Felling, Tim; Weber, Christoph

Working Paper

Consistent and robust delimitation of price zones under uncertainty with an application to Central Western Europe

HEMF Working Paper, No. 06/2017

Provided in Cooperation with:
University of Duisburg-Essen, Chair for Management Science and Energy Economics

Suggested Citation: Felling, Tim; Weber, Christoph (2017) : Consistent and robust delimitation of price zones under uncertainty with an application to Central Western Europe, HEMF Working Paper, No. 06/2017, University of Duisburg-Essen, House of Energy Markets & Finance, Essen

This Version is available at:
https://hdl.handle.net/10419/172440

House of Energy Markets & Finance

Consistent and robust delimitation of price zones under uncertainty with an application to Central Western Europe

HEMF Working Paper No. 06/2017

by

Tim Felling

and

Christoph Weber

June 2017

UNIVERSITÄT DUISBURG ESSEN

Open-Minded

Consistent and robust delimitation of price zones under uncertainty with an application to Central Western Europe by Tim Felling and Christoph Weber

Abstract

New and alternative delimitations of price zones for Central Western Europe (CWE) might constitute a mid-term solution to cope with the increasing congestion in the electricity transmission grids. The significantly growing infeed from renewable energy sources puts more and more pressure on the grid and emphasizes the need for improved congestion management. Thus, a new delimitation of price zones is frequently considered in current discussions and research. The present paper applies a novel hierarchical cluster algorithm that clusters locational marginal prices and weights nodes depending on their demand- and supply situation to identify possible new price zone configurations. The algorithm is applied in a scenario analysis of six scenarios reflecting main drivers that influence the future development of European Electricity markets in line with the trilemma of energy policy targets. Robustness of the new configuration is an important criterion for price zone configurations according to the European Guideline on Capacity Allocation and Congestion Management (CACM). Therefore, a robust price zone configuration is computed taking into account all the six individual scenarios. Results show that shape, size and price variations of price zones on the one hand strongly depend on the individual scenario. On the other hand, the identified robust configuration is shown to outperform other configurations, particularly also the current price zone configuration in CWE.

Keywords: Cluster Analysis, Electricity Market Design, Nodal Pricing, Congestion Management, Energy Markets and Regulation; Bidding zones, Price Zone Configuration, Bidding Zone Configuration

JEL-Classification: C38, C61, D47, L51, Q41, Q48

Tim Felling
House of Energy Markets and Finance
University of Duisburg-Essen, Germany
Universitätsstr. 12, 45117 Essen
+49-(0)201 / 183-xxx
Tim.felling@uni-due.de
www.hemf.net

Christoph Weber
House of Energy Markets and Finance
University of Duisburg-Essen, Germany
Christoph.Weber@uni-due.de

The authors are solely responsible for the contents which do not necessarily represent the opinion of the House of Energy Markets and Finance.

Content

1 Introduction and Literature Review

The face of electricity markets is constantly evolving. E.g. in 2015, Flow-Based-Market-Coupling has been introduced in Central Western Europe (CWE) and the extension to Central Eastern Europe is already in planning. The continuously growing capacities of renewables imply also a shift in generation locations and increasing fluctuating infeed. This has a severe impact on the congestion situation and grid operations. In Germany, redispatch costs tripled from 2012 to 2015 (Bundesnetzagentur and Bundeskartellamt, 2015, 2014, 2013, 2012). Therefore, new frameworks for the European Electricity markets are currently discussed in the literature and in the political arena. A potential solution is to reshape present price zones (bidding zones). Currently, a large ENTSO-E bidding zone study is undertaken that shall give insights into the effects of redesigned price zones in Europe. So far, national borders often align with borders of price zones. That might not be the optimal solution, as national borders do not necessarily reflect congestions in the grid.

The optimal solution for congestion management is often considered to be obtained via locational marginal pricing (nodal pricing), as nodal prices do not only reflect demand and supply characteristics but also congestions in the electricity grid. ((Hogan, 1992), (Stoft, 1997), (Egerer et al., 2016)). According to (Egerer et al., 2016), a single, uniform price for a zone might reflect wrong price information, since internal congestions and bottlenecks are not transparent. (Neuhoff et al., 2013) support the preceding findings stating that locational marginal prices (LMPs) constitutes to a more efficient grid utilization resulting in significant cost savings. Also, (Bertsch et al., 2015) investigate a zonal and nodal approach and conclude that LMPs are the best solution. Other configurations, e.g. zonal pricing or uniform pricing, would cause an increase of system costs of up to 4.6%. (Ding and Fuller, 2005) describe nodal pricing as the economically most efficient method as well, but also state that locational marginal pricing goes along with complex and complicated processes like data processing, accounting and financial settlements. (Walton and Tabors, 1996) investigate a variance criterion, namely the variance of LMPs between and within aggregated zones to evaluate zonal configurations, i.e. to reduce the amount of nodes in the Western Systems Coordinating Council (WSCC) system from 3500 to 20. The most well-known example for a system using LMP is Pennsylvania New Jersey Maryland Interconnection LLC (PJM) in the United States, but also most other deregulated electricity markets in the US use LMPs. Several price zones within one country can be found in Europe in Scandinavia where Norway, Sweden and Denmark are split into different price zones.

Currently, a zonal approach that aggregates similar nodal prices to zones seems to be more readily applicable in Europe, since the implementation of nodal pricing generally requires the establishment of an independent system operator (ISO) who combines the role of market

operator with (at least) part of the grid operation. The European guideline on Capacity Allocation and Congestion Management for Electricity (CACM) identifies evaluation criteria for future price (bidding) zone configurations. The main criteria are liquidity, market power, stability, robustness, network security and unbiasedness of prices in the new price zones. Before applying the criteria and analysing the results, a new configuration of price zones obviously has to be identified.

Two major methods, along with a few other approaches, have been developed in recent publications to delimitate these new price zones. The first one is to cluster the aforementioned LMPs to zones with similar prices, the other method refers to clustering of Power-Transfer-Distribution-Factors (PTDF). Within these two possibilities, various types of cluster algorithms have been applied, e.g. hierarchical, genetic or partition algorithms such as fuzzy-k-means ((Yang and Zhou, 2006)) . Also, underlying models vary from large scale applications to small examples like IEEE-test cases.

Clustering of LMPs is applied by (Imran and Bialek, 2008), (Burstedde, 2012), (Breuer et al., 2013), (Wawrzyniak et al., 2013), (Breuer and Moser, 2014). (Breuer and Moser, 2014) uses a genetic algorithm and applies the algorithm to a large scale model of the European transmission system for the years 2016 and 2018. He investigates redispatch costs, network security and also changing price zones depending on seasons. In contrast, (Burstedde, 2012) applies a hierarchical algorithm based on Ward's criterion to a simplified model of the European transmission system with 72 nodes. In addition, two different scenario years (2015 and 2020) are investigated and evaluated e.g. using total system costs. (Imran and Bialek, 2008) present three different approaches to cluster LMPs notably geographical clustering, fuzzy-c-means and price differential clustering. (Wawrzyniak et al., 2013) investigate zonal solutions based on LMPs for different wind scenarios on a Polish nodal system.

In contrast to the LMP-method (Duthaler, 2012), (Kang et al., 2013), (Klos et al., 2014), (Kłos et al., 2015), (Sarfati et al., 2015) and (Bergh et al., 2016) apply cluster algorithms based on PTDF-values. (Kłos et al., 2015) aim to reduce loop effects by clustering PTDF values. Their methodology refers to the mentioned goal of CACM to minimize adverse effects of internal transactions on other price zones. A similar approach is utilized by (Bergh et al., 2016). The authors cluster PTDF values on selected critical branches and, after investigating a base scenario, analyse several delimitations with different amounts of price zones. (Klos et al., 2014) identify critical branches using a clustering based on KKT-multipliers first and cluster PTDF-values afterwards. (Kang et al., 2013) investigate the IEEE-39 system for a given number of zones. (Sarfati et al., 2015) consider new delimitations based on five indicators, e.g. loop flows or price convergence on a 32-node model of the Nordic system. Also, three different wind-infeed scenarios are investigated.

Obviously, uncertainties strongly influence the delimitation of bidding zones. The dispatch of power plants, which depends mostly on variable costs, but also the development of demand, grid development and expansion of renewable energy sources (RES) affect congestions in the electrical grid, which in turn affect the LMPs and PTDF-values. Developments of all aforementioned factors may be related to the political choices made, which notably reflect the priority accorded to the different objectives within the triangle of energy policy targets (energy policy trilemma), namely security of supply, sustainability and economic efficiency (Spiecker and Weber, 2014).

Given these uncertainties and under consideration of the different clustering approaches and underlying models, the major contributions of the present paper are three-fold. First, the applied cluster algorithm has a clear economic foundation and objective, namely the minimization of price variations within the newly formed price zones. Second, a novel hierarchical cluster algorithm is applied, that weights nodes according to their relevance in terms of infeed and demand. By doing so, the importance of different nodes is acknowledged. All this is applied on a large-scale electricity system, namely the CWE system. Third, several different scenarios are considered for a single year (2020). The scenarios are derived by varying not only one but five different key drivers for electricity markets. This corresponds to an operationalisation of the robustness criterion for price zones referred to in Article 33 of the CACM guideline. The guideline mentions robustness and stability of price zones as relevant criteria without providing clear definitions nor a delimitation between the two. Hence robustness is understood here as referring to uncertainties within one period (year), whereas stability is interpreted as absence of (or limited) changes over time. Stability is not considered further here, yet could be treated within the same framework.

The remainder of the article is organized as follows: Section 2 outlines the developed methodology. First a general overview is given. Then the cluster algorithm is described, followed by the scenario construction and the evaluation methodology. Section 3 presents the application and the utilized grid- and generation models as well as the parameters retained for the different scenarios. The obtained results are then presented and discussed in section 4. Notably standard bidding zone configurations are compared to the robust one. Furthermore, the results are also compared to the current price zone configuration with five price zones in the (extended) CWE region (Switzerland, Netherlands, Belgium, France and Germany-Austria-Luxembourg). Note that Switzerland being not part of the EU is to date not member of the electric CWE region. Given the strong electrical interconnections it has both with France and Germany (lines to Austria are less developed), it is yet subsequently included in the analysis.

2 Methodology

The overall methodology of this paper is sketched in Figure 1. The starting point is the construction of scenarios reflecting the relevant uncertainties (cf. section 2.3).

For each of the investigated scenarios, nodal hourly prices (LMPs) for the year 2020 are then computed by applying a DC-OPF on a detailed grid of Central Western Europe (CWE). In view of the later clustering, the infeed and demand situation at each node in each scenario is used to calculate nodal weights. A more detailed description of the grid model and data will be given in sections 3.1 and 3.3. The computed LMPs for each scenario are then clustered into aggregated price zones under consideration of the infeed and demand situation at the nodes using a hierarchical cluster algorithm. Thus, new price zones are computed for each of the scenarios resulting in optimal price zone configurations for each scenario for different numbers of zones.

Additionally, the hourly prices of all six scenarios are clustered together to determine the so-called robust configuration (cf. section 4.5.3 for a discussion whether this configuration is robust in a mathematical sense). Details on the developed algorithm can be found in section 2.1 and 2.2.

The results of both the optimal price zones for the individual scenarios and for the robust scenario may then be analysed (cf. section 4).

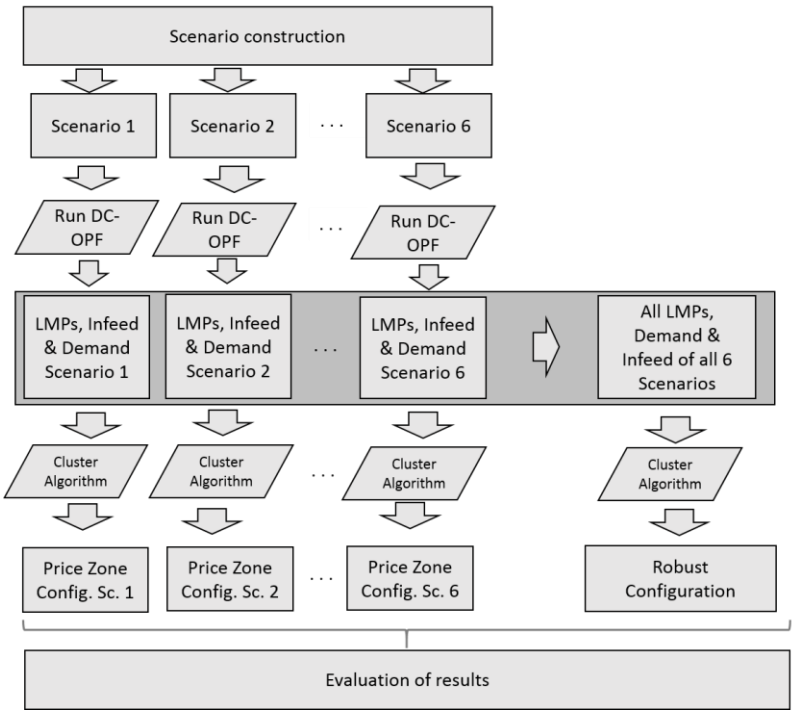

Figure 1: Methodology of this paper

- Background of scenario-analysis in section 2.3
- Description of the six scenarios and the identified set of parameters in section 3.2
- Application and description of DC-OPF in section 3.1
- Methodology of cluster algorithm in section 2.1
- Application in section 3.3
- Evaluation methodology in section 2.4
- Evaluation of results in section 4

As previously mentioned, the remainder of this section is organized as follows: First, the cluster algorithm is presented in sections 2.1 and 2.2. The scenario approach (2.3) and the evaluation methodology (2.4) are discussed subsequently.

2.1 Cluster Algorithm

The cluster algorithm may be used to determine new price zone configurations for individual scenarios as well as for a combination of scenarios. In the latter case, a robust configuration is obtained (c.f. (Felling and Weber, 2016)).

Problem formulation

The main objective of the algorithm is to minimize price variations within a zone. Therefore, we take the total weighted variation (*V)* of prices in the system under study as a starting point (1). Thereby $p_{n,h}$ refers to the hourly price at a node whereas $\overline{p_h}$ equals the average hourly price of all nodes in the system. *N* is the set of all nodes and *H* the set of all hours considered. In case a robust configuration is looked for, the set *H* includes the hours from the different scenarios considered.

$$\mathrm{V} = \sum_{\mathrm{n}\in\mathrm{N}} \sum_{\mathrm{h}\in\mathrm{H}} \left(\left(\mathrm{p_{h,n}} - \overline{\mathrm{p_h}} \right)^2 \cdot \mathrm{W_n} \right) \tag{1}$$

The last multiplier W_n corresponds to the weight of a node. To account for the different relevance of nodes in the system, weights are assigned according to the overall supply and demand quantities in the nodes. Weights are normalized so that their average equals one to facilitate interpretation. Numerical singularities are avoided by assigning a weight of 0.001 to nodes that have neither load nor infeed. Details of the computation are given in the appendix, cf. equations (A21) – (A23).

Equation (2) shows the computation of the average hourly prices:

$$\overline{p}_h = \left(\sum_{n=1}^{N} \left[p_{h,n} \cdot W_n \right] \right) \cdot \frac{1}{\sum_N W_n} \tag{2}$$

As the main objective is to identify price zones minimizing within-zone price variations, the total weight of a zone is set equal to the sum of the weights of all nodes within that particular zone, cf. equation (3). Thereby the assignment of nodes to zones is of importance. The nodes of a zone (cluster) c are represented by $n \in N_c$.

$$W_c = \sum_{n\in N_c} W_n \tag{3}$$

Analogously to (2), the average hourly price of a zone is the weighted average of the nodal prices for the nodes within the cluster.

$$\overline{p}_{h,c} = \left(\sum_{n \in N_c} [\, p_{h,n} \cdot W_n] \right) \cdot \frac{1}{\sum_{n \in N_c} [W_n]} \tag{4}$$

The within-zone variation (V^{within}) presented in equation (5) is the variation of hourly prices of a zone compared to the average price of the particular zone. Thus, the total within-variation over all zones is computed using (3) and (4) as follows:

$$V^{within} = \sum_{c \in C} \sum_{n \in N_c} \sum_{h \in H} \left(\left(p_{h,n} - \overline{p_{h,c}} \right)^2 \cdot W_n \right) \tag{5}$$

Based on equation (5) the within-zone variation in a single zone (V_C^{within}) is presented in equation (6) .

$$V_C^{within} = \sum_{n \in N_c} \sum_{h \in H} \left(\left(p_{h,n} - \overline{p_{h,c}} \right)^2 \cdot W_n \right) \tag{6}$$

The total variation in the system (1) can then be decomposed into the variations between- and within-zone as shown in equation (7).

$$V = V^{within} + V^{between} \tag{7}$$

Thereby, the between-zone variation is derived analogously to the within-zone variations. In contrast to the within-zone variation, the between zone variation is given by the squared difference of the hourly average zonal prices to the hourly overall mean weighted prices, cf. equation (8).

$$V^{between} = \sum_{c \in C} \sum_{h \in H} \left(\overline{p_{h,c}} - \overline{p_h} \right)^2 \cdot W_C \tag{8}$$

The objective is to determine a new price zone configuration that has a low within-zone variation and a respective higher between-zone variation, so that price differences occur rather between zones than within zones. The clustering approach used here can hence be viewed as a generalization of Ward's method (cf. e.g., (Härdle and Simar, 2003) pp. 396-398). (Batagelj, 1988) provides some properties this and even more general versions of Ward's method.

Iterative price zone aggregation

The aforementioned equations form the basis of the clustering algorithm. It clusters the single nodes stepwise to aggregated zones. In the first step, each node corresponds to a single zone, in which case the within-zone variation is zero. All variations are between the nodes. During the next steps, the increase (delta) of the within-zone price variations is computed for every possible merger of two zones. For example, when two nodes with different prices are merged to a new zone, the within-zone variation of the whole system increases. This increase (delta) is computed for all possible mergers. The two zones (nodes) with the least increase are

merged. This procedure is repeated until no zones are left to merge, i.e. when all nodes are grouped into one single zone and the between-zone variation becomes zero.

The computation of the increase in weighted within-zone variations is shown in equation (9) based on (5). The within-zone variations for the new zone $C_i \cup C_j$, including all nodes from both old zones, are computed first. Then the weighted within-zone variations of the two old zones C_i and C_j are deducted to obtain the increase.

$$\Delta V_{C_i \cup C_j}^{within} = \sum_{n \in C_i \cup C_j} \sum_{h \in H} \left[\left(p_{h,n} - \overline{p}_{h,C_i \cup C_j} \right)^2 \cdot W_n \right] - V_{C_i}^{within} - V_{C_j}^{within} \quad (9)$$

A merger is thereby only possible if the two old zones are physically connected by at least one transmission line. This ensures that any newly formed zone will only consist of nodes that are connected physically. The respective information is obtained from the node adjacency matrix (NAM).

2.2 Flowchart of the Cluster Algorithm

The algorithm can be divided into four main parts as Figure 2 visualizes.

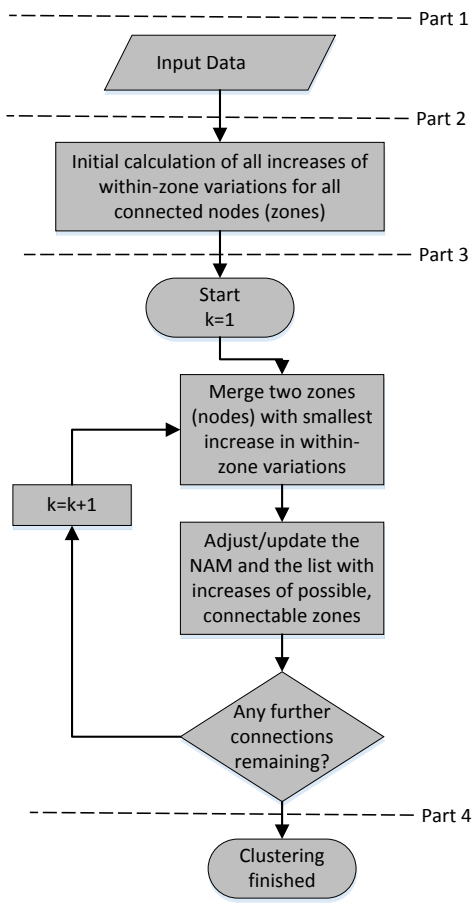

Figure 2: Flowch1art of the algorithm

- **Part 1: Input data preparation:** The algorithm requires three types of data.

  Prices: For each node, an array of prices (e.g. one price per node per hour) is handed over. Thus, by inserting hourly LMPs for one year, all grid load cases of a year are represented by the variation of demand and infeed in the separate hours.

  Demand and infeed values: For each node, the load situation is characterized by one value (the sum of average infeeds and loads). These are used to compute the nodal weights.

  Node adjacency matrix of the grid: This matrix contains the information about which nodes are physically connected.

- **Part 2: List of increases of new configurations:** For all nodes that are physically connected, the algorithm computes the increase (delta) in within-zone variation according to (9) once. For example, in a network with three nodes, where all nodes are connected, the algorithm would initially compute the increase of all three possible new zonal configurations.
- **Part 3: Iterative merging:** Subsequently the smallest increase is chosen and the associated nodes (zones) are merged to a new zone. The computed list of increases from step 2 is then updated for all zones that are part of the two recently merged zones or are their neighbours using an adapted version of the Lance-Williams-formula (cf. (Lance and Williams, 1967; Murtagh and Legendre, 2014)). The remaining increases (deltas) are not affected and remain the same. Also, the NAM has to be adjusted.
- **Part 4: End:** When only one zone is left, the algorithm stops.

The overall result is a sequence or set of optimal price zone configurations, because the result of each single iteration corresponds to the optimal zonal configuration for the particular number of zones. Each configuration is optimal regarding the within-zone variation.

As presented in Figure 1, the cluster algorithm may be applied for each scenario (cf. section 2.3). Additionally a robust configuration may be determined by taking the particular LMPs of the different scenarios together into one price array.

2.3 Scenario Analysis

2.3.1 General considerations

The location and shape of price zones as identified through the clustering algorithm mostly depend on the underlying locational marginal prices (LMPs). This minimization of within-zone price variation corresponds to the welfare-maximization objective of a (perfect) regulator under the hypothesis that price variations within zones lead to welfare losses. This hypothesis seems plausible since within zone price variations will imply inefficient redispatch to remove congestions and inadequate incentives for investments. Yet the challenge for the regulator is that he has to make his choice of price zones under (irreducible) uncertainty – especially when this choice has to be made several years in advance to allow market participants to adapt. The resulting decision situation is visualized in Figure 3.

In order to identify the optimal configuration, future LMPs have to be forecasted. Those are influenced by various parameters such as demand, grid development and generation characteristics. They are, as the future energy markets in general, not trivial to predict as they are influenced by further factors as well. These factors are usually outside the control of the regulator and include world energy market developments, general economic growth and energy policies. E.g., generation at a node and in a network depends not only on different fuel

prices but also on the amount of conventional und renewable capacities. Those again are mostly affected by energy policies. Furthermore, energy policies not only influence generation capacities but also grid development.

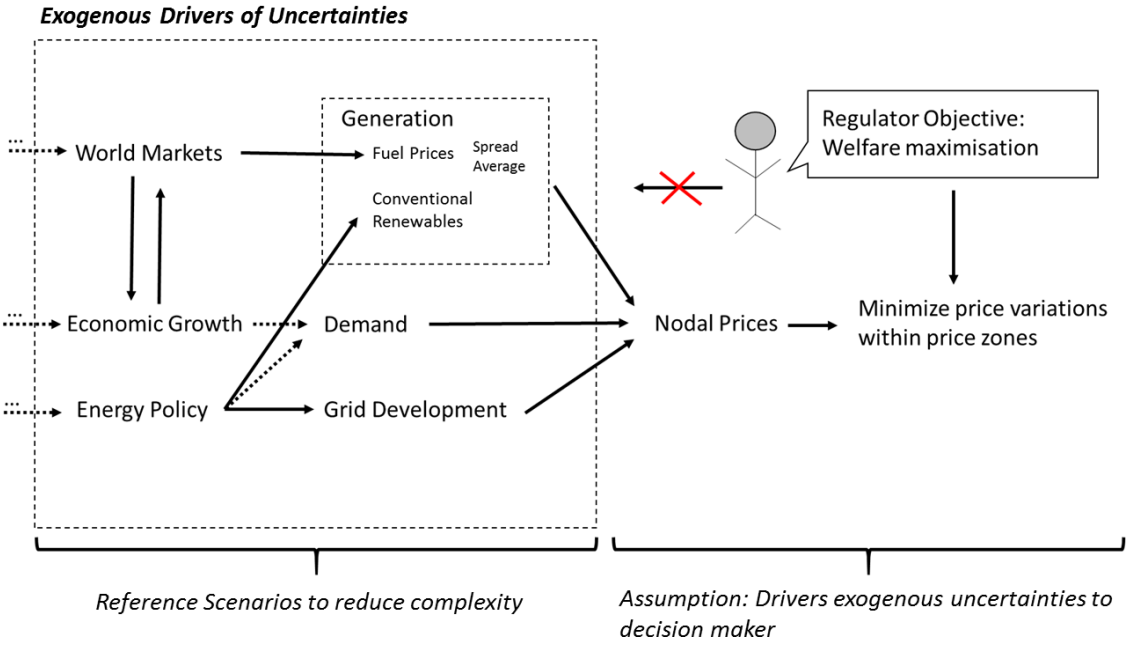

Figure 3: Decision Situation and Exogenous Drivers of Uncertainties

In conclusion, a regulator will have to consider different, descriptive scenarios when it comes to fixing price zones several years in advance. Identifying suitable scenarios is a challenging task since a limited number of scenarios should be chosen to represent as good as possible the range of future uncertainties.

This task can hardly be accomplished from scratch within this paper, especially as the focus of this paper is on the clustering methodology for price zone configurations. Therefore, we base our selection of scenarios on an existing approach (Spiecker and Weber, 2014), where systematically different development paths for the European energy system have been investigated. The starting point for the scenario description is there the triangle of energy policy targets, also called the energy policy trilemma.

From the scenario outline one has then to derive a consistent set of parameters that are both relevant for the question at hand and distinguishable between the scenarios. Since local marginal prices are used as input for the clustering algorithm, the parameters used to characterize the scenarios should have a major impact on the nodal prices. At the same time they should be linked back to the economic and political developments describing the scenarios (cf. Figure 3).

Therefore we use the following parameters to characterize and differentiate the scenarios:

- Average fuel price level
- Fuel price spread
- Expansion of renewables
- Demand development
- Status of grid development

The parameter choices for the selected scenarios are introduced in section 3.2. Yet in order to ensure that choices are consistent with observed developments in the past, we develop subsequently an approach to derive price parameters based on empirical estimates.

2.3.2 Determination of price parameters

Figure 3 illustrates that fuel prices are one key parameter of uncertainty that strongly influences LMPs. For the merit order as well as the congestion costs, not only the absolute fuel price level is relevant but also the price or rather cost spreads between different generation technologies. Of particular relevance are the spreads between coal-fired and gas-fired generation units, since those are frequently setting the prices in the European market. Therefore we use two parameters to characterize the fuel price development, on the one hand the average fuel price and on the other hand the spread in generation costs for coal and gas units. Subsequently we briefly explain the methodology to determine possible future variations in these parameters.

As starting point we take the wide-spread assumption that commodity prices like other asset prices follow a geometric Brownian motion (Hull, 2015). This implies that log price changes are independently, identically and normally distributed. Since we are ultimately interested in marginal electricity generation cost, we do not model fuel and CO_2 prices separately. Rather we model directly the variable generation cost for a typical coal plant c_{coal} (including fuel and CO_2 certificate cost) and for a typical gas plant c_{coal}. Those may be written

$$\mathrm{c_{i,t}} = \frac{\mathrm{c_{i,t}^{f}} + \varepsilon_\mathrm{i} \cdot \mathrm{c_t^{CO2}}}{\eta_\mathrm{i}} \qquad \mathrm{i} \in \{\mathrm{coal}, \mathrm{gas}\} \tag{10}$$

With the plant efficiencies η_i and carbon intensities of fuels ε_i.

Then logarithmic averages and spreads may be computed:

$$ln\,\mathrm{c_{m,t}} = \frac{ln\,\mathrm{c_{gas,t}} + ln\,\mathrm{c_{coal,t}}}{2} \tag{11}$$

$$ln\,\Delta\mathrm{c_t} = \frac{ln\,\mathrm{c_{coal,t}} - ln\,\mathrm{c_{gas,t}}}{2} \tag{12}$$

And for these price factors now independent, identically increments are assumed, i.e.:

$$\Delta\left(ln\,\mathrm{c_{m,t}}\right) \sim \mathrm{N}(\mu_\mathrm{cm}, \sigma_\mathrm{cm}) \tag{13}$$

$$\Delta(ln\,\Delta c_t) \sim N(\mu_{\Delta c}, \sigma_{\Delta c}) \quad (14)$$

Thereby the parameters for the normal distributions are obtained from an appropriate sample of historical observations. By modelling the mean price development and the spread development as separate, orthogonal factors, we automatically include the observable correlation between coal and gas prices in the modelling.

When it comes to derive future parameter values, we may use the following well-known property of Brownian motions:

$$ln\,c_{m,t+T} - ln\,c_{m,t} \sim N\left(T \cdot \mu_{cm}, \sqrt{T} \cdot \sigma_{cm}\right) \quad (15)$$

$$ln\,\Delta c_{t+T} - ln\,\Delta c_t \sim N\left(T \cdot \mu_{\Delta c}, \sqrt{T} \cdot \sigma_{\Delta c}\right) \quad (16)$$

i.e. the drift scales linearly with the forecast time horizon whereas the standard deviation increases by the square root of the time difference. If within our scenarios we consider three hypotheses *high, mid* and *low* for these prices, then a rather straight forward choice is to select $T \cdot \mu + \sqrt{T} \cdot \sigma$ for the price change in the *high* hypothesis, $T \cdot \mu$ for the *mid* case and $T \cdot \mu - \sqrt{T} \cdot \sigma$ for the *low* case. Those correspond than roughly to the 85%, 50% and 15% quantile of a distribution that has been empirically derived instead of being chosen arbitrarily.

2.4 Evaluation Methodology

As previously described, the cluster algorithm computes the optimal price zone delimitations for either an individual scenario or for a combination of all scenarios. The latter price zone delimitation is called the robust configuration (cf. Figure 1). The question then arises what the relative merits or disadvantages of these different configurations are. In addition, it might be of interest to compare those optimized configurations to others, exogenously defined ones, notably the current CWE configuration.

For evaluation, obviously the within-zone variations according to equation (5) may be assessed for each of the configurations. In order to evaluate the configurations, it is however not only of interest to investigate the performance of each configuration in the particular scenario for which it has been optimized. Rather their performance may be checked by evaluating them also under all other possible scenarios. For example, the configuration determined in the first scenario may be fed with the LMPs computed for the second scenario and then the within-variation may be computed. This then indicates whether the first configuration still performs reasonably well if scenario 1 instead of scenario 2 materializes. This is of particular importance if the selection of the configuration has to be done well-ahead of the actual implementation. Then severe uncertainties persist at the moment of configuration selection which will be resolved when the configuration is actually implemented. The current European procedures are likely to induce such a time gap of two to three years.

To perform the corresponding evaluation, the assignment of nodes to zones in equation (5) is changed. Thus, for the same LMPs the within-zone variation changes with a different assignment of nodes to zones ($n \in N_c$) as the zonal average price $\overline{p_{h,c}}$ changes. Ergo the variations of multiple configurations for a scenario, which materializes as a set of LMPs, can be investigated.

Furthermore, the dependence of the within and between variations on the number of zones is of interest. This provides an indication on the potential benefits of smaller over larger zones.

Besides these assessments based on the selection criterion itself, further criteria are of interest. Notably the size of zones is an indication of balanced design. As a measure for the size of zones, the percentage of the weight of a cluster (W_C, c.f. equation (3)) in the total weight of the system is assessed.

3 Application

The presented methodology and developed cluster algorithm is applied to a model of the (future) electricity grid of Central Western Europe and its neighboring countries as subsequently described. Covered countries are Austria, Switzerland, Belgium, France, Germany, Luxembourg and the Netherlands. Before describing the scenarios and parameters used in section 3.2, the underlying models for demand, renewables and grid modelling are briefly described in section 3.1. Then some computation details on the clustering algorithm are given in section 3.3.

3.1 Underlying Models

The grid model is based on public available data. Envisaged network extensions until 2020 are modelled according to the European Ten-Year-Network-Development-Plan (TYNDP) (ENTSO-E, 2014) and the German grid development plan (Bundesnetzagentur, 2014). The status of the network extension, as mentioned in section 2.3.1 and discussed in more detail below, is one key variable that differs across the scenarios. Data for load, renewable infeed and distributed conventional generation are simulated by a vertical load model developed in the research group based on public available data (Osinski, 2016). The model of the CWE transmission grid compromises over 2200 nodes, 3600 branches and 600 transformers of the voltage levels 220 and 380 kV. In Germany, relevant 110 kV nodes are represented as well. Furthermore, phase shifters are incorporated into the model.

Market related relevant values are derived from the Joint Market Model (Meibom et al., 2011; Tuohy et al., 2009). The model simulates day-ahead and intra-day markets and is capable of optimizing hydro- and combined heat production. Data like cross-border flows to non-CWE countries, infeed and shadow prices for hydro power stations and combined heat production are integrated into the grid model via an interface.

3.2 Scenarios

The background of the scenario analysis have been outlined in section 2.3. The scenario family taken here as starting point (Spiecker and Weber, 2014) is exploring different priority settings within the energy policy trilemma (cf. above). In the following two sections, the selected six scenarios are presented. Afterwards the corresponding parameters are discussed.

3.2.1 Scenario Outlines

The scenario storylines for the first five scenarios correspond to those presented by (Spiecker and Weber, 2014). A sixth scenario has been added, which reflects a policy vision of improved control over the relevant variables. This scenario describes what policy makers aspire at. All scenario outlines have been adjusted to reflect particularly the parameters of interest in the context of this study.

Scenario 1: "Realistic Scenario"

This scenario is characterized by continued conflicts between the different energy policy priorities. On the one hand, extension of renewable capacities will progress as planned and reach the targeted goals of SOAF for 2020 (Scenario B) (ENTSO-E, 2015). On the other hand, grid development is expected to be delayed by four years, which is considered a realistic assumption given the current gap between the targeted commissioning dates and actual progress of grid extensions. Prices, demand and the spread between coal and gas prices are assumed to develop according to medium forecasts.

Scenario 2: "Climate Policy"

The emphasis in this scenario is on the extension of renewable capacities and energy efficiency measures. A reduction of greenhouse gas emissions constitutes the main objective. This is reflected by an increasing and accelerated extension of renewable capacities and decreasing demand due to efficiency measures. Less effort is put on the reinforcement of the grid, which is thus expected to be delayed by two years. The shrinking demand for conventional resources will lead to a drop of prices for resources and CO2-certificates.

Scenario 3: "Climate Market"

In comparison to scenario 2, this scenario focuses on a market for greenhouse gas emissions (CO2). Besides an extension of renewables, greenhouse gas emissions shall be reduced by raising costs for emission certificates. Thus, the spread between coal and gas prices will decrease and the price for CO2 will rise.

Scenario 4: "Security of Supply"

The scenario "Security of Supply", in contrast to the climate scenarios, sets the focus on grid development rather than extension of renewable capacities. In fact, extension of capacities is

slowed down to protect the grid from too high quantities of fluctuating renewable infeed. This causes, again in contrast to the climate scenarios, higher demand for conventional generation and therefore increasing prices.

Scenario 5: "Economic Growth"

This scenario expects economic growth that is reflected in rising demand in line with rising prices for fuels. The further parameters remain on a medium level.

Scenario 6: "Alternative Scenario"

This "alternative" scenario illustrates a rather unrealistic path assuming that some of the mentioned drivers are not exogenous to the decision maker. Therefore, both stability and sustainability goals are achieved. The grid development progress as scheduled and expansion of renewable capacities overachieves the current aims. The further parameters remain on a medium level.

The following section presents the chosen parameters that are varied to determine the scenarios.

3.2.2 Scenario Parameters

In section 2.3.2, the methodology to derive meaningful and empirically validated parameters for fuel prices has already been discussed. For the empirical application, weekly fuel price changes for the year 2015 have been analysed. The expected drift has been set to zero, since empirical estimates have been insignificant. and the corresponding annualized standard deviations have been determined as:

$$\sigma_{\mathrm{cm}} = 0{,}117 \tag{17}$$

$$\sigma_{\Delta \mathrm{c}} = 0{,}066 \tag{18}$$

The CO2 prices are chosen so as to support the change in the cost spread.

For the further parameters mentioned in section 2.3.1, grid development, demand and the extension of renewables the retained hypotheses are explained subsequently:

Grid Development

The status of the grid development is differentiated in: progress as scheduled (*no delay*), delay of two years (*delayed*) and delay of four years (*very delayed*).

Expansion of renewables

Regarding the build-up of renewable capacities, capacities of the SOAF for 2020 (Scenario B) (ENTSO-E, 2015) for the year 2020 are chosen as *mid* case. An accelerated (*high*) or slowed down (*low*) build-up is translated into a 50% increase/decrease of the additional capacities from 2015 on.

Demand

The demand case *mid* refers to the same demand as in 2015. An increased demand (*high*) is translated into 5% increase in demand, a decrease (*low*) into minus 5%, respectively. This equals 1%-point per year.

An overview over the parameters selection for the scenarios is given in Table 1.

Table 1: Parameter selection for the scenarios

	Scenario 1 "Realistic Scenario"	Scenario 2 "Climate Policy"	Scenario3 "Climate Market"	Scenario4 "Security of Supply"	Scenario 5 "Economic Growth"	Scenario 6 "Alternative Scenario"
Grid Development	Very delayed	Delayed	Delayed	No delay	Delayed	No delay
Extension of Renewable Capacities	Mid	High	High	Low	Mid	High
Demand	Mid	Low	Low	Mid	High	Mid
Average Prices	Mid	Low	Low	High	High	Mid
Spread Coal vs. Gas	Mid	High	Low	Low	Mid	Mid
Price for CO2	Mid	Low	High	High	Mid	Mid

The costs for fuel oil and light oil are adapted to the development of gas prices. Overall, this results in the following prices shown in Table 2.

Table 2: Fuel prices for scenarios

	Scenario 1 "Realistic Scenario"	Scenario 2 "Climate Policy"	Scenario 3 "Climate Market"	Scenario 4 "Security of Supply"	Scenario 5 "Economic Growth"	Scenario 6 "Alternative Scenario"
Gas [€/MWh]	22,4	20,4	13,5	24,5	29,6	22,4
Coal [€/MWh]	8,0	5,3	5,4	12,0	11,2	8,0
Fueloil [EUR/MWh]	23,6	21,4	14,2	25,7	31,1	23,6
Lightoil [EUR/MWh]	37,5	23,5	23,5	59,7	59,7	37,5
CO2 [€/t]	8,1	5,3	12,3	12,3	8,1	8,1
Lignite[EUR/MWh th]	1,5	1,5	1,5	1,5	1,5	1,5
Nuclear[EUR/MWh]	1	1	1	1	1	1

3.3 Price zone computation

The parameters in the different scenarios affect the LMPs and the nodal weights in the scenarios (cf. Figure 3 and Table 1). Both the LMPs and the weights are inserted to the cluster algorithm. Overall, the algorithm is applied seven times. Once for each scenario and one time for the robust configuration. For the robust configuration, the hourly prices of all six scenarios are handed over as input data, as presented in Figure 1 and Figure 2.

As a result, six price zone configurations, each optimal in the particular scenario, are computed. In addition the robust configuration that is based on the LMPs of all scenarios is determined. Thereby the algorithm does not compute the optimal number of zones for a scenario but the optimal configuration for a given number of zones. This is because a hierarchical cluster algorithm is applied. During each step, starting with a nodal setup where each node is a zone, the optimal zonal solution is computed. Thus, the sequence of configurations from each cluster-application offers the optimal solution for two, three, four etc. zones.

Thereby the algorithm is able to compute the configurations quite fast. The computation time of the robust configuration is higher than the time for the single scenario configurations, simply due to the amount of data. Table 3 shows the computation time for a single scenario and the robust configuration. The calculation is undertaken on an 8 GB RAM System with an Intel® Core™ i5-4690 CPU with 3.5 GHz. Despite the amount of data, the algorithm is able to find the solution for the single scenarios in about 2 minutes. Since the amount of LMPs is six times higher for the robust configuration, the computation takes about a factor five longer.

Table 3: Computation Time

	Configurations 1 – 6	Robust Configuration
Computation Time	110 s– 121 s	530 s– 564 s

4 Results

This section presents the results of the scenario analysis and the cluster algorithm. It organizes as follows: At first, the LMPs and the standard deviation of prices at nodes are presented since the nodal prices are the key input parameter of the algorithm. The LMP values and their standard deviation already offer a good first impression why the later presented price zone configurations differ between the scenarios.

Second, the resulting price zone configurations of each scenario for a fixed number of price zones is presented. Five zones are chosen as this equals the current number of price zones in the considered extended CWE region. Also the robust five zone configuration is presented that takes into account the different possible scenarios.

Third, results regarding varying numbers of zones are shown for the robust configuration. Section 2.1 has already explained that the algorithm computes a sequence of optimal configurations for different numbers of zones. Thus, results for selected numbers of zones and their shape, size and their values for within and between variations are shown.

Finally, the benefits of the robust configuration are evaluated using the evaluation methodology outlined in section 2.4. This is done by comparing it to the single scenario configurations and

to the current configuration in CWE. Moreover, the size and within-zone variations are analysed.

4.1 Resulting LMPs

This section illustrates the LMPs for the scenarios by showing their nodal average and standard deviations. The LMPs are computed from a DC-OPF using MATPOWER based on the underlying models and scenario assumptions explained in section 3. The detailed optimization problem formulation can be found in (Zimmerman et al., 2011). N-1-secure prices are approximated by reducing line capacities to 85% of the original values.

Figure 4 presents the average LMPs for each of the six scenarios. The average nodal LMPs are computed according to equation (19).

$$\overline{p_n} = \frac{1}{|H|} \cdot \sum_{h \in H} p_{h,n} \tag{19}$$

The colors indicate price differences from low prices (blue) to higher prices (red). Thereby obviously unrealistic prices in a few of the simulated hours (eventually related to inaccuracies of the grid and load model and the regional distribution) of under 0 €/MWh and above 150 €/MWh were set to 0 €/MWh respectively 150 €/MWh. This approach has been applied to less than 1 % of all prices.

The first obvious observation is that resulting LMPs differ significantly between the scenarios. This observation is confirmed when looking at the nodal standard deviations presented in the following section.

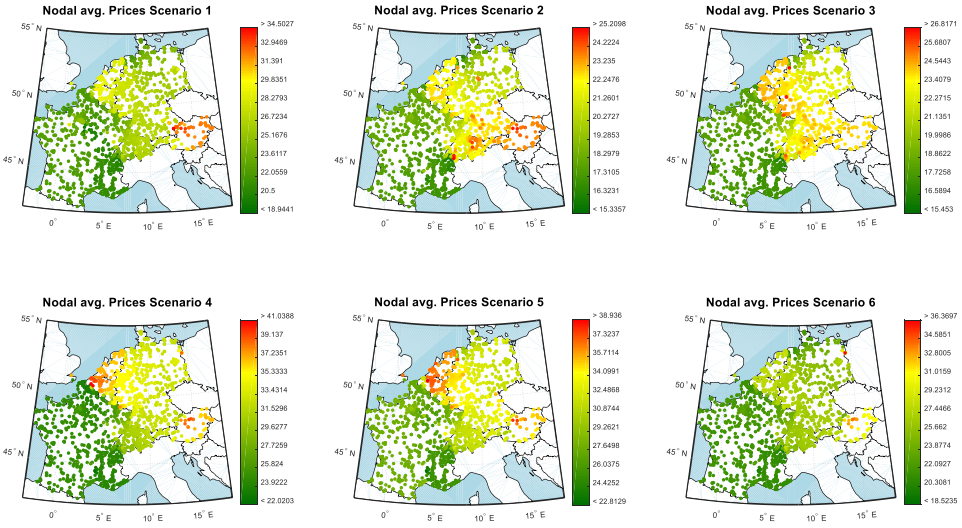

Figure 4: Average nodal prices for the six scenarios

4.2 Resulting Standard Deviation at Nodes

Figure 5 illustrates the standard deviation at nodes since price variations are detected by the algorithm according to equation (9). The presented standard deviation at nodes are derived according to equation (20).

$$\sigma_n = \sqrt{\frac{1}{|H|} \cdot \sum_{h \in H} (p_{h,n} - \overline{p_n})^2} \tag{20}$$

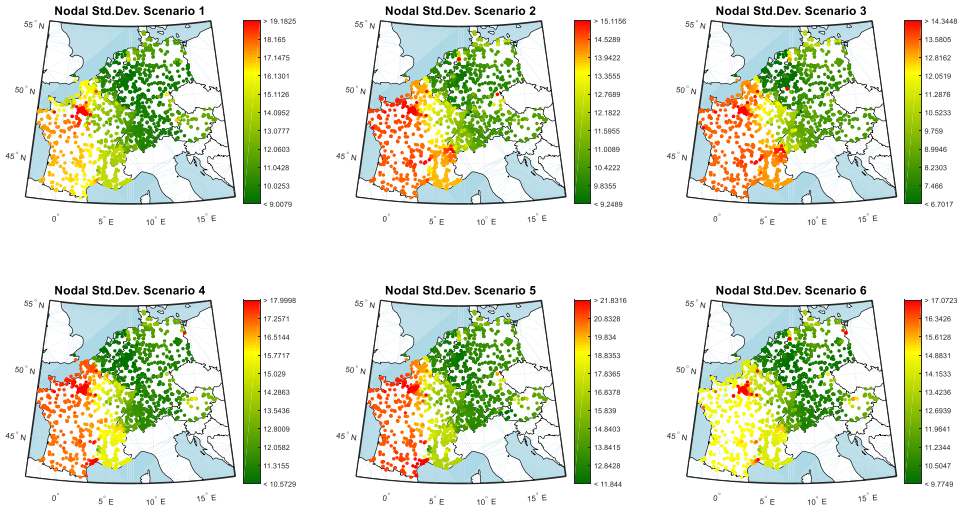

Figure 5: Standard deviations of nodal prices for the six scenarios

The cluster algorithm detects both the difference in average prices and in the variations between nodes as shown in the next section.

4.3 Resulting Price Zone Configurations

This section presents exemplary price zone configurations for the scenarios based on the LMPs described in the previous section. Results for configurations with five zones are presented, as there are also currently five price zones in the considered extended CWE region. First, the six configurations corresponding to the six scenarios are shown. Then the robust configuration with five price zones is presented.

4.3.1 Price Zone configurations of the scenarios

In Figure 6, the shapes of the resulting price zones are visualized. Two major observations should be highlighted when considering the visual shape of price zones.

First, borders of price zones do not necessarily align with national borders any more. Each country is split into more than one zone that are merged with parts of other countries. The individual price zones of Belgium and the Netherlands disappear. Both zones are united in all scenarios and further merged with parts of Germany. The eastern part of Austria is split apart

from the remaining four price zones in three out of six scenarios (scenario one, four and six). In each of the three cases it is also the most expensive price zone. In reference to that, the German-Austrian-Luxembourgian price zone is at least split into three new price zones. Throughout all six scenarios, a price zone in northern or north-eastern Germany emerges that has lower average prices than the other, more southern zones, what is due to the high wind feed-in in that area.

In comparison to that, France is at most only split in two parts and almost remains a zone by its own in scenario four and six. Eventually, one reason for this is the applied methodology of detecting price differences. Since nuclear power stations set marginal generation cost at similar levels at most nodes in France (especially, as time restrictions like ramp rates are not modelled in the DC-OPF), the algorithm rather adjusts power plant outputs than to detect price differences. Furthermore, Switzerland is always incorporated into a South-German-Austrian zone.

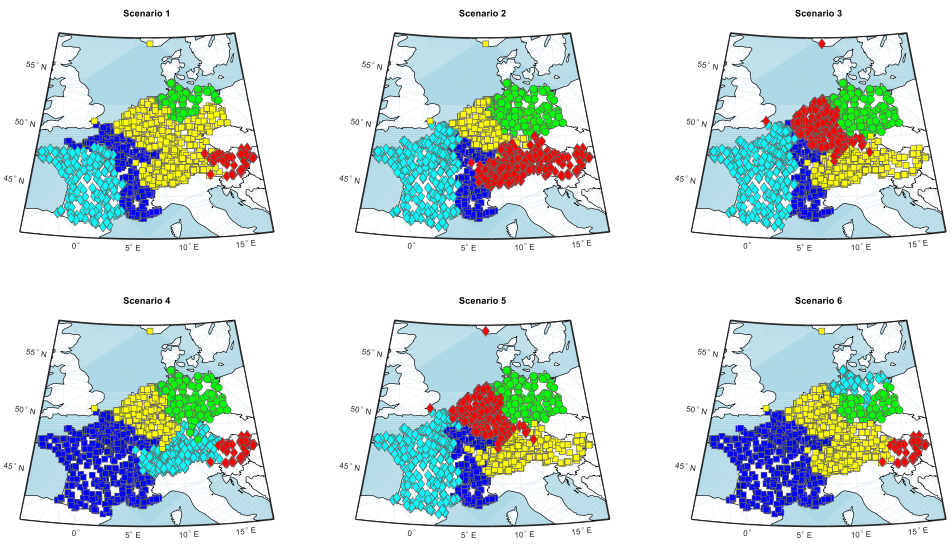

Figure 6: Shape of price zones for 5 zones in CWE for the six scenarios

Second, the price zones differ between the scenarios although similarities are detectable. This observation is in line with Figure 4 and Figure 5 that already illustrated the difference of the average LMPs between the scenarios. Between scenario two and three, the resulting price zones differ, although only the price spread between coal and gas is varied. The difference appears small, yet the difference in average prices of the two German-BeNeLux-Swiss-zones is reversed. In scenario 2, where the spread is high, the zone including Netherlands and Belgium has a lower average price than the other partly German-Austrian-Swiss-zone. In case of a low spread and a high CO2-price, the BeNeLux-zone grows and the average price rises above the average price of the Austrian-German-Swiss-zone. The main reason is that the power plants in the BeNeLux and West-German region are more affected by the high CO2-

price than those in the other zone where also hydro power stations and nuclear power plants exist. In scenario five, which is based on the same status of grid development as two and three, the rising demand appears to affect the Belgium-Dutch zone more than the other zones. In this particular scenario and in scenario 3, the Dutch-Belgium-Western-German price zone has the highest average prices. In all other scenarios, the highest prices are (as previously mentioned) found in the southern German- and/or- Austrian price zone.

In conclusion, none of the existing price zones remains the same. The big German-Austrian-Luxembourgian price zone is split into at least two separate zones and France is split in four out of six scenarios into two zones. The rather small zones like Switzerland, Belgium and Netherlands (and Austria) are merged and join parts of Germany. In addition, borders of price zones differ between the six scenarios.

4.3.2 Resulting Robust Price Zone Configuration

In parallel to Figure 6, the results for the robust price zone configuration for five price zones are presented in Figure 7. As mentioned before, the robust configuration is determined based on all the LMPs of all six scenarios.

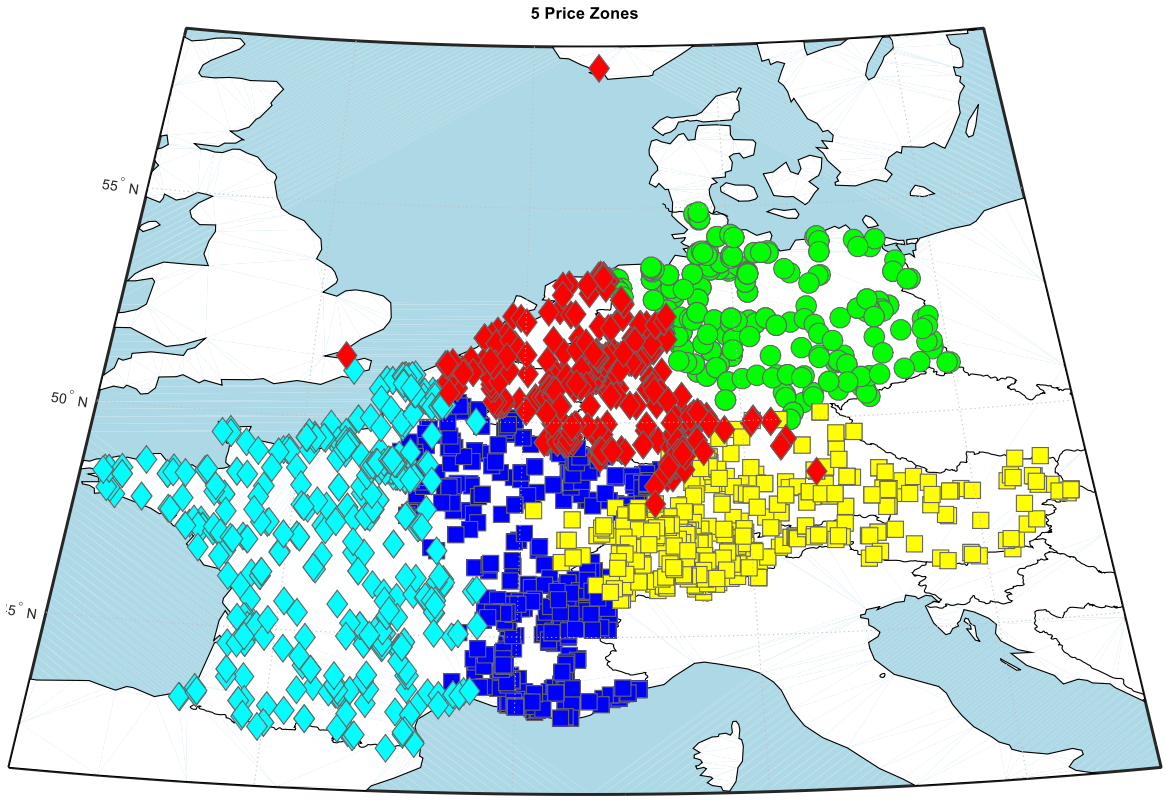

Figure 7: Robust Configuration for 5 Zones

In line with the observations made in section 4.3.1 and Figure 6, price zone borders do not align with national borders anymore. Germany is split into three parts. The northeastern zone has lower average prices than the other two regions. The western region merges with a Belgium-Dutch price zone. In comparison to the other regions, this appears to be the most expensive one on average. The southern zone merges with Austria, Switzerland and a few nodes of France. France is split in two parts. The split is more or less straight in north-south direction passing just east of Paris. Thereby the eastern French zone has the lowest average price of all zones.

4.4 Results regarding number of zones

The previous two sections have illustrated the scenario specific optimal configurations and the robust configuration in case of five price zones. As mentioned in sections 2.1 and 3.3, the algorithm computes not one optimal configuration per scenario but a sequence of optimal configurations for different numbers of price zones. Therefore, the robust configurations for selected numbers of price zones are presented subsequently. In addition, key parameters like the within- and between-zone variations are investigated.

4.4.1 Resulting robust configurations

Figure 8 shows the results for robust price zone configurations for different numbers of zones. Thereby the resulting configurations are presented in a decreasing order starting from 20 price zones. This mirrors the hierarchical, stepwise aggregation of zones by the cluster algorithm as described in sections 2.1 and 2.2. During the step from eight to seven price zones the algorithm for example detects the least increase in within-variation from merging two north-eastern German Zones to one. From seven to six zones, a Dutch-Belgium price zone is merged with a small Western-German price zone.

Finally, the optimal result for two zones presents a (almost only) French and a unified zone of the other remaining countries. Thereby the French zone would have the lower average price.

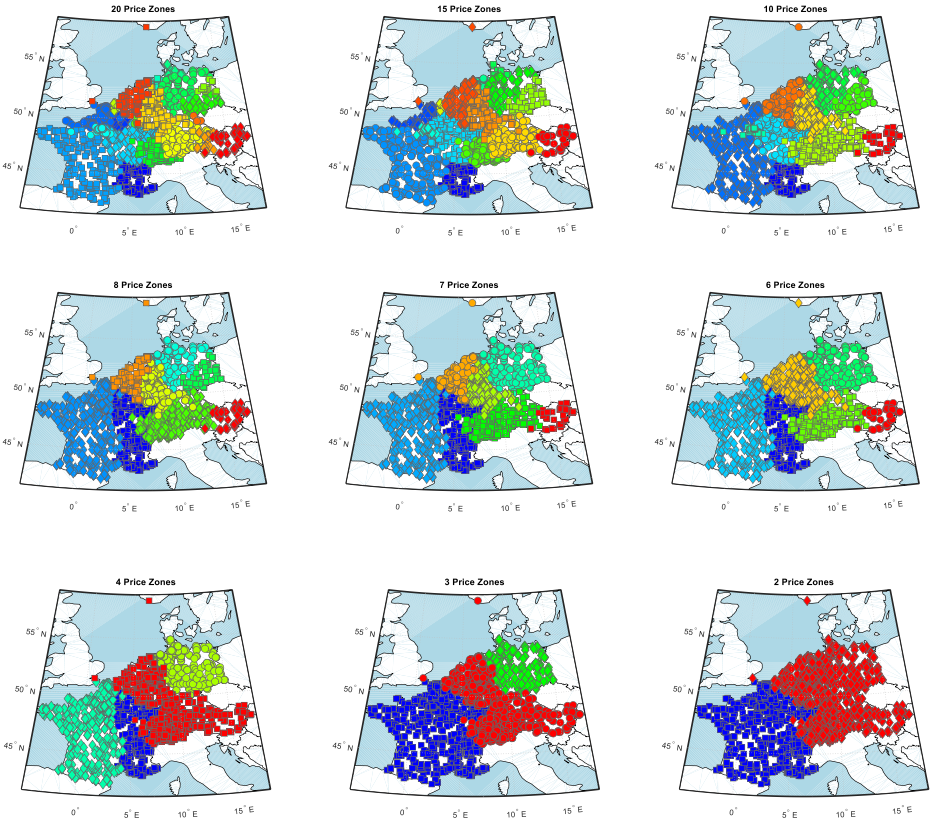

Figure 8: Robust Price Zone Configurations for various number of zones

4.4.2 Development of within & between zone variations

For each of the presented configurations, the share of within- and between variations is computed. Figure 9 shows the results depending on the total amount of price zones. The x-axis shows the number of price zones in a log-scale. The y-axis indicates the share of within- and between-zone variations in percent. When each node is a price zone, all variations are located between the nodes (zones). For the case of one big price zone, all variations are located within the zone, respectively. For the presented case of five price zones around 20.5% of the variations are within the zones and 79.5% between them.

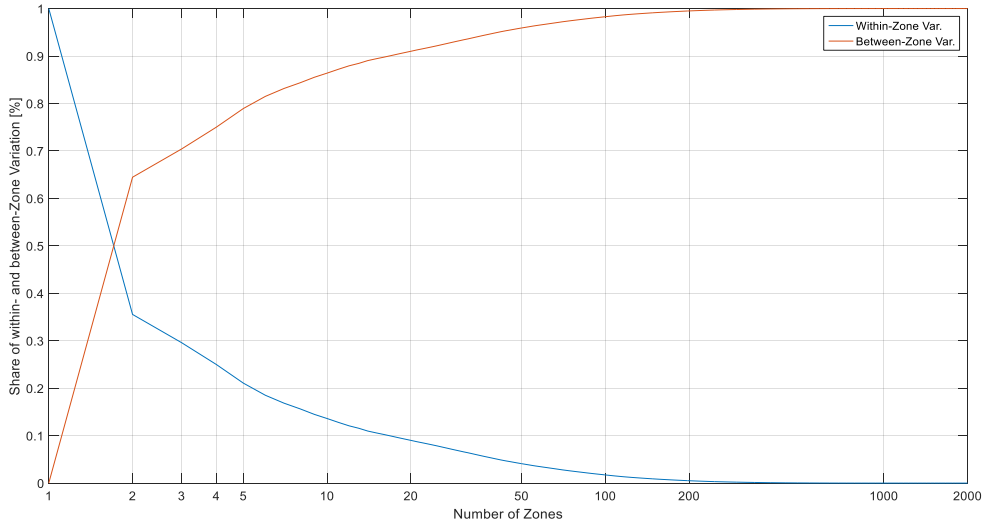

Figure 9: Within-and between-zone variation of the robust configuration depending in number of zones

To ease the interpretation of these values, the following section compares these parameters to the variations in the current CWE configuration. In addition, further benefits of the robust configuration in comparison to the current CWE configuration are presented.

4.5 Benefits of the robust configuration

Whereas the previous discussion has mostly focused on the qualitative and graphical interpretation of different price zone configurations, the focus is now on assessing quantitatively the merits of a robust configuration. In line with Figure 9 the share of within-zone variation of each scenario is compared to the robust and the current CWE configuration. In addition, the size of price zones, in particular the number of nodes per zone, is assessed and analysed. Finally, the within-zone variation that results from applying each of the eight configurations (six scenario configurations, the current CWE and the robust configuration) in each of the single scenarios, is presented.

4.5.1 Comparison to current CWE configuration and scenario optimized configurations

Figure 10 presents a comparison of the within-zone variation of each scenario configuration to the robust and CWE-configuration in each particular scenario. E.g. the top left bar chart shows the share of within-zone price variations for the optimal 5-zone-configuration of scenario one. The second bar shows the share of within-zone price variations that would occur when the robust configuration would be applied in scenario 1. The third bar represents the resulting within-zone share when the current CWE-configuration is used in the first scenario. In each of the six bar charts, i.e. under all scenarios, the CWE-configuration is the worst choice. The robust configuration does not beat the configuration optimized for a given scenario - this is impossible by virtue of the solved optimization problem. Yet its share of within-zone

price variations is throughout only slightly higher than the share for the scenario-specific optimal configuration. Throughout, the robust configuration is moreover performing substantially better than the current CWE-configuration.

Noticeable is scenario 4, where overall the lowest shares of within-zone variations occur. This scenario refers to the “system security” scenario. Both the reduced infeed from renewables and the grid development as scheduled without delays lead to less volatile prices and less congestions, respectively.

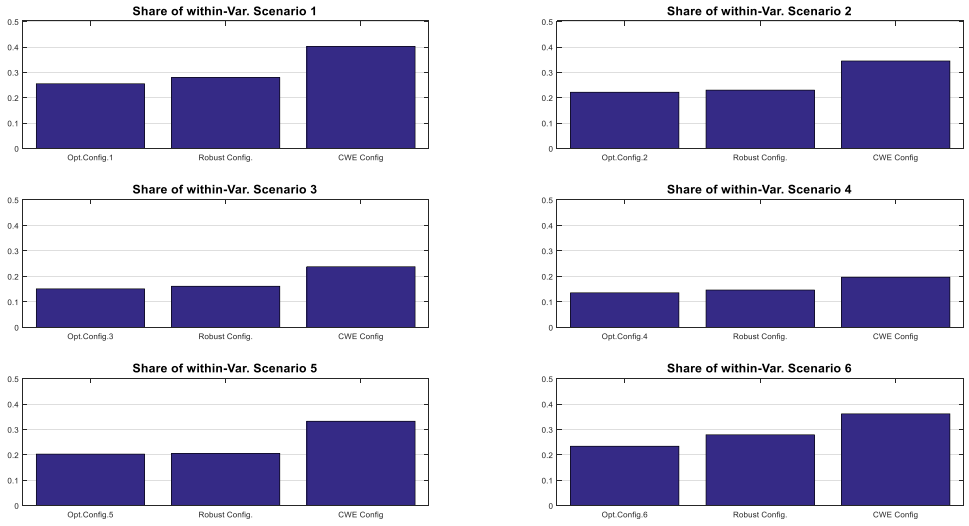

Figure 10: Comparison of within-zone variation between configurations

4.5.2 Size of new price zones

In addition to the benefits regarding the within-zone variation, this section investigates the resulting size of price zones. Thereby size refers to the weight of a zone in line with equation (3). Figure 11 visualizes the share of the zonal weight (W_c) in each of the five zones ordered by increasing weight. The fifth and largest zone is on the right and the smallest zone, the first zone, on the left. The y-axis indicates how many nodes, in relation to all nodes in the system, are located in each particular price zone. The single bars represent the six configurations of the scenarios (cf. Figure 6) whereas the horizontal grey line shows the sizes of the current CWE configuration and the red line the sizes of the robust configuration.

The size of zones in the current CWE configuration seems more uneven in comparison to the other configurations. The two big French and German-Austrian-Luxembourgian price zones sum up to around 82% of total CWE size. The scenario-specific configurations mostly avoid such extremely large zones, especially scenario 2 shows rather equally sized zones. Nevertheless, the other five scenario configurations still have a rather wide range of sizes in comparison to the robust configuration. In the robust configuration, the weights of zones are by contrast rather equal. The three small zones are of quite similar size. Eventually, the robust

configuration turns out to level out zone sizes by not giving weight to strong congestions and corresponding deviating LMPs in particular scenarios.

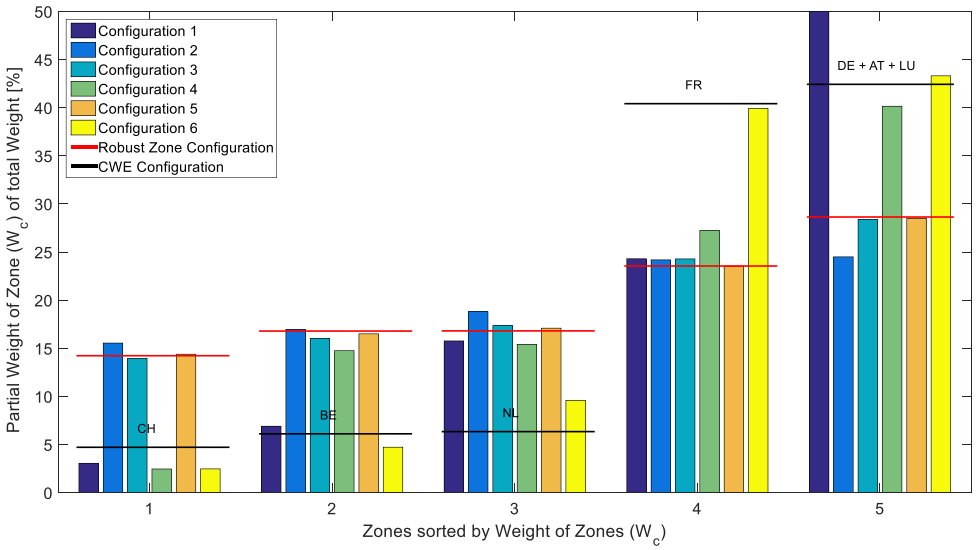

Figure 11: Size of price zones

4.5.3 Evaluation of the benefits of robust configuration

From the previous two figures, we have already concluded that the robust configuration has benefits in terms of more equally sized zones and that its within-zone variation in each scenario is only slightly higher in comparison to the scenario-specific configuration. Figure 12 extends these observations and shows that the robust configuration indeed beats the other configurations both in terms of average within-zone variation and in worst-case performance.

Whereas Figure 10 has compared the within-zone variation of the robust configuration to each scenario-specific configuration, Figure 12 visualizes the resulting total within-variation of all configurations in each scenario. So the robust configuration is no longer only compared to the best case within each scenario, but the more realistic decision situation is represented: The decision maker has to make a single configuration choice now and the performance will only be revealed later. As seen of today, different scenarios are possible and the decision maker wants to compare the merits of his configuration choice against these scenarios. Consequently, the eight retained configuration choices are aligned on the x-axis. These include the six optimal configurations with five zones for each scenario, the robust configuration and the existing CWE configuration. On the y-axis, the performance in terms of within-zone price variations is visible. For each configuration, the stars indicate the within-zone variation that result from the LMPs in all scenarios. The blue bar presents the average of these six values. Note that we choose to present the absolute within-zone price variations here instead of the

shares shown in figure 10. Hence, scenarios with higher overall variations get a higher weight than scenarios with lower overall price variations since they are considered as more critical.

In this figure, we may thus see the consequences of relying on erroneous predictions when choosing a configuration. E.g., configuration 1 would be the best choice, if scenario 1 was anticipated – the dark blue square in column 1 is lower than the dark blue squares in the other columns (other possible configuration choices). Yet the other squares in the first column indicate the consequences if another scenario actually became reality. The worst case would result from a development according to scenario five (yellow square). Then the within-zone-variation would increase by about 50% in comparison to the expected value of scenario one. The best case, as it is in most scenarios, would be scenario three.

In comparison between all eight configurations, the current CWE configuration appears to be the worst choice. This is true both in terms of the average within-zone variation (grey bar) and in terms of the worst case outcome: the highest value of within-variation occurs when the CWE configuration is chosen (or rather maintained) and scenario 5 materializes.

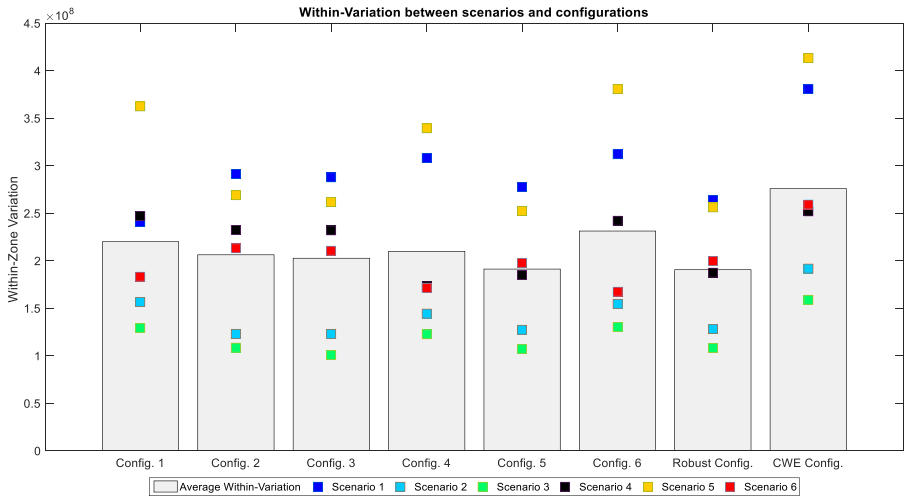

Figure 12: Comparison of within-zone variations over all scenarios and configurations

In contrast to that, the robust configuration exhibits the lowest average within-zone variation. So under the given uncertainties, the robust configuration performs best “on average”. This goes beyond the observations made in Figure 10, where we only could state that it would not perform much worse than the scenario-specific best choice. Pushing further, we can even assert that the robust configuration is robust in the usual minimax sense of the word: the worst case that can occur with the robust configuration corresponds to the square in the seventh column with the highest value, i.e. scenario 1. This worst case is however better than the worst case of any other configuration. The top squares in columns 1 to 6 and 8 rank higher than the

one in column 7.[1] Consequently, the uncertainties of future development are dealt with best by choosing the robust configuration.

5 Conclusions and Outlook

The present paper commenced with a review of relevant literature and the actuality and relevance of the topic. Price zone configurations for European countries are highly relevant and a frequently addressed topic in academic literature and politics. Nevertheless, the impact of uncertainties on price zone configurations has not been addressed in that manner yet.

Therefore, a new and consistent methodology has been developed based on a generalized version of Ward's method. At first, the decomposition of the total variation into a variation within- and between zones is presented. Based on that a hierarchical cluster algorithm that weights nodes according to their infeed and demand situation and has an objective function with a clear economic foundation, the minimization of within-zone price variations, is applied.

Finally, descriptive scenarios are derived in order to investigate the shape of price zones and the effect on within-zone price variations. Based on these scenarios, an approach to compute a robust configuration is presented by applying the same cluster algorithm on the results (LMPs, weights) of all six scenarios.

The results provide interesting insights regarding the shape of price zones and the benefits of a robust configuration. First, zones differ significantly between the six scenarios. Not only grid development has a major impact but also the other parameters like average prices and price spreads. Second, national borders do not align with price zone borders anymore. Third, price variations within price zones can be significantly lowered by reshaping current price zones.

The benefits of a robust configuration become apparent both when comparing to scenario-specific configurations and in comparison to the current CWE configuration. One advantage is the similar size of price zones. The robust configuration has more equally sized zones than most of the scenario configurations, including notably the current CWE configuration – this may help to limit the market power of major players in the electricity market. Also, the within-zone price variations have been investigated. In each scenario, the robust configuration has lower within-zone price variations than the current CWE-configuration. The within-zone

[1] Note that we cannot assert these findings for any possible application of the method. By design, the method ensures that at each step increases in the average within-variations are minimized. Given the underlying quadratic penalty function, the simultaneous application to several scenarios will normally also ensure in the case of the robust configuration that there are no outliers for certain scenarios – hence usually the results are robust in a minimax sense. But this cannot be proven formally – even the global optimality with respect to the minimization of within-variations cannot be proved for the general case. This is generally acknowledged for Ward's method (e.g. (Govaert, 2013), section 7.3.5.4) as for other clustering methods.

variations are only slightly higher than the variations in the scenario-specific configurations. In total, the average variation of prices within a zone is even the lowest of all eight configurations. Finally, the robustness of the configuration is confirmed by the improved worst-case property regarding within-zone variations in different scenarios.

Beyond the presented observations and results, further aspects and improvements may be considered to extend this paper's work. The paper has investigated impacts of price zone configurations for the year 2020 for an extended CWE region. An extension could focus on both an extension of the simulated countries and of the simulated years. In particular, results for more European countries and for years more in the future are of interest to policy makers. Here a key issue is that investment decisions may depend on the zonal configuration. This leads to endogeneity issues as changing power plant characteristics affect the zonal configuration and vice versa, the zonal configuration influences the investment decision.

Finally, a detailed market simulation with the configured price zones could give further insights on the role of price zone configurations and their effect on welfare and redispatch amounts and costs.

References

Batagelj, V., 1988. Generalized Ward and Related Clustering Problems Ward. Classif. Relat. Methods Data Anal. 67–74.

Bergh, K. Van Den, Wijsen, C., Delarue, E., William, D., Bergh, K. Van Den, Wijsen, C., Delarue, E., William, D., 2016. The impact of bidding zone configurations on electricity market outcomes. doi:10.1109/ENERGYCON.2016.7514031

Bertsch, J., Hagspiel, S., Just, L., 2015. Congestion management in power systems - Long-term modeling framework and large-scale application. EWI Work. Pap.

Breuer, C., Moser, A., 2014. Optimized bidding area delimitations and their impact on electricity markets and congestion management, in: International Conference on the European Energy Market, EEM. doi:10.1109/EEM.2014.6861218

Breuer, C., Seeger, N., Moser, A., 2013. Determination of alternative bidding areas based on a full nodal pricing approach. IEEE Power Energy Soc. Gen. Meet. doi:10.1109/PESMG.2013.6672466

Bundesnetzagentur, 2014. ENTWICKLUNGSPLAN.

Bundesnetzagentur, Bundeskartellamt, 2015. Monitoringbericht 2015.

Bundesnetzagentur, Bundeskartellamt, 2014. Monitoringbericht 2014.

Bundesnetzagentur, Bundeskartellamt, 2013. Monitoringbericht 2013.

Bundesnetzagentur, Bundeskartellamt, 2012. Monitoringbericht 2012.

Burstedde, B., 2012. From nodal to zonal pricing: A bottom-up approach to the second-best. 9th Int. Conf. Eur. Energy Mark. EEM 12 1–8. doi:10.1109/EEM.2012.6254665

Ding, F., Fuller, J.D., 2005. Nodal, uniform, or zonal pricing: Distribution of economic surplus. IEEE Trans. Power Syst. 20, 875–882. doi:10.1109/TPWRS.2005.846042

Duthaler, C.L., 2012. A Network- and Performance-based Zonal Configuration Algorithm for Electricity Systems. doi:10.5075/epfl-thesis-5387

Egerer, J., Weibezahn, J., Hermann, H., 2016. Two Price Zones for the German Electricity Market – Market Implications and Distributional Effects. Energy Econ. 59, 365–381. doi:10.1016/j.eneco.2016.08.002

ENTSO-E, 2015. 2015 Scenario outlook & adequacy forecast. Cagnasso, Rosa Lovera, Arianna Natale, Simone Carnevale, Elisabetta Bova, Giovanna Spina, Giulia.

ENTSO-E, 2014. 10-Year Network Development Plan 2014 1–506.

Felling, T., Weber, C., 2016. Identifying price zones using nodal prices and supply & demand

weighted nodes. 2016 IEEE Int. Energy Conf. ENERGYCON 2016. doi:10.1109/ENERGYCON.2016.7514113

Govaert, G., 2013. Data Analysis. John Wiley & Sons.

Härdle, W., Simar, L., 2003. Applied Multivariate Statistical Analysis Applied Multivariate Statistical Analysis, Technometrics. doi:10.1198/tech.2005.s319

Hogan, W., 1992. Contract networks for electricity power transmission. J. Regul. Econ. 242, 211–242.

Hull, J.C., 2015. Optionen, Futures und andere Derivate. Pearson Studium.

Imran, M., Bialek, J.W., 2008. Effectiveness of zonal congestion management in the European electricity market. 2008 IEEE 2nd Int. Power Energy Conf. 16–20. doi:10.1109/PECON.2008.4762432

Kang, C.Q., Chen, Q.X., Lin, W.M., Hong, Y.R., Xia, Q., Chen, Z.X., Wu, Y., Xin, J.B., 2013. Zonal marginal pricing approach based on sequential network partition and congestion contribution identification. Int. J. Electr. Power Energy Syst. 51, 321–328. doi:10.1016/j.ijepes.2013.02.033

Kłos, M., Wawrzyniak, K., Jakubek, M., 2015. Decomposition of power flow used for optimizing zonal configurations of energy market. Int. Conf. Eur. Energy Mark. EEM 2015–Augus. doi:10.1109/EEM.2015.7216779

Klos, M., Wawrzyniak, K., Jakubek, M., Orynczak, G., 2014. The scheme of a novel methodology for zonal division based on power transfer distribution factors. Proceedings, IECON 2014 - 40th Annu. Conf. IEEE Ind. Electron. Soc. 3598–3604. doi:10.1109/IECON.2014.7049033

Lance, G.N., Williams, W.T., 1967. A general theory of classificatory sorting strategies: II. Clustering systems. Comput. J. 10, 271–277. doi:10.1093/comjnl/10.3.271

Meibom, P., Barth, R., Hasche, B., Brand, H., Weber, C., O'Malley, M., 2011. Stochastic optimization model to study the operational impacts of high wind penetrations in Ireland. IEEE Trans. Power Syst. 26, 1367–1379. doi:10.1109/TPWRS.2010.2070848

Murtagh, F., Legendre, P., 2014. Ward ' s Hierarchical Agglomerative Clustering Method : Which Algorithms Implement Ward ' s Criterion ? J. Classif. 31, 274–295. doi:10.1007/s00357-

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

doi:10.1016/j.eneco.2013.09.004

Osinski, P., 2016. Mimeo Vertical Load Model. Chair for Management Science and Energy Economics, University Duisburg-Essen.

Sarfati, M., Hesamzadeh, M.R., Canon, A., 2015. Five Indicators for Assessing Bidding Area Configurations in Zonally-Priced Power Markets 0–4. doi:10.1109/PESGM.2015.7286517

Spiecker, S., Weber, C., 2014. The future of the european electricity system and the impact of fluctuating renewable energy - A scenario analysis. Energy Policy 65, 185–197. doi:10.1016/j.enpol.2013.10.032

Stoft, S., 1997. Transmission pricing zones: simple or complex? Electr. J. 10, 24–31. doi:10.1016/S1040-6190(97)80294-1

Tuohy, A., Meibom, P., Denny, E., O'Malley, M., 2009. Unit commitment for systems with significant wind penetration. IEEE Trans. Power Syst. 24, 592–601. doi:10.1109/TPWRS.2009.2016470

Walton, S., Tabors, R.D., 1996. Zonal transmission pricing: methodology and preliminary results from the WSCC. Electr. J. 9, 34–41. doi:10.1016/S1040-6190(96)80456-8

Wawrzyniak, K., Orynczak, G., Klos, M., Goska, A., Jakubek, M., 2013. Division of the energy market into zones in variable weather conditions using Locational Marginal Prices. IECON Proc. (Industrial Electron. Conf. 2027–2032. doi:10.1109/IECON.2013.6699443

Yang, H., Zhou, R., 2006. Monte Carlo simulation based price zone partitioning considering market uncertainty. 2006 9th Int. Conf. Probabilistic Methods Appl. to Power Syst. PMAPS 1–5. doi:10.1109/PMAPS.2006.360219

Zimmerman, R.D., Murillo-Sanchez, C.E., Thomas, R.J., 2011. Matpower: Steady- State Operations, Planning and Analysis Tools for Power Systems Research and Education. IEEE Trans. Power Syst. 26, 12–19.

Appendix

Calculation of weights:

$$W_n^0 = \sum_H \left(\left| q_{gen,h,n} \right| + \left| q_{dem,h,n} \right| \right) \quad \text{(A21)}$$

$$\overline{W_n^0} = \frac{\sum_n W_n^0}{card\ N} \quad \text{(A22)}$$

$$W_n = \frac{W_n^0}{\overline{W_n^0}} \quad \text{(A23)}$$

Correspondence

M.Sc. Tim Felling

(Corresponding Author)

Tel. +49 201 183-5728

Fax +49 201 183-2703

E-Mail tim.felling@.uni-due.de

Prof. Dr. Christoph Weber

Christoph.weber@uni-due.de

House of Energy Markets and Finance

University of Duisburg-Essen, Germany

Universitätsstr. 12, 45117

E-Mail web.hemf@wiwi.uni-due.de

Web www.hemf.net