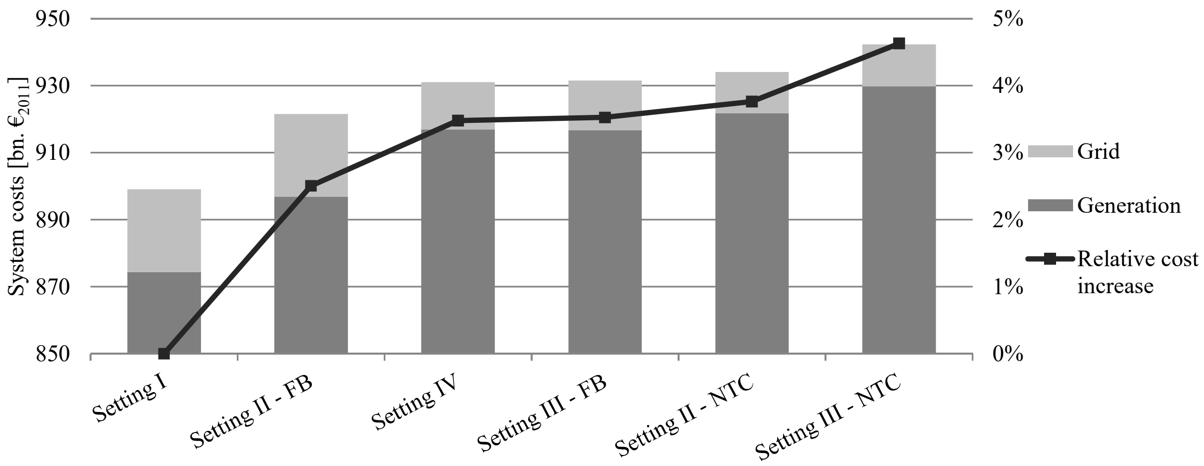

Figure 3: Total costs and relative performance of the different settings

As expected, nodal pricing (*Setting I*) is most efficient, with total costs summing up to 899.0 bn. €$_{2011}$ (874.3 bn. for generation and 24.7 bn. for the grid). Overall, costs increase by up to 4.6% relative to *Setting I* for the other settings. Thereby, NTC-based market coupling induces highest inefficiencies of 3.8% and 4.6% for one single TSO or zonal TSOs, respectively, both with the possibility to do redispatch on a national basis (*Setting II-NTC* and *Setting III-NTC*).[26] Hence, offering few amounts of trading capacity to the generation market, as implied by NTC-based market coupling, induces significant inefficiencies. In fact, by increasing trading capacities via flow-based market coupling, system costs can be lowered and inefficiencies amount to 2.5% for the single TSO, respectively 3.5% for zonal TSOs compared to nodal pricing (*Setting II-FB* and *Setting III-FB*). Hence, efficiency gains of 1.1-1.3 % of total system costs can be achieved by switching from NTC to flow-based market coupling. In turn, enhanced trading activities induced by flow-based market coupling entail greater TSO activity, both in the expansion as well as in the redispatch. For this reason, TSO costs are higher for flow-based than for NTC-based market coupling. However, these additional costs are overcompensated by lower costs in the generation sector. The net effect of a switch from NTC to flow-based market coupling is beneficial for the overall system.

Somewhat surprisingly, the national g-component (*Setting IV*) hardly performs better than the same setting with redispatch (*Setting III-FB*). Hence, the optimal allocation of power generation within market

[26]Since topology control (as, e.g., in Kunz (2013) is not considered, costs of redispatch could possibly be lower. However, since topology control would also be available in the market clearing of the nodal pricing, efficiency gains would persist for all regimes. Hence, the reported differences between the inefficiencies should be similar.

zones is hardly influenced by grid restrictions within that zone. In contrast, the optimal allocation induced by nodal prices throughout the CWE region entails substantial gains in efficiency due to reduced system costs. The setting that comes closest to nodal pricing consists of flow-based coupled zonal markets with a single TSOs and induces an inefficiency of 2.5% in comparison to nodal pricing (*Setting II-FB* vs. *Setting I*).

Even though the share of TSO costs on total costs is very small compared to the share of generation costs in all settings (1.3-2.7%)[27], the amount of grid capacities varies greatly between the different settings. Figure 4 shows the aggregated high voltage (HV) AC and HVDC line capacities.

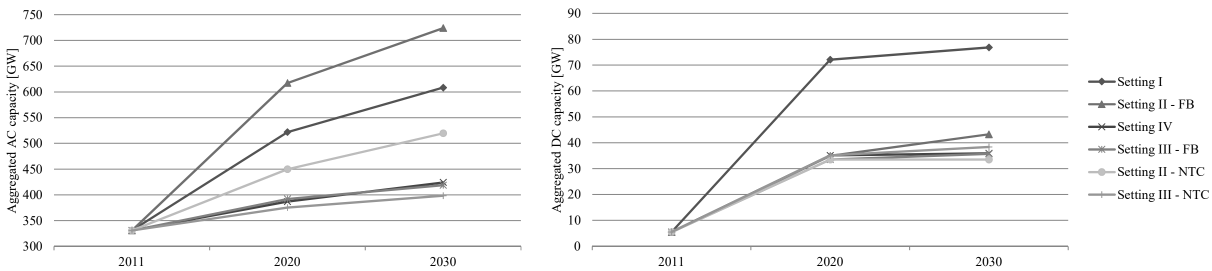

Figure 4: Aggregated line capacities AC and DC

Grid capacities are generally lower in the case of zonal TSOs where they only agree on the smaller of the two proposed expansion levels for cross-border lines (*Setting III-FB* and *Setting III-NTC*). In these cases, overall AC grid capacities increase from 331GW in 2011 to 398GW (*Setting III-NTC*) respectively 418GW (*Setting III-FB*) in 2030, corresponding to an increase of 20-28%. In case of a single TSO, cross-border along with overall line expansions are significantly higher compared to zonal TSOs, with 2030 levels reaching 519GW (*Setting II-NTC*) to 724GW (*Setting II-FB*). Especially in *Setting II-FB*, the TSO is obliged to cope with inefficiently allocated generation plants by excessively expanding the grid, while not being able to avoid those measures with suitable price signals. DC line expansions appear to be crucial for an efficient system development, especially towards the UK where large wind farms help to reach CO_2-targets and to supply the UK itself as well as the continent with comparatively cheap electricity. Thereby, the high DC expansion level in the nodal pricing regime is remarkable. Whereas in zonal markets prices are "averaged" across the zone, nodal prices reveal the true value of connecting specific nodes via DC-lines and thus enable efficient investments in those projects. In consequence, in the nodal pricing regime, DC line capacities are about double as high as in the other settings. This helps to reduce overall costs to a minimum (*Setting I*).

Besides the overall level of grid and generation capacities, their regional allocation also differs between the various settings, mainly due to differences in the (local) availability of transmission upgrades. As has been seen, higher grid expansion levels result from a single TSO (*Setting I and Setting II*), enabling a better utilization of renewable energies at favorable sites (i.e., sites where the specific costs of electricity generation are lowest). In Figure 5, we exemplarily illustrate this effect based on a cross-border line between France and Germany (line 80 in our model). However, the same effect is observable for other interconnections, e.g.,

[27]The rather minor role of grid costs compared to costs occurring in the generation sector has already been identified, e.g., in Fürsch et al. (2013).

between France and Belgium. Higher grid capacities allow the use of high wind speed locations in Northern France and thus foster more expansion of wind capacities in this area. In case of zonal TSOs (*Setting III and Setting IV*) only low amounts of wind capacity are built in France (e.g., in node FR-06) as these areas cannot be connected with the rest of the system. To still meet the European CO_2-target, PV power plants are built in the southern part of Germany (e.g., in node DE-27). Obviously, these locations are non-optimal with respect to other options as they are not used in the setting with one TSO. Thus, implemented market designs significantly influence the amount and location of renewable energies within the system.[28]

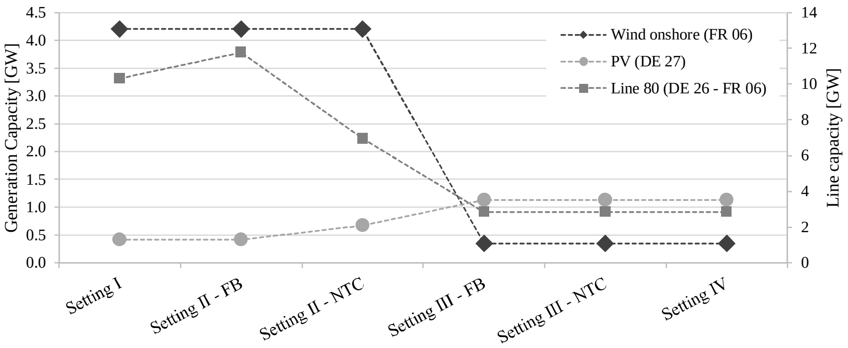

Figure 5: Exemplary grid expansion and regional allocation of renewable energies

5. Conclusions

In the context of liberalized power markets and unbundled generation and transmission services, the purpose of this paper was to develop a modeling framework for different regulatory designs regarding congestion management including both, the operation as well as the investment perspective in the generation and transmission sector. We have presented an analytical formulation that is able to account for different regulatory designs of market areas, a single or zonal TSOs, as well as different forms of measures to relieve congestion, namely grid expansion, redispatch and g-components. We have then proposed an algorithm to numerically solve these problems, based on the concept of decomposition. This technique has shown to entail a number of characteristics that work to our advantage, especially flexible algorithmic implementation as well as consistency of the grid flow representation through PTDF update.

Calibrating our model to the CWE region, we have demonstrated the applicability of our numerical solution algorithm in a large-scale application consisting of 70 nodes and 174 lines along with a detailed bottom-up representation of the generation sector. Compared to nodal pricing as the efficient benchmark, inefficiencies induced by alternative settings reach additional system costs of up to 4.6%. Major deteriorative factors are TSOs activities restricted to zones as well as low trading capacities offered to the market. These findings may serve as a guideline for policymakers when designing international power markets. For instance, our results confirm ongoing efforts to implement flow-based market coupling and to foster a closer cooperation of TSOs in the CWE region. In fact, we find that such a regulatory design could come close

[28] Conventional capacities are also affected. However, the effect is less pronounced as the differences between the site-specific costs of generation are smaller.

to the nodal pricing benchmark, with an efficiency difference of only 2.5%. Reported cost differences might be impacted by numerical imprecision in the solution algorithm, although empirical observations of the convergence behavior suggest that the general effects as well as the order of magnitude persist. Noticeably, the magnitude of these results should be interpreted as the lower bound of efficiency gains, since we focus on frictions in the congestion management only.

More generally, we find that a single TSO (or enhanced coordination between the zonal TSOs) is key for an efficient development of both, grid and generation infrastructures. Whereas the expansion of grid infrastructure is immediately affected, the generation sector indirectly takes advantage of increased grid capacities and hence, can develop more efficiently. Better allocation of generation units with respect to grid costs through high resolution price signals gains importance for larger geographical areas and larger differences between generation costs and expansion potentials (such as wind or solar power). This has been found for the CWE region, and may prove even more important for the whole of Europe. It should be noted, however, that efficiency gains need to be put into the context of transaction costs occurring from the switch to a different congestion management design. In addition, socio-economic factors such as acceptance for grid expansion are not considered in the analysis, but might also play a role considering the large differences of necessary grid quantities.

Limitations of our approach that leave room for extensions and improvement stem from the fact that we assume linear transmission investments, and do not consider strategic behavior of individual agents, imperfectly regulated TSOs, or uncertainty about future developments (e.g., delays in expansion projects). The assumption of an inelastic demand probably reduces the magnitude of the measured inefficiencies, since demand does not react to any price changes and hence only supply-side effects are captured. Algorithmically, the effectiveness of our solution process could be further improved, e.g., through better usage of numerical properties of the problem (such as gradients, etc.). Nevertheless, in its present form, our framework may serve as a valuable tool to assess a number of further relevant questions, such as the tradeoff between different flexibility options (such as grids, storages or renewable curtailment), the impact of different forms of congestion management in other European regions, or the valuation of grid expansion projects.

References

ACER (2014, March). Report on the influence of existing bidding zones on electricity markets.

Ackermann, T., S. Cherevatskiy, T. Brown, R. Eriksson, A. Samadi, M. Ghandhari, L. Söder, D. Lindenberger, C. Jägemann, S. Hagspiel, V. Cuk, P. F. Ribeiro, S. Cobben, H. Bindner, F. R. Isleifsson, and L. Mihet-Popa (2013, May). Smart Modeling of Optimal Integration of High Penetration of PV - Smooth PV. Final Report.

Andersson, G. (2011). *Power System Analysis.* Eidgenössische Technische Hochschule Zürich (ETH).

Baldick, R. and H. Niu (2005). *Electricity Deregulation: Choices and Challenges*, Chapter Lessons Learned: The Texas Experience. University of Chicago Press.

Bazaraa, M. S., H. D. Sherali, and C. M. Shetty (2006). *Nonlinear Programming - Theory and Algorithms.* John Wiley & Sons.

Benders, J. F. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik 4*, 238–252.

Bertsch, J., T. Brown, S. Hagspiel, and L. Just (2015). The relevance of grid expansion under zonal markets. *EWI Working Paper 15/07.*

Boyd, S., L. Xiao, A. Mutapcic, and J. Mattingley (2008). Notes on Decomposition Methods.

Brunekreeft, G., K. Neuhoff, and D. Newbery (2005). Electricity transmission: An overview of the current debate. *Utilities Policy 13*(2), 73–93.

Burstedde, B. (2012). *Essays on the Economic of Congestion Management - Theory and Model-based Analysis for Central Western Europe.* Ph. D. thesis, Universität zu Köln.

Capacity Allocating Service Company (2014, May). Documentation of the CWE FB MC solution - As basis for the formal approval-request.

Caramanis, M. (1982, December). Investment decisions and long-term planning under electricity spot pricing. *IEEE Transactions on Power Apparatus and Systems PAS-101*(12), 4640–4648.

Chao, H.-p., S. Peck, S. Oren, and R. Wilson (2000). Flow-based transmission rights and congestion management. *The Electricity Journal 13*(8), 38–58.

Conejo, A. J., E. Castillo, R. Minguez, and R. Garcia-Bertrand (2006). *Decomposition Techniques in Mathematical Programming - Engineering and Science Applications*. Springer.

Daxhelet, O. and Y. Smeers (2007). The EU regulation on cross-border trade of electricity: A two-stage equilibrium model. *European Journal of Operations Research 181*, 1396–1412.

Ehrenmann, A. and Y. Smeers (2005). Inefficiencies in European congestion management proposals. *Utilities policy 13*(2), 135–152.

European Commission (2013, March). Green Paper - A 2030 framework for climate and energy policies. COM(2013) 169 final.

European Commission (2014, January). Impact Assessment - A 2030 framework for climate and energy policies. SWD(2014) 15 final.

Fürsch, M., S. Hagspiel, C. Jägemann, S. Nagl, D. Lindenberger, and E. Tröster (2013). The role of grid extensions in a cost-effcient transformation of the European electricity system until 2050. *Applied Energy 104*, 642–652.

Glachant, J.-M. (2010). The achievement of the EU electricity internal market through market coupling. *EUI Working Papers, RSCAS 2010/87*.

Green, R. (2007). Nodal pricing of electricity: how much does it cost to get it wrong? *Journal of Regulatory Economics 31*(2), 125–149.

Hagspiel, S., C. Jägemann, D. Lindenberger, T. Brown, S. Cherevatskiy, and E. Tröster (2014). Cost-optimal power system extension under flow-based market coupling. *Energy 66*, 654–666.

Höffler, F. and A. Wambach (2013). Investment coordination in network industries: The case of electricity grid and electricity. *Journal of Regulatory Economics 44*(3), 287–307.

Hogan, W., J. Rosellón, and I. Vogelsang (2010). Toward a combined merchant-regulatory mechanism for electricity transmission expansion. *Journal of Regulatory Economics 38*, 113–143.

Hogan, W. W. (1992). Contract networks for electric power transmission. *Journal of Regulatory Economics 4*(2), 211–242.

Huppmann, D. and J. Egerer (2014). National-strategic investment in European power transmission capacity. *DIW Discussion Papers, No. 1379*.

Jägemann, C., M. Fürsch, S. Hagspiel, and S. Nagl (2013). Decarbonizing Europe's power sector by 2050 - analyzing the implications of alternative decarbonization pathways. *Energy Economics 40*, 622–636.

Joskow, P. and J. Tirole (2005, June). Merchant transmission investment. *The Journal of Industrial Economics LIII*(2), 233–264.

Kunz, F. (2013). Improving congestion management: How to facilitate the integration of renewable generation in germany. *The Energy Journal 34*(4), 55–78.

Kurzidem, M. J. (2010). *Analysis of Flow-based Market Coupling in Oligopolistic Power Markets*. Ph. D. thesis, ETH Zurich.

Leuthold, F., H. Weigt, and C. von Hirschhausen (2008). Efficient pricing for European electricity networks - the theory of nodal pricing applied to feeding-in wind in Germany. *Utilities Policy 16*, 284–291.

Murty, K. G. (1983). *Linear Programming*. John Wiley & Sons.

Neuhoff, K., R. Boyd, T. Grau, J. Barquin, F. Echabarren, J. Bialek, C. Dent, C. von Hirschhausen, B. F. Hobbs, F. Kunz, H. Weigt, C. Nabe, G. Papaefthymiou, and C. Weber (2013). Renewable electric energy integration: Quantifying the value of design of markets for international transmission capacity. *Energy Economics 40*, 760–772.

Oggioni, G., Y. S. E. Allevi, and S. Schaible (2012). A generalized nash equilibrium model of market coupling in the european power system. *Networks & Spatial Economics 12*, 503–560.

Oggioni, G. and Y. Smeers (2012). Degree of coordination in market coupling and counter-trading. *The Energy Journal 33*(3), 39–90.

Oggioni, G. and Y. Smeers (2013). Market failures of market coupling and counter-trading in europe: An illustrative model based discussion. *Energy Economics 35*, 74–87.

Ozdemir, O., F. D. Munoz, J. L. Ho, and B. F. Hobbs (2015). Economic analysis of transmission expansion planning with price-responsive demand and quadratic losses by successive lp. *IEEE Transactions on Power Systems PP*(99), 1–12.

Platts (2009, December). UDI World Electric Power Plants Data Base (WEPP).

Richter, J. (2011). DIMENSION - a dispatch and investment model for European electricity markets. EWI WP 11/3.

Rious, V., P. Dessante, and Y. Perez (2009). Is combination of nodal pricing and average participation tariff the best solution to coordinate the location of power plants with lumpy transmission investment? *EUI Working Papers, RSCAS 2009/14*.

Sauma, E. E. and S. S. Oren (2006). Proactive planning and valuation of transmission investment in restructured electricity markets. *Journal of Regulatory Economics 30*, 261–290.

Schweppe, F. C., M. C. Caramanis, R. D. Tabors, and R. E. Bohn (1988). *Spot Pricing of Electricity*. Norwell, MA: Kluwer.

van der Weijde, A. H. and B. F. Hobbs (2011). Locational-based coupling of electricity markets: benefits from coordinating unit commitment and balancing markets. *Journal of Regulatory Economics 39*, 223–251.

Appendix A. Notation list

Abbreviation	Dimension	Description
Model sets		
$i \in \mathbf{I}, j \in \mathbf{J}, k \in \mathbf{K}, l \in \mathbf{L}$		Nodes, $\mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L} = [1, 2, ...]$
$m, n \in \mathbf{M}$		Zonal markets, $\mathbf{M} = [1, 2, ...]$
$i \in \mathbf{I_m}, j \in \mathbf{J_m}$		Nodes that belong to zonal market m, $\mathbf{I_m} \subset \mathbf{I}$, $\mathbf{J_m} \subset \mathbf{J}$
$i \in \mathbf{I_{m,cb}}, j \in \mathbf{J_{m,cb}}$		Nodes that belong to zonal market m and are connected to a another zone n by a cross-border line, $\mathbf{I_{m,cb}} \subset \mathbf{I_m}$, $\mathbf{J_{m,cb}} \subset \mathbf{J_m}$
$t \in \mathbf{T}$		Point in time for dispatch decisions (e.g., hours)
Model parameters		
δ_i	EUR/kW	Investment and FOM costs of generation capacity in node i
$\gamma_{i,t}$	EUR/kWh	Variable costs of generation capacity in node i
$\mu_{i,j}$	EUR/kW	Investment costs of line between node i and node j
$d_{i,t}$	kW	Electricity demand in node i
$PTDF_{k,i,j}$	$-$	Power Transfer Distribution Factor (impact of the power balance in node k on flows on line i, j)
$\sigma_{i,j}$	%	Cost share for an interconnector capacity between node i and node j, $i \in \mathbf{I_{m,cb}}$, $j \in \mathbf{J_{m,cb}}$
Model primal variables		
$\overline{G}_i$	kW	Generation capacity in node i, $\overline{G}_i \geq 0$
$G_{i,t}$	kW	Generation dispatch in node i, $G_{i,t} \geq 0$
$T_{i,j,t}, T_{m,n,t}$	kW	Electricity trade from node i to node j, or market m to market n
X	EUR	Costs of generation
Y	EUR	Costs of TSO
$\overline{P}_{i,j}$	kW	Line capacity between node i and node j, $\overline{P}_{i,j} \geq 0 \;\; \forall\, i, j \neq j, i$
$P_{i,j,t}$	kW	Electricity flow on line between node i and node j
$R_{i,t}$	kW	Redispatch in node i
α	EUR	Helping variable to include transmission costs of the current iteration in the master problem
Model dual variables		
$\kappa_{i,j,t}, \kappa_{m,n,t}$	EUR/kW	price for transmission between nodes (i and j) or zones (m and n)
$\lambda_{i,t}, \lambda_{m,t}$	EUR/kW	nodal or zonal price for electricity

Table A.2: Model sets, parameters and variables

Appendix B. Derivation of the load flow equations by means of PTDFs

Power Transfer Distribution Factors (PTDFs) are a well-established method to account for load flows in meshed electricity networks by means of linearization. They can be derived from the network equations in an AC power network that write as follows:[29]

$$P_i = U_i \sum_{j\in\mathcal{I}} U_j(g_{i,j}\cos(\varphi_i-\varphi_j)+b_{i,j}\sin(\varphi_i-\varphi_j)) \tag{B.1}$$

$$Q_i = U_i \sum_{j\in\mathcal{I}} U_j(g_{i,j}\sin(\varphi_i-\varphi_j)-b_{i,j}\cos(\varphi_i-\varphi_j)) \tag{B.2}$$

$$P_{i,j} = U_i^2 g_{i,j} - U_iU_jg_{i,j}\cos(\varphi_i-\varphi_j) - U_iU_jb_{i,j}\sin(\varphi_i-\varphi_j) \tag{B.3}$$

$$Q_{i,j} = -U_i^2(b_{i,j}+b_{i,j}^{sh}) + U_iU_jb_{i,j}\cos(\varphi_i-\varphi_j) - U_iU_jg_{i,j}\sin(\varphi_i-\varphi_j). \tag{B.4}$$

P_i and Q_i represent the net active and reactive power infeed (i.e., nodal power balances), and $P_{i,j}$ and $Q_{i,j}$ the active and reactive power flows between node i and j. Voltage levels U and phase angles φ of the nodes as well as series conductances g and series susceptances b of the transmission lines determine active and reactive power flows in a highly nonlinear way.

In order to linearize the above equations, a number of assumptions are made:

- All voltages are set to 1 p.u.
- Voltage angles are all similar (and hence, $\sin(\varphi_i-\varphi_j) \approx \varphi_i-\varphi_j$).
- Reactive power is neglected (i.e., $Q_i = Q_{i,j} = 0$).
- Losses are neglected and line reactances are much larger than their resistance, such that $x \gg r \approx 0$.

Under these assumptions and using Kirchoff's power law, the network equations can be simplified to

$$P_{i,j} \approx \frac{1}{x_{i,j}}(\varphi_i-\varphi_j) \tag{B.5}$$

$$P_i \approx \sum_{j\in\Omega_i}\frac{1}{x_{i,j}}(\varphi_i-\varphi_j), \tag{B.6}$$

with Ω_i representing the nodes adjacent to i. If there are multiple nodes and branches, this can be written in a more convenient matrix notation as $\tilde{\boldsymbol{P_i}} = \tilde{\boldsymbol{B}} \cdot \tilde{\boldsymbol{\Theta}}$, with $\tilde{\boldsymbol{P_i}}$ being the vector of net active nodal power balances P_i, $\tilde{\boldsymbol{\Theta}}$ the vector of phase angles, and $\tilde{\boldsymbol{B}}$ the nodal admittance matrix with the following entries:

$$\tilde{B}_{i,j} = -\frac{1}{x_{i,j}} \tag{B.7}$$

$$\tilde{B}_{i,i} = \sum_{j\in\Omega_i}\frac{1}{x_{i,j}}. \tag{B.8}$$

By deleting the row and column belonging to the reference node (thus assuming a zero reference angle at this node), the previously singular matrix $\tilde{\boldsymbol{B}}$ becomes $\boldsymbol{B}$, the vector of phase angles $\boldsymbol{\Theta}$, and the vector of net active nodal power balances $\boldsymbol{P_i}$. We can now solve for $\boldsymbol{\Theta}$ by matrix inversion:

$$\boldsymbol{\Theta} = \boldsymbol{B}^{-1} \cdot \boldsymbol{P_i}. \tag{B.9}$$

[29] The following is based on Andersson (2011), even though the general approach can be found in most electrical engineering textbooks.

Defining $H_{ki} = 1/x_{i,j}$, $H_{kj} = -1/x_{i,j}$ and $H_{km} = 0$ for $m \neq i,j$ (with k running over the branches i,j), Equation (B.5) can be rewritten in matrix form as $\boldsymbol{P_{i,j}} = \boldsymbol{H} \cdot \boldsymbol{\Theta}$. Inserting $\boldsymbol{\Theta}$ from Equation (B.9) finally yields

$$\boldsymbol{P_{i,j}} = \boldsymbol{H} \cdot \boldsymbol{\Theta} = \boldsymbol{H} \cdot \boldsymbol{B}^{-1} \cdot \boldsymbol{P_i} = PTDF \cdot \boldsymbol{P_i} \tag{B.10}$$

The elements of $PTDF$ are the power transfer distribution factors that constitute a linear relationship between nodal power balances and load flows. Note that the size of the $PTDF$ matrix is determined by the size of the system, with the number of matrix lines corresponding to the number of transmission lines, and the number of matrix columns representing the number of nodes. The matrix entry $PTDF_{k,i,j}$ represents the impact of the power balance in node k on power flows on line between node i and j. Also note that $PTDF$ essentially depends (only) on the line impedances $x_{i,j}$ in the system that in turn depend primarily on the respective line capacities $\overline{P}_{i,j}$. Hence, as done, e.g., in Hogan et al. (2010), we apply the law of parallel circuits to adjust line reactances when altering transmission capacities, i.e.,

$$x_{i,j} = \frac{\overline{P}^0_{i,j}}{\overline{P}_{i,j}} x^0_{i,j}, \tag{B.11}$$

where $\{\overline{P}^0_{i,j}, x^0_{i,j}\}$ is a point of reference taken from the original configuration of the transmission network. Overall, this yields a functional dependency of power flows on nodal balances (determined by generation G_k and load d_k in all nodes) as well as line capacities $\overline{P}_{k,l}$ of all lines in the system, i.e., $P_{i,j} = P_{i,j}(\overline{P}_{k,l}, G_k, d_k)$.

Appendix C. Equivalence of Problem P1 and P1'

To show the equivalence of the optimal solution of *P1* and *P1'*, we compare the problems by means of their Karush-Kuhn-Tucker (KKT) conditions. If they are equal, the optimal solution has to be equal, too (e.g., Bazaraa et al. (2006)). For the derivations, note that trade is a function of line capacity, generation and demand, i.e., $T_{i,j,t} = T_{i,j,t}(\overline{P}_{k,l}, G_{k,t}, d_{k,t})$, and that $T_{i,j,t} = -T_{j,i,t}$. The following is the Lagrangian function belonging to Problem *P1*:

$$\begin{aligned} L(\overline{G}_i, G_{i,t}, T_{i,j,t}, \overline{P}_{i,j}, \lambda_{i,t}, \tau_{i,t}, \kappa_{i,j,t}) = & \sum_i \delta_i \overline{G}_i + \sum_{i,t} \gamma_{i,t} G_{i,t} + \sum_{i,j} \mu_{i,j} \overline{P}_{i,j} \\ & + \sum_{i,t} (\lambda_{i,t}(G_{i,t} - \sum_j T_{i,j,t} - d_{i,t}) + \tau_{i,t}(G_{i,t} - \overline{G}_i)) \\ & + \sum_{i,j,t} (\kappa_{i,j,t}(|T_{i,j,t}| - \overline{P}_{i,j})) \end{aligned} \tag{C.1}$$

The corresponding KKT conditions are:

$$\frac{\partial L}{\partial \overline{G}_i} = \delta_i - \sum_t \tau_{i,t} \le 0,\ \ \overline{G}_i \ge 0,\ \ \overline{G}_i(\frac{\partial L}{\partial \overline{G}_i}) = 0 \quad \forall i \tag{C.2a}$$

$$\frac{\partial L}{\partial G_{i,t}} = \gamma_{i,t} + \lambda_{i,t}(1 - \sum_j \frac{\partial T_{i,j,t}}{\partial G_{i,t}}) + \tau_{i,t} + \sum_j \kappa_{i,j,t}\frac{\partial T_{i,j,t}}{\partial G_{i,t}} \le 0,\ G_{i,t} \ge 0,\ G_{i,t}(\frac{\partial L}{\partial G_{i,t}}) = 0\ \ \forall i, t \tag{C.2b}$$

$$\frac{\partial L}{\partial \overline{P}_{i,j}} = \mu_{i,j} - \sum_t \lambda_{i,t}\frac{\partial T_{i,j,t}}{\partial \overline{P}_{i,j}} + \sum_t \kappa_{i,j,t}(\frac{\partial T_{i,j,t}}{\partial \overline{P}_{i,j}} - 1) \le 0,\ \ \overline{P}_{i,j} \ge 0,\ \ \overline{P}_{i,j}(\frac{\partial L}{\partial \overline{P}_{i,j}}) = 0 \quad \forall i, j \tag{C.2c}$$

$$\frac{\partial L}{\partial \kappa_{i,j,t}} = |T_{i,j,t}| - \overline{P}_{i,j} \le 0, \quad \kappa_{i,j,t} \ge 0, \quad \kappa_{i,j,t}(\frac{\partial L}{\partial \kappa_{i,j,t}}) = 0 \qquad \forall i, j, t \tag{C.2d}$$

$$\frac{\partial L}{\partial \tau_{i,j}} = G_{i,t} - \overline{G}_i \le 0, \quad \tau_{i,j} \ge 0, \quad \tau_{i,j}(\frac{\partial L}{\partial \tau_{i,j}}) = 0 \qquad \forall i, j \tag{C.2e}$$

$$\frac{\partial L}{\partial \lambda_{i,t}} = G_{i,t} - \sum_j T_{i,j,t} - d_{i,t} = 0 \qquad \forall i, t \tag{C.2f}$$

$$\frac{\partial L}{\partial T_{i,j,t}} = \kappa_{i,j,t} - \lambda_{i,t} + \lambda_{j,t} = 0 \qquad \forall i, j, t \tag{C.2g}$$

The Langragian functions for *P1'* are:

$$\begin{aligned} L'^a(\overline{G}_i, G_{i,t}, T_{i,j,t}, \lambda_{i,t}, \tau_{i,t}) = &\sum_i \delta_i \overline{G}_i + \sum_{i,t} \gamma_{i,t} G_{i,t} + \sum_{i,j,t} \kappa_{i,j,t} T_{i,j,t} \\ &+ \sum_{i,t} (\lambda_{i,t}(G_{i,t} - \sum_{j,t} T_{i,j,t} - d_{i,t}) + \kappa_{i,j,t}(G_{i,t} - \overline{G}_i)) \end{aligned} \tag{C.3}$$

$$L'^b(\overline{P}_{i,j}, \kappa_{i,j,t}) = \sum_{i,j} \mu_{i,j}\overline{P}_{i,j} + \sum_{i,t}(\lambda_{i,t}(G_{i,t} - \sum_{j,t} T_{i,j,t} - d_{i,t})) + \sum_{i,j,t}(\kappa_{i,j,t}(|T_{i,j,t}| - \overline{P}_{i,j})) \tag{C.4}$$

The KKT conditions of *P1'a* are:

$$\frac{\partial L'^a}{\partial \overline{G}_i} = \delta_i - \sum_t \tau_{i,t} \le 0,\ \ \overline{G}_i \ge 0,\ \ \overline{G}_i(\frac{\partial L}{\partial \overline{G}_i}) = 0 \quad \forall i \tag{C.5a}$$

$$\frac{\partial L'^a}{\partial G_{i,t}} = \gamma_{i,t} + \lambda_{i,t}(1 - \sum_j \frac{\partial T_{i,j,t}}{\partial G_{i,t}}) + \tau_{i,t} + \sum_j \kappa_{i,j,t}\frac{\partial T_{i,j,t}}{\partial G_{i,t}} \le 0,\ G_{i,t} \ge 0,\ G_{i,t}(\frac{\partial L}{\partial G_{i,t}}) = 0\ \ \forall i, t \tag{C.5b}$$

$$\frac{\partial L'^a}{\partial \tau_{i,j}} = G_{i,t} - \overline{G}_i \le 0, \quad \tau_{i,j} \ge 0, \quad \tau_{i,j}(\frac{\partial L'^a}{\partial \tau_{i,j}}) = 0 \qquad \forall i, j \tag{C.5c}$$

$$\frac{\partial L'^a}{\partial \lambda_{i,t}} = G_{i,t} - \sum_j T_{i,j,t} - d_{i,t} \qquad \forall i, t \tag{C.5d}$$

$$\frac{\partial L'^a}{\partial T_{i,j,t}} = \kappa_{i,j,t} - \lambda_{i,t} + \lambda_{j,t} = 0 \qquad \forall i, j, t \tag{C.5e}$$

The KKT conditions of *P1'b* are:

$$\frac{\partial L'^b}{\partial \overline{P}_{i,j}} = \mu_{i,j} - \sum_t \lambda_{i,t}\frac{\partial T_{i,j,t}}{\partial \overline{P}_{i,j}} + \sum_t \kappa_{i,j,t}(\frac{\partial T_{i,j,t}}{\partial \overline{P}_{i,j}} - 1) \le 0,\ \ \overline{P}_{i,j} \ge 0,\ \ \overline{P}_{i,j}(\frac{\partial L}{\partial \overline{P}_{i,j}}) = 0 \quad \forall i, j \tag{C.6a}$$

$$\frac{\partial L'^b}{\partial \kappa_{i,j,t}} = |T_{i,j,t}| - \overline{P}_{i,j} \le 0, \quad \kappa_{i,j,t} \ge 0, \quad \kappa_{i,j,t}(\frac{\partial L}{\partial \kappa_{i,j,t}}) = 0 \qquad \forall i, j, t \tag{C.6b}$$

Comparing the KKT conditions of problem *P1* to the ones of *P1a* and *P1b*, we can conclude that the problems are indeed equivalent.

Appendix D. Model of Setting III: coupled zonal markets with zonal TSOs and zonal redispatch

Mathematically, the model of *Setting III*, representing coupled zonal markets with zonal TSOs and zonal redispatch, is formulated as follows:

P3a *Generation*

$$\min_{\overline{G}_i, G_{i,t}, T_{m,n,t}} \quad X = \sum_i \delta_i \overline{G}_i + \sum_{i,t} \gamma_{i,t} G_{i,t} + \sum_{m,n,t} \kappa_{m,n,t} T_{m,n,t} \tag{D.1a}$$

$$\text{s.t.} \quad \sum_{i \in \mathbf{I_m}} G_{i,t} - \sum_{n,t} T_{m,n,t} = \sum_{i \in \mathbf{I_m}} d_{i,t} \qquad \forall m,t \qquad | \, \lambda_m \tag{D.1b}$$

$$G_{i,t} \leq \overline{G}_i \qquad \forall i,t \tag{D.1c}$$

$$T_{m,n,t} = -T_{n,m,t} \qquad \forall m,n,t \tag{D.1d}$$

P3b *Transmission*

$$\min_{\overline{P}_{i,j \in \mathbf{I_m}}, R_{i \in \mathbf{I_m},t}} \quad Y_m = \sum_{i,j \in \mathbf{I_m}} \mu_{i,j} \overline{P}_{i,j} + \sum_{i,j \in \mathbf{I_{m,cb}}} \sigma_{i,j} \mu_{i,j} \overline{P}_{i,j} + \sum_{i \in \mathbf{I_m},t} \gamma_{i,t} R_{i,t} \qquad \forall m \tag{D.1e}$$

$$\text{s.t.} \sum_{i \in \mathbf{I_m}} G_{i,t} - \sum_{n} T_{m,n,t} = \sum_{i \in \mathbf{I_m}} d_{i,t} \qquad \forall m,t \tag{D.1f}$$

$$|T_{i,j,t}| = |P_{i,j,t}(\overline{P}_{k,l}, R_{k,t}, G_{k,t}, d_{k,t})| \leq \overline{P}_{i,j} \qquad \forall t, i, j \in \mathbf{I_m} \qquad | \, \kappa_{i,j \in \mathbf{I_m}} \tag{D.1g}$$

$$\sum_{i \in \mathbf{I_m},t} R_{i,t} = 0 \tag{D.1h}$$

$$0 \leq G_{i,t} + R_{i,t} \leq \overline{G}_i \qquad \forall t, i \in \mathbf{I_m} \tag{D.1i}$$

$$\kappa_{m,n,t} = g(\kappa_{i,j,t}) \tag{D.1j}$$

$$T_{m,n,t} = -T_{n,m,t} \qquad \forall m,n,t \tag{D.1k}$$

In problem *P3*, there are now separate optimization problems for each zonal TSO (indicated by Y_m), with the objective to minimize costs from zonal grid and cross-border capacity extensions as well as from zonal redispatch measures (Equation (D.1e)). For the redispatch, TSOs have to consider the same restrictions as in the previous setting (Equations (D.1h) and (D.1i)). TSOs are assumed to negotiate about the extension of cross-border capacities according to some regulatory rule that ensures the acceptance of a unique price for each cross-border line by both of the neighboring TSOs. For instance, the regulatory rule may be specified such that both TSOs are obliged to accept the higher price offer, or, equivalently, the lower of the two capacities offered for the specific cross-border line. Corresponding costs from inter-zonal grid extensions are assumed to be shared among the TSOs according to the cost allocation key $\sigma_{i,j}$. According to Equation (D.1j), prices for transmission between zones that are provided to the generation stage ($\kappa_{m,n,t}$) are determined just as in the previous *Setting II* with only one TSO, depending on the type of market coupling, i.e., the specification of function g. The only difference is that line-specific prices $\kappa_{i,j,t}$ may now deviate from *Setting II* as they result from the separated activities of each zonal TSO (specifically, from Equation (D.1g), i.e., the restriction of flows on intra-zonal and cross-border lines).

Appendix E. Model of Setting IV: coupled zonal markets with zonal TSOs and g-component

Mathematically, the model of *Setting IV*, representing coupled zonal markets with zonal TSOs and g-component, is formulated as follows:

P4a *Generation*

$$\min_{\overline{G}_i, G_{i,t}, T_{m,n,t}} \quad X = \sum_i \delta_i \overline{G}_i + \sum_{i,t} \gamma_{i,t} G_{i,t} + \sum_{i,j,t} \kappa_{i,j,t} T_{i,j,t} \tag{E.1a}$$

$$\text{s.t.} \quad \sum_{i \in \mathbf{I_m}} G_{i,t} - \sum_n T_{m,n,t} = \sum_{i \in \mathbf{I_m}} d_{i,t} \qquad \forall m, t \qquad | \, \lambda_m \tag{E.1b}$$

$$G_{i,t} \leq \overline{G}_i \qquad \forall i, t \tag{E.1c}$$

$$T_{m,n,t} = -T_{n,m,t} \qquad \forall m, n, t \tag{E.1d}$$

P4b *Transmission*

$$\min_{\overline{P}_{i,j \in \mathbf{I_m}, \mathbf{I_{m,cb}}}} \quad Y_m = \sum_{i,j \in \mathbf{I_m}} \mu_{i,j} \overline{P}_{i,j} + \sum_{i,j \in \mathbf{I_{m,cb}}} \sigma_{i,j} \mu_{i,j} \overline{P}_{i,j} \qquad \forall m \tag{E.1e}$$

$$\text{s.t.} \sum_{i \in \mathbf{I_m}} G_{i,t} - \sum_n T_{m,n,t} = \sum_{i \in \mathbf{I_m}} d_{i,t} \qquad \forall m, t \tag{E.1f}$$

$$|T_{i,j,t}| = |P_{i,j,t}(\overline{P}_{k,l}, G_{k,t}, d_{k,t})| \leq \overline{P}_{i,j} \quad \forall t, i, j \in \mathbf{I_m}, \mathbf{I_{m,cb}} \quad | \, \kappa_{i,j \in \mathbf{I_m}, \mathbf{I_{m,cb}}, t} \tag{E.1g}$$

$$T_{m,n,t} = -T_{n,m,t} \qquad \forall m, n, t \tag{E.1h}$$

Problem *P4a* is almost identical to *P2a* (and *P3a*), with the exception of one term in the objective function (E.1a). With a g-component, generators pay nodal instead of zonal prices for transmission ($\kappa_{i,j,t}$ instead of $\kappa_{m,n,t}$), depending on the impact of their nodal generation level on the grid infrastructure (by means of $T_{i,j,t} = T_{i,j,t}(G_{k,t}, d_{k,t})$). These prices are determined by the zonal TSOs via their flow-restriction (E.1g).

Appendix F. Numerical algorithm for NTC-coupled zonal markets, zonal TSOs, and zonal redispatch

In Section 3, we have shown the numerical implementation of the nodal pricing regime. For the sake of clarifying the major changes needed to represent the alternative Settings *II-IV*, we here present the model for m zonal (instead of nodal) markets that are coupled via NTC-based capacity restrictions, along with multiple zonal TSOs (instead of only one), all having the possibility to deploy zonal redispatch as an alternative to grid expansion. Hence, the model corresponds to *Setting III* with NTC-based market coupling. Compared to nodal pricing, no more nodal or time-specific information about grid costs is provided. Instead, an aggregated price $\kappa^{(v)}_{m,n,t}$ for each border is calculated via a function g_{NTC} and passed on to generation level. The model with flow-based market coupling works in the same way, only that the price $\kappa^{(v)}_{m,n,t}$ is calculated via a different function g_{FB}.

$v = 1$; convergence=false

While(convergence=false) {

Master problem: generation

$$\min_{\overline{G}_i, G_{i,t}, T_{m,n,t}, \alpha} \quad X = \sum_i \delta_i \overline{G}_i + \sum_{i,t} \gamma_{i,t} G_{i,t} + \alpha \tag{F.1a}$$

$$\text{s.t.} \quad \sum_{i \in \mathbf{I_m}} G_{i,t} - \sum_n T_{m,n,t} = \sum_{i \in \mathbf{I_m}} d_{i,t} \qquad \forall m, t \tag{F.1b}$$

$$G_{i,t} \leq \overline{G}_i \qquad \forall i, t \tag{F.1c}$$

$$T_{m,n,t} = -T_{n,m,t} \qquad \forall m, n, t \tag{F.1d}$$

$$\sum_m Y_m^{(u)} + \sum_{m,n,t} \kappa_{m,n,t}^{(u)} \cdot (T_{m,n,t} - T_{m,n,t}^{(u)}) \leq \alpha \qquad \forall u = 1, ..., v-1 | v > 1 \tag{F.1e}$$

− −

$$G_{i,t}^{(v)} = \text{Optimal value of } G_{i,t} \qquad \forall i, t \tag{F.1f}$$

Sub-problem: transmission

$$\min_{\overline{P}_{i,j \in \mathbf{I_m}, \mathbf{I_{m,cb}}}, R_{i \in \mathbf{I_m}, t}, T_{m,n,t}} \quad Y_m = \sum_{i,j \in \mathbf{I_m}} \mu_{i,j} \overline{P}_{i,j} + \frac{1}{2} \sum_{i,j \in \mathbf{I_{m,cb}}} \mu_{i,j} \overline{P}_{i,j} + \sum_{i \in \mathbf{I_m}, t} R_{i,t} \gamma_{i,t} \ \forall m \tag{F.1g}$$

$$\text{s.t.} \quad |P_{i,j,t}| = \left| \sum_k PTDF_{k,i,j} \cdot (G_{k,t}^{(v)} + R_{k,t} - d_{k,t}) \right| \leq \overline{P}_{i,j} \qquad \forall i, j, t \qquad | \, \kappa_{i,j,t}^{(v)} \tag{F.1h}$$

$$0 \leq R_{i,t} + G_{i,t}^{(v)} \leq \overline{G}_i \qquad \forall i, t \in \mathbf{I_m} \tag{F.1i}$$

$$\sum_{i \in \mathbf{I_m}} R_{i,t} = 0 \tag{F.1j}$$

− −

$$Y_m^* = \text{Optimal value of } Y_m \tag{F.1k}$$

$$PTDF^{(v)} = \text{New PTDF matrix calculated based on } \overline{P}_{i,j} \kappa_{m,n,t} = g_{NTC}(\kappa_{i,j,t}^{(v)}) \tag{F.1l}$$

$$\overline{P}_{i,j \in \mathbf{I_{m,cb}}} = \overline{P}_{i,j \in \mathbf{I_{n,cb}}} = \min \left\{ \overline{P}_{i,j \in \mathbf{I_{m,cb}}} ; \overline{P}_{i,j \in \mathbf{I_{n,cb}}} \right\} \tag{F.1m}$$

if(convergence criterion < threshold; convergence=true)
$v = v + 1$
};

Appendix G. Numerical assumptions for the large-scale application

To depict the CWE region in a high spatial resolution, we split the gross electricity demand per country among the nodes belonging to this country according to the percentage of population living in that region.

Appendix H. Convergence analysis

To illustrate the convergent behavior of our problem, Figure H.6, left hand side, shows the development of the optimality error (relative difference between the upper and lower bound of the optimization), along with the (absolute) rate of change of the lower bound obtained during the iterative solution of the nodal pricing setting. The lower bound is observed to change only slightly, reaching change rates smaller than 0.01% after some 40 iterations. Moreover, as can be derived from the interpolation curves presented in Figure H.6, left

Country	2011	2020	2030
Belgium	87	98	105
Germany	573	612	629
France	466	524	559
Luxembourg	7	8	8
Netherlands	113	128	137
Eastern	276	328	366
Northern	387	436	465
Southern	450	528	594
Southwest	317	378	433
United Kingdom	400	450	481

Table G.3: Assumptions for the gross electricity demand [TWh]

Technology	2020	2030
Wind Onshore	1253	1188
Wind Offshore (<20m depth)	2800	2350
Wind Offshore (>20m depth)	3080	2585
Photovoltaics (roof)	1260	935
Photovoltaics (ground)	1110	785
Biomass gas	2398	2395
Biomass solid	3297	3295
Biomass gas, CHP	2597	2595
Biomass solid, CHP	3497	3493
Geothermal	10504	9500
Compressed Air Storage	1100	1100
Pump Storage	1200	1200
Lignite	1500	1500
Lignite Innovative	1600	1600
Coal	1200	1200
Coal Innovative	2025	1800
IGCC	1700	1700
CCGT	711	711
OCGT	400	400
Nuclear	3157	3157

Table G.4: Assumptions for the generation technology investment costs [€/kW]

Fuel type	2011	2020	2030
Nuclear	3.6	3.3	3.3
Lignite	1.4	1.4	2.7
Oil	39.0	47.6	58.0
Coal	9.6	10.1	10.9
Gas	14.0	23.1	25.9

Table G.5: Assumptions for the gross fuel prices [€/MWh$_{th}$]

hand side, the relative error decreases at much faster rates with a ratio of approximately 200 for an estimated exponential trend and an iteration count of 60. Based on the fact that in a Benders decomposition the lower bound is non-decreasing (i.e., change rates are always positive as demonstrated in Figure H.6, left hand side), and the empirically observed behavior of the lower bound, it can be concluded that the error further

Grid Technology	Extension costs	FOM costs
AC overhead line incl. compensation	445 €/(MVA*km)	2.2 €/(MVA*km)
DC overhead line	400 €/(MW*km)	2.0 €/(MW*km)
DC underground	1250 €/(MW*km)	6.3 €/(MW*km)
DC submarine	1100 €/(MW*km)	5.5 €/(MW*km)
DC converter pair	150000 €/MW	750.0 €/MW

Table G.6: Assumptions for the grid extension and FOM costs

decreases mainly due to changes in the upper bound. Hence, we argue that the lower bound can be taken as a good approximation of the optimal objective value as soon as our convergence criterion is met. To support this argument and to deepen our insights, we closely analyzed optimized levels of the variables, observing that they reach fairly stable levels in the last iterations before reaching the convergence criterion.[30] As an example, the right hand side of Figure H.6 shows aggregated AC line capacities obtained in the final runs of the nodal pricing setting.

Based on the interpolation curves estimated from the observed changes in the optimality error, a 1% threshold is expected to be reached after around 150 iterations. The estimated increase of the lower bound and hence, the improvement of the optimal solution, will then be around 0.21% higher compared to our obtained value. At around 300 iterations, the optimal solution will deviate by about 0.24% from our obtained value, and further improvements of the optimal solution would be negligible. Considering the extensive computational burden as well as the expected limited improvements, we do not consider a smaller convergence threshold and rather accept some level of uncertainty regarding the different levels of optimality achieved in the different settings.

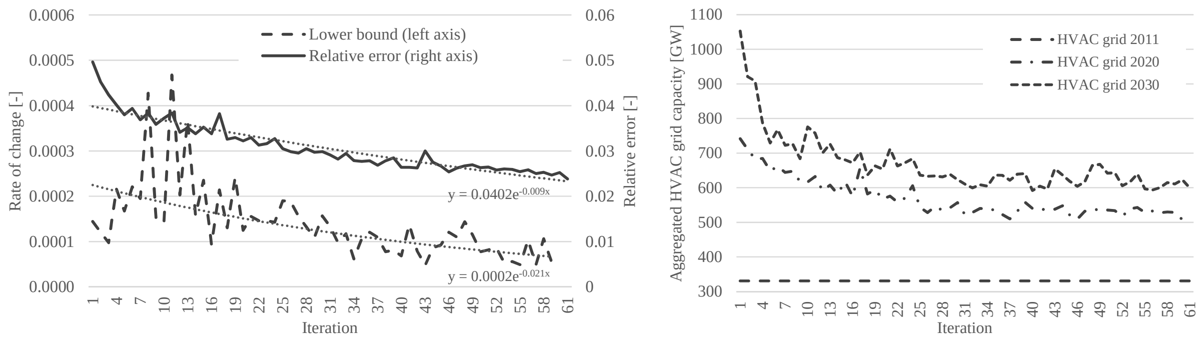

Figure H.6: Development of lower bound, optimality error and aggregated AC line capacities during the iteration in *Setting I*

[30]Note that this argument is also supported by the analysis of convergence in a very similar setting published in Hagspiel et al. (2014).