Sustainable Economy Research Group (S.E.R.G.)
@ CentraleSupélec Paris-Saclay University
Industrial Engineering Department (LGI)

https://www.serg.paris/

WORKING PAPER SERIES N°1

« Regulating TSO Interaction in Bid Filtering for European Balancing Markets »

by

Marie Girod, Viktor Terrier, Virgine Dussartre, Jean-Yves Boiurnaud, Yannick Perez and Benjamin Donnot

03-2025 WP

Regulating TSO interaction in bid filtering for European balancing markets

Marie Girod[1,2,*], Viktor Terrier[1], Virginie Dussartre[1], Jean-Yves Bourmaud[1], Yannick Perez[2], and Benjamin Donnot[1]

[1] *RTE, 7C Place du Dme, 92073 Paris, La Dfense, France*
[2] *Laboratoire de Gnie Industriel, CentraleSuplec, Universit Paris-Saclay, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette, France*
[*] *Corresponding author, e-mail address: marie.girod@rte-france.com*

Abstract

Europe is undertaking projects for near real-time common balancing markets to meet the flexibility needs induced by renewable deployment. A new congestion management method, bid filtering, has been authorized by regulation to prevent unsolvable last minute congestion. It is designed to manage internal congestion and is performed by each Transmission System Operator (TSO) separately without knowledge of bids in other zones. Bids from all zones are shared in the same market, which means filtering from one TSO could affect welfare in other zones, depending on its objective and on regulation. This paper evaluates the potential effects of multiple TSOs interacting with different filtering strategies. Three TSO strategies are considered - Benevolent, Local, and Conservative - and different combinations are tested using multi-agent reinforcement learning. Results show that although several TSOs filtering benevolently leads to the highest net Social Welfare, it is unlikely that all TSOs will adopt this strategy considering political and social constraints in EU27 countries. We discuss several regulatory options to create the conditions for a Social Welfare-maximizing filtering and foster coordination between TSOs.

Keywords— Electricity networks, balancing markets, congestion management, European integration, filtering, multi-agent reinforcement learning

Highlights

- Real-time European balancing market are a new source of congestion
- Filtering is a congestion management method that screens bids pre-market
- Filtering is performed separately by TSOs but results are shared in the same market
- The interaction and different strategies of TSOs within filtering are studied
- Regulation is advised to foster coordination and ensure Social Welfare-maximization

1 Introduction

Integrating electricity markets and managing congestion are two of the main challenges in Europe to successfully deploy renewables and decarbonize the energy sector (International Energy Agency, 2024). The next steps in market integration are the current projects for common balancing markets, which will help manage renewables and new flexibility technologies near real-time. These markets will be run on the European scale, a few minutes before real-time. Balancing markets and network management will take place almost simultaneously, which has never been done before on so large a scale. This technical exploit may thus create new constraints in terms of congestion management.

The European projects for manual balancing markets are TERRE[1] for Replacement Reserve (RR) and MARI[2] for manual Frequency Restoration Reserve (mFRR). In these markets, bids and TSO needs from participating zones are shared and selected according to a common merit order list. TERRE went live in January 2020, and 6 TSOs have connected since (ENTSO-E, 2023). The technical go-live of MARI was in September 2022, but most TSOs are expected to connect in 2024 (ENTSO-E, 2020).

A new congestion management method, bid filtering, was introduced by regulation to face this unprecedented situation. It enables Transmission System Operators (TSOs) to remove some bids from the market if they are expected to create congestion. The method aims to avoid major unresolved congestion post-market, as, given the short timing, no levers may be left to solve it. Although the market clearing is common across Europe, filtering is designed to manage internal constraints, and performed by each TSO separately. TSOs are allowed to filter bids only in their zone, with no information about bids in other zones.

Several methodologies for filtering in the context of European balancing markets are described in the literature. (Guntermann et al., 2018) add bids one by one in the merit order to a network forecast and test if the network remains secure. If it does not remain secure, the bid is removed from the merit order. A similar method was used as a baseline in (Girod et al., 2024). (Doorman et al., 2022) run a DC Optimal Power Flow over the zone with all submitted bids and exchange forecasts: if a bid is activated outside of the merit order, it helps reduce congestion and should be advantaged; if a bid is not activated although it should have been given the merit order, it increases congestion and should be disadvantaged. (Papavasiliou et al., 2020) present an alternative to filtering, where TSOs run Optimal Power Flows to create a residual function for the zone: the function replaces individual bids and defines a price per activated volume in the zone. In (Girod et al., 2024), a new, continuous method for filtering was presented and evaluated. It adds a price delta to bid prices, reflecting the cost of physical delivery to advantage/disadvantage bids that reduce/increase congestion. The price deltas are determined using reinforcement learning (RL). (Girod et al., 2024) showed on a year of data on an updated IEEE-96 network that this filtering method increases net Social Welfare compared to a baseline filtering or no filtering and helped avoid load shedding. It is therefore the filtering method applied in this paper, and hereafter named 'Proposed filtering'.

Even if it is zone-specific and performed by each TSO separately, zones are not isolated and filtering necessarily interacts with the rest of the balancing process. As bids modified by filtering are shared in a common market, filtering in one zone may heavily impact the merit order list and the resulting dispatch for all zones, thereby affecting both their market welfare and congestion management costs. Furthermore, if several TSOs filter, these processes interact: a bid in one zone may for example have more value than initially expected and may not be worth filtering if the neighboring TSO has heavily filtered at that hour, leading to less liquidity on the market and potentially lower congestion risk. The interaction between

[1] Trans European Replacement Reserves Exchange
[2] Manually Activated Reserves Initiative

TSOs in the case of filtering may be more pronounced with Proposed filtering as TSOs are given more freedom - they have a continuous set of actions and can advantage a bid as well as disadvantage it. Several TSOs filtering separately could thereby lead to complementary or conflicting positions. This could be reinforced by different filtering strategies: some TSOs may have other objectives than maximizing Social Welfare over all zones. To the best of our knowledge, the interaction and strategies of TSOs in filtering has not yet been studied.

(Girod et al., 2024) studied a case with one TSO filtering, and the TSO aimed to maximize net Social Welfare over all zones. In this paper, we evaluate the impact of several TSOs filtering and with different objectives. We evaluate whether or not several TSOs filtering is also beneficial for the global system, for separate zones and for individual TSOs. We also determine which conditions are more favorable and if regulation is necessary to bring them about.

We set up a simulation where several TSOs apply Proposed filtering separately and results are shared in the same balancing market. Multi-agent RL is used to train separate TSOs that impact the same simulation environment. In practice, this entails that TSOs can learn to react to strategies employed by other TSOs. Three TSO strategies are defined: a Benevolent TSO that aims to maximize global Social Welfare, a Local TSO that aims to maximize Social Welfare in its zone, and a Conservative TSO that aims to minimize its costs. Different combinations of strategies are tested on the simulation environment.

There are three contributions in this paper: (i) we define three objective functions for TSOs; (ii) we simulate several TSOs filtering simultaneously using multi-agent RL and test different combinations of TSO strategies; (iii) given the results, we recommend regulating filtering to create conditions for TSOs to maximize Social Welfare and favour TSO coordination and discuss different solutions.

This paper is organized as follows: section 2 describes the methodology, section 3 introduces the case study, section 4 presents and discusses results, and section 5 concludes and details policy implications.

2 Methodology

In this paper, we analyze the impact of several TSOs filtering, with different objectives. Filtering is performed separately by TSOs in their zone: each TSO can only filter bids in its zone and does not have information about bids in the other zones. TSOs can choose to filter bids in their zone with different strategies. Section 2.1 describes the studied strategies, section 2.2 details the filtering method used and section 2.3 presents the simulation environment. All acronyms and notations are detailed in appendix A.

2.1 Studied strategies

Three TSO strategies are studied: the Benevolent TSO with limited information, the Local TSO and the Conservative TSO. The strategies are detailed respectively in sections 2.1.1, 2.1.2 and 2.1.3. The Benevolent TSO aims to maximize net Social Welfare over all zones in the market, the Local TSO maximizes local Social Welfare i.e. net Social Welfare in its own zone, and the Conservative TSO only minimizes its costs. A fourth strategy, Do Nothing, is added as a baseline. In this case, the TSO does not filter. Table 1 summarizes the objective and reward given to the RL agent for different TSO strategies.

2.1.1 Benevolent TSO with limited information

Net Social Welfare is defined here as market welfare minus congestion management costs. This metric reflects the balance between finding the dispatch that best satisfies market players and maintaining

Table 1: Studied variants

	Do Nothing	Benevolent TSO	Local TSO	Conservative TSO
Objective	No filtering	Maximize net Social Welfare over all zones	Maximize net Social Welfare in own zone	Minimize TSO costs
Reward	NA	Market welfare - congestion management costs	Local market welfare - local congestion management costs	- (TSO balancing costs + compensation + local congestion management costs)

security, considering remedial actions to solve potential congestion and load shedding in the worst case. Market welfare is computed without TSO demand as the demand price is an arbitrary value[3] that has a strong impact on numerical results but no impact on variant ranking. Congestion management costs are the sum of redispatch costs and load shedding costs. Because demand is not included, the only varying parameter in net Social Welfare is production cost (including redispatch and load shedding costs). Maximizing net Social Welfare in this case is therefore equivalent to minimizing total costs. The objective function of the Benevolent TSO is the same as in (Girod et al., 2024) and can be written as:

$$\max_{s_0 \to a_1 \to s_1 \to a_2 \to \ldots \to a_{N-1} \to s_N} \sum_{t=1}^{N} nSW_t = f(s_t, a_t) \tag{1}$$

with

$$f(s_t, a_t) = M(Q_t, P_t + a_t, ATC_t) - SA(M(t, a_t), L_t, MC). \tag{2}$$

Where for time step i, nSW_i is the net Social Welfare, s_i is the network state, $a_i = (a_0, \ldots, a_b)$ is the filtering action[4] a price delta for all b bids emitted in all zones at $t = i$. If bid j is not in the zone of the considered TSO, $a_j = 0$. In the RL framework, a state leads to an action, which leads to a new state etc. The RL agent gets an observation - a snapshot of the state - in order to take its action.

$Q_i = (q_0, \ldots, q_b)$ are the volumes offered and $P_i = (p_0, \ldots, p_b)$ are the bid prices for all b bids emitted in all zones, $ATC_i = (atc_0, \ldots, atc_{2z})$ are cross-border capacities between z borders and $L_i = (l_0, \ldots, l_n)$ is the load at n nodes of the network. N is the number of hours in a year and $MC = (mc_0, \ldots, mc_g)$ are marginal costs for the g generators of the network.

$M(t)$ is the market outcome, which consists of market welfare, market price and production levels of generators. $SA(t)$ is the security analysis outcome. It depends on the production level of generators determined by the market.

Initial prices P_t are used to compute net Social Welfare as we want to represent actual generator costs, reflected in initial prices, and not financial transactions, which is what modified prices $P_t + a_t$ represent.

Market welfare can be written as

$$M(t, a(t)) = -\sum_{b \in B} q_b^* \times p_b \times \sigma_b \tag{3}$$

[3] Demand is always met - if not by the market, by the security analysis (with load shedding in the worst case) so adding TSO demand welfare in the market welfare computation would only shift results by a significant yet irrelevant constant.
[4] In the interest of clarity, the choice has been made not to display the time dependence on variables internal to sets.

with B the set of bids emitted in all zones, q_b^* the accepted volume of bid b and σ_b its sign: 1 for an upward bid, -1 for a downward bid. Here, q_b^* depends on the filtering action a_t, which modified the position of bids in the merit order.

2.1.2 Local TSO

The Local TSO aims to maximize local welfare, defined here as local market welfare minus local congestion managements costs. Local market welfare is the welfare of balancing bids accepted in the zone, as well as the demand of the local TSO. TSO demand is not included in net Social Welfare but is in local Social Welfare because the local system is not closed: TSO demand may be met by a bid or redispatch from another zone that would not be counted in local market welfare otherwise. This point carries substantial weight and its impact is discussed in section 4.1. Local congestion management costs are the sum of redispatch and load shedding costs in the zone. Thus, the objective of the Local TSO can be written as:

$$\max_{s_0 \to a_1 \to s_1 \to \ldots \to s_N} \sum_{t=1}^{N} lW_t = g(s_t, a_t) \tag{4}$$

with

$$g(s_t, a_t) = lM_z(M(t, a_t)) - lSA_z(M(t, a_t), L_t, MC) \tag{5}$$

and

$$lM_z(t) = \sum_{b \in B_z} (-\sigma_b \times q_b^* \times (p_b - C_z^*)) - \sigma_{TSO} \times q_{TSO}^* \times (p_{TSO} - C_z^*) \tag{6}$$

For time step i, lW_t is the local welfare, $lM_z(i)$ is the local market welfare and $lSA_z(i)$ are the local congestion management costs. B_z is the set of accepted bids accepted in zone z. C_z^* is the market clearing price in zone z. σ_{TSO}, q_{TSO}^* and p_{TSO} are respectively the direction (same as for Balancing Service Provider bids), the accepted volume and the price of TSO demand in zone z.

2.1.3 Conservative TSO

The Conservative TSO aims only to minimize its own costs. TSO costs consist in balancing costs i.e. the price paid on the market for TSO demand, local congestion management costs and filtering compensation costs.

In this paper, filtering compensation takes place when a bid is advantaged by filtering and activated by the balancing market at a loss[5]: the market price does not allow it to recover its costs. The TSO should compensate the difference between the initial price and the market price. Figure 1 illustrates the compensation mechanism for an upward bid. The bid in blue has been advantaged and its price has been updated to $p_b + a_t(b)$ with $a_t(b) < 0$. The market clears at a price C_z^* lower than the advantaged bid's initial price p_b. The TSO compensates the part of the bid that is not covered by the market $q_b^* \times (p_b - C_z^*)$. The same applies for an advantaged downward bid if its initial price is lower than the market price. Compensation is counted in market welfare; it does not appear in other TSO strategies

[5]Other compensation schemes were possible. For example, bids that should have been accepted by the market and were not because they were disadvantaged by filtering could also be compensated. The scheme chosen in this paper matches current filtering practices.

because it is a financial transaction and other strategies consider net Social Welfare, which does not reflect internal financial transactions.

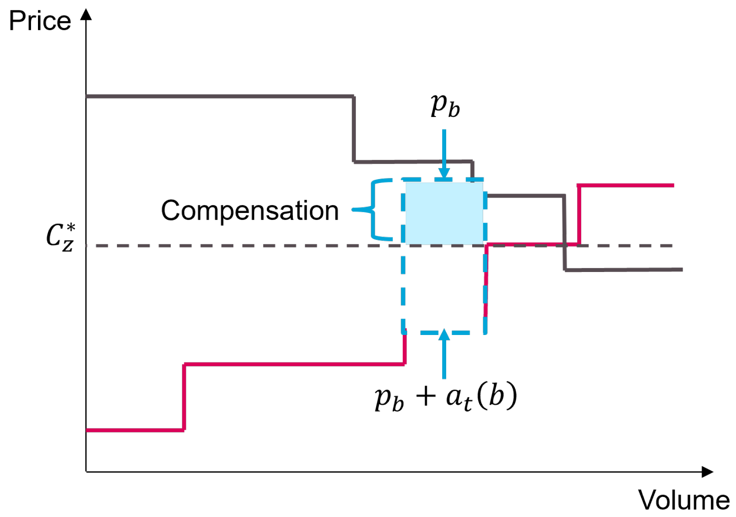


Figure 1: TSO compensation mechanism

The objective function of the Conservative TSO can be written as:

$$\max_{s_0 \to a_1 \to s_1 \to \ldots \to s_N} \sum_{t=1}^{N} cW_t = -h(s_t, a_t) \tag{7}$$

with

$$h(s_t, a_t) = BC_z(M(t, a_t)) + Comp_z(M(t, a_t), Q_t, P_t) + lSA_z(M(t, a_t), L_t, MC) \tag{8}$$

At time step i, $BC_z(i)$ and $Comp_z(i)$ are respectively the TSO balancing costs and compensation costs in zone z. They can be written as

$$BC_z(M(t, a_t)) = -\sigma_{TSO} \times q^*_{TSO} \times C^*_z \tag{9}$$

$$Comp_z(M(t, a_t), Q_t, P_t) = \sum_{b \in B_z} q^*_b \times \delta_b \tag{10}$$

δ_b is a variable that indicates the compensation price if the bid was advantaged at a loss:

$$\delta_b = max(0, \sigma_b \times (p_b - C^*_z)) \tag{11}$$

Although the TSO compensation formula 10 is the one used in the results analysis to compute real compensation costs, in the objective function of the Conservative RL agent, a slightly different formula was used, as described below:

$$Comp'_z(M(t, a_t), Q_t, P_t) = \sum_{b \in B_z} q^*_b \times \delta'_b \tag{12}$$

and

$$\delta'_b = max\Big(0, \sigma_b \times (-a_t(b))\Big) \tag{13}$$

In this formulation, the whole price delta is compensated, not just the difference with the market price. It encourages the Conservative TSO to reduce the price deltas that advantage bids.

2.2 Filtering methodology

2.2.1 Proposed filtering

In this paper, the filtering method used is the one presented in (Girod et al., 2024). The action of the filtering method is to add a price delta $a = (a_0, \ldots, a_b)$ to bid prices $P_t = (p_0, \ldots, p_b)$ in order to advantage/disadvantage bids that decrease/increase congestion. The price delta aims to represent the congestion management costs linked to the activation of the bid. As each TSO can only apply filtering in its own zone, if bid j is not in the zone where filtering is applied, $a_j = 0$. The balancing market clearing is then run with these new prices, which has an impact both on bid activation and market prices. The security analysis is run with the modified market outcome.

2.2.2 Multi-agent modeling

In (Girod et al., 2024), only one TSO filtered in its zone. In this paper, several TSOs filter separately in their own zone. Each TSO i applies a set of price deltas $a_i = (a_{i,0}, \ldots, a_{i,b})$ to bid prices P_t with $a_{i,j} = 0$ if bid j is not in the zone of TSO i. Although TSOs determine the price deltas for their zone separately, they are all applied to the same simulation environment: the same market clears with all sets of price deltas.

2.2.3 Resolution with reinforcement learning

RL is used in order to determine the price deltas. RL is a section of non-supervised machine learning where an agent performs an action, which has an impact on an environment, and the agent is rewarded accordingly to this impact. Here, more specifically, multi-agent RL is used to simulate several TSOs acting separately on the same environment.

Multi-agent RL has been used in power systems extensively with different objectives. A few examples follow. (Daneshfar and Bevrani, 2010) apply multi-agent RL for load-frequency control. (Du et al., 2021) approximate a Nash equilibrium in day-ahead electricity market bidding.(May and Huang, 2023) simulate a peer-to-peer community of prosumers.(Harder et al., 2023) apply multi-agent RL to model electricity markets with 145 revenue-maximizing participants.

In this framework, the n agents' actions $(a_1, \ldots, a_n)$ are applied to an environment, which returns n rewards $(r_1, \ldots, r_n)$ linked to the actions of all agents on the environment. Figure 2 summarizes the RL framework and information flows in a case with two agents.

An RL agent's observation space is the set of values it can access, and its action space is the set of actions. The observation space and action space are distinct for each RL agent. Here, an agent's observation space is made up of two sets of variables: flows on all lines for a predicted network (productions set to the output of the day-ahead market, load and renewable forecasts updated with H-1 predictions - the same predictions as in the balancing market and security analysis, distributed slack); maximum power of upward and downward bids for each generator in the zone where the agent filters. Predictions of flows on all lines are common to all agents; bid information is different for each one. The Grid2op framework described in (Marot et al., 2021) is used for network computation and observations.

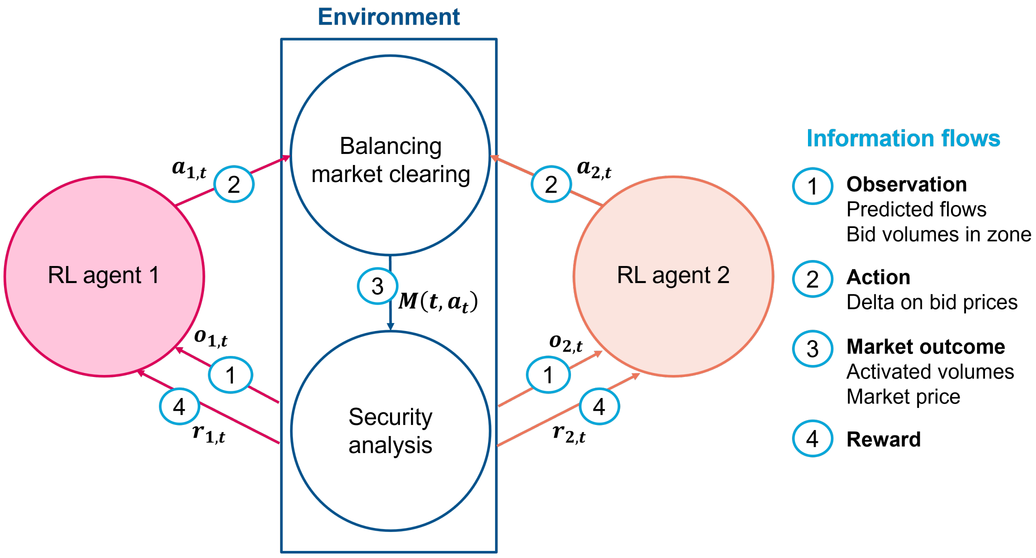


Figure 2: Multi-agent Reinforcement Learning framework

An RL agent's action space is a different price delta for each generator in his zone, both for an upward bid and a downward bid. If the generator has not submitted a bid at that time, the bid's maximum power is set to zero. The action space is restricted by an upward and downward limit, set slightly higher and lower than the TSO's upward and downward demand price respectively. This limit enables the agent to remove a bid from the market by setting a delta too high for the bid to be met by TSO demand.

Each RL agent receives its reward at the end of a week of training. Its objective is to learn to maximize its reward. The reward depends on its strategy (Benevolent, Local or Conservative, described in section 2.1) and its zone: if the TSO is a Local TSO in zone x, it will receive the sum of Local Welfare in zone x over the week. The reward also indirectly depends on the filtering actions of other TSOs, as they impact both the same balancing market and security analysis. Therefore an agent will learn the filtering scheme that maximizes its objective function given the other agents' strategies.

The RL-algorithm used is Soft Actor-Critic, a state-of-the-art deep RL algorithm. It outperforms other classic algorithms (Haarnoja et al., 2019) for multi-agent RL with continuous action and observation spaces. It was also the closest algorithm to the one used in (Girod et al., 2024) that worked for multi-agent. Using a similar algorithm should help recover comparable results. Soft Actor-Critic is particularly adapted to the model, with its large and continuous action and observation spaces. Parameters such as learning rate, neural network architecture and fragment length were tuned. The same parameters were used in all variants to avoid advantaging one over the other. Variants were trained on 900 000 iterations over one year of data at an hourly time step, at which point they had all converged. The converged model was used to evaluate results (last saved checkpoint). In future works, other RL algorithms could be tested and parameters further refined.

2.3 Evaluation framework

All variants are evaluated in the simulation environment presented in Figure 3. First, each TSO can filter bids in its area, given a network forecast and according to its own filtering strategy. Then they send the bids from their area and their own demand to the market, which clears considering the cross-border capacities, and returns market welfare and price. After the market, the network is updated with volumes sold. The security analysis runs considering the updated network, solves existing congestion and computes congestion management costs.

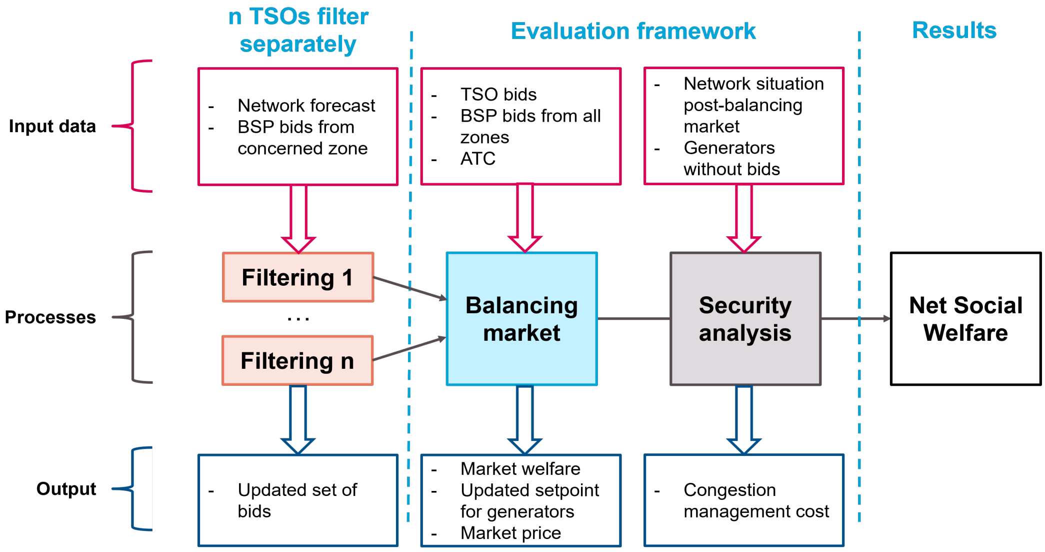


Figure 3: Evaluation framework

The balancing market and security analysis are the same as in (Girod et al., 2024). Balancing Service Providers offer all their available power upward and downward at marginal cost considering their technical constraints (gradient, minimum on/off time and minimum stable time for thermal generators). TSO demand is the imbalance in their zone, formulated at all costs. The zonal market clearing maximizes market welfare considering Available Transfer Capacity constraints. Market prices are the price of the marginal bid, using modified - not initial - bid prices, and considering ATC constraints. This pricing method was chosen as it is coherent with the market dispatch.

The security analysis simulates the actions of an operator at the balancing market outcome: it aims to restore network security at minimal cost. It is modeled here using a security-constrained Optimal Power Flow. The objective function and several constraints of the security-constrained Optimal Power Flow are presented in (Girod et al., 2024).

3 Case study

3.1 Network

The model was applied to the RTS GMLC (Barrows et al., 2020), a modified IEEE-96 network with a high share of renewables. The RTS GMLC includes one year of load and renewable production data, as well as H-24 forecasts. The network is slightly simplified in this paper: storage is removed, minimum power constraints for generators are not considered and only marginal cost is included. Load shedding cost is 33 k/MWh, the value used by the French regulator (Commission de regulation de l'Energie, 2022). Contingencies for N-1 states are a subset of 17 lines selected using a sensitivity analysis.

A limit in feasible volume of redispatch was included in the security analysis and evaluated in (Girod et al., 2024) to represent operational constraints near real-time. This limit was previously set to 200MW of feasible redispatch. The lower the limit is set, the more constrained the network. A sensitivity analysis showed that in a constrained network, the RL agent became efficient in critical situations and neglected the situations with lower congestion management costs. The main objective of (Girod et al., 2024) was to assess the capacity of filtering to resolve insecure situations, and therefore needed a range of insecure cases. In this paper, we focus on the impact of different TSO strategies on the market and the welfare of different zones and wanted to test the RL agent on more varied situations. We chose a less congested network, and set the limit in feasible volume of redispatch to 300MW.

3.2 Data generation

A balancing market aims to resolve an imbalance. The imbalance is created here with a day-ahead market simulation using the ATLAS model (Little et al., 2024). The market clearing, simulating DA and ID jointly, is run with flow-based constraints. The flow-based computation is detailed in (Girod et al., 2022). Renewable production and load forecasts are updated at different stages of the simulation and create imbalances. In order for the simulated day-ahead market to fill both the roles of the day-ahead and intraday markets, uncertainties are set to H-2 forecasts instead of the actual H-24 forecast for day-ahead bids. Flow-based computation, usually run with H-48 forecasts, is run here with H-24 forecasts for coherence.

Balancing bids are also computed in the ATLAS model (Cogen et al., 2024). Balancing bids are computed with H-1 forecasts. Generators bid all their available capacity considering technical constraints for thermal generators: gradient, minimum on/off time and minimum stable time constraints. Generators bid at marginal cost. In particular, Renewables bid at 0.1€ for upward bids and -0.1€ for downward bids to avoid unrealistic netting. TSO demand is the hourly zonal imbalance, computed considering exchanges from the day-ahead market. TSOs bid at "all costs" - in this case 300€ /MWh, which is higher than the most expensive generator's marginal cost at 180€ /MWh. The balancing market clears with ATC constraints. The balancing ATC is derived from the day-ahead flow-based domain using the operational method (Creos, Tennet, Amprion, RTE, Transnet BW, Elia, 50Hertz, APG, 2020) and is thereby based on H-24 forecasts.

The RL models are trained and tested on two different data sets with one year of data each, at an hourly time step[6]. In operational conditions, training would necessarily take place on a different data set than the one that is observed in real time. The reference situation and balancing bids are created using the method described above in both cases, but two changes are applied to the network for the training set in order to create different but consistent data sets. First, the RTS GMLC forecasts were switched: the H-24 forecast was considered as real-time data and real-time data was considered as an H-24 forecast. H-1 and H-2 forecasts were updated accordingly. This approach preserved coherent forecasts and imbalances in the training set. Second, four thermal generators were removed in the training data set: they were very active in the testing set, so removing them had a strong impact on dispatch and created a substantial difference between the two data sets; they were too slow to take part in redispatch, so removing them did not impact the TSO real-time action modeled by the security analysis.

3.3 Zone characteristics

The network includes three interconnected zones. They have a similar network topology but different production and load. Table 2 presents some characteristics. Local congestion management costs and market welfare are computed in the simulation environment described in section 2.3 in the Do Nothing case. Zone 3 is the most impactful: it has the highest installed renewable capacity, highest congestion management costs and highest local balancing market welfare. Zone 2 is the least impactful in all three categories.

[6] The code and data sets are being reviewed and will be published in open access if the paper is accepted

Table 2: Zone characteristics in balancing Do Nothing simulation

	Zone 1	Zone 2	Zone 3
Installed renewable capacity (GW)	1.5	1.0	4.0
Local congestion management costs (M€)	3.9	2.0	27
Local balancing market welfare (M€)	32	14	51

In order to have varied cases, zone 2 and 3 are chosen as the zones where filtering is applied. Zone 1 is chosen to not be part of the filtering process (no filtering is applied to this zone) to limit the number of possible combinations and make it easier to understand what is happening.

3.4 Studied strategy combinations

In all the studied variants, one TSO does not cooperate (Do Nothing objective) and the two others filter separately in their own zone. All three strategies - Benevolent, Local and Conservative - are applied once in both zones and once in combination with a Benevolent TSO, to understand the effects of common and diverging strategies. Table 3 summarizes the studied variants. Each combination is only tested once: the zone 1 TSO does not filter and when strategies diverge, the TSO in zone 3 is the Benevolent one. In future works, the roles between TSOs could be exchanged to test all possible combinations.

Table 3: Studied TSO strategy combinations

Variant	Zone 2	Zone 3	Zone 1
Do Nothing	Do Nothing	Do Nothing	Do Nothing
Two Benevolent TSOs	Benevolent	Benevolent	Do Nothing
Local + Benevolent TSOs	Local	Benevolent	Do Nothing
Two Local TSOs	Local	Local	Do Nothing
Conservative + Benevolent TSOs	Conservative	Benevolent	Do Nothing
Two Conservative TSOs	Conservative	Conservative	Do Nothing

4 Results and Discussion

All variants were evaluated over a year on the test data, which was never seen during the RL training. In this section, first, results for all considered variants are provided, aggregated over the year. Observations on filtering efficiency, impacts on TSO costs, coordination and result coherence are made. The impact of coordination is discussed. Second, time series of total costs and TSO costs in zones 2 and 3 are presented. Third, TSO costs are broken down to outline incentives TSOs are given with the current filtering design.

4.1 General results

Figure 4 presents total costs, local costs and TSO costs for every variant in Table 3. Total costs and local costs are respectively the opposite of net Social Welfare and local welfare. Local costs are negative because they take into account TSO demand, which increases welfare and counts as negative costs. Several observations can be made when considering Figure 4, as detailed in the following subsections.

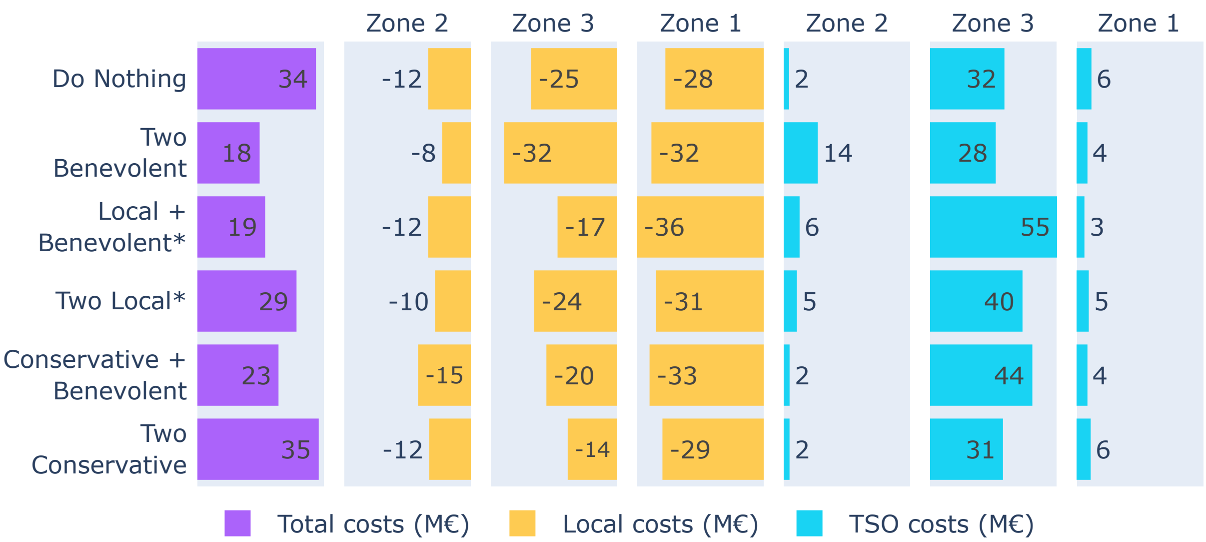


Figure 4: Total costs, local costs and TSO costs by zone for each variant
The variants marked with a * have converged in a local maximum and are discussed below.

4.1.1 Global filtering efficiency

When Two Benevolent TSOs filter, total costs are highly reduced compared to Do Nothing. The Two Benevolent TSOs variant is the most efficient in terms of total cost minimization. More generally, as long as the zone 3 TSO is Benevolent (Two Benevolent, Local + Benevolent, Conservative + Benevolent variants), costs are significantly reduced with filtering. As seen in section 3.3, zone 3 has the highest congestion management costs and market welfare, so it makes sense that filtering in zone 3 has a strong impact.

4.1.2 Impact of strategies on TSO costs

Filtering can be expensive for TSOs. In the Two Benevolent TSOs variant, although total costs are reduced, TSO costs for zone 2 are multiplied by 7. If a TSO only considers its own costs, it is not necessarily in its interest to filter, even if filtering reduces total costs: over all studied variants, TSO costs in zone 2 are either stable or increase with filtering; in zone 3 they either slightly decrease or significantly increase. TSO costs will be further analyzed in section 4.3. The fact that filtering is not necessarily beneficial for individual TSOs is all the more true given that TSO costs in zone 1, where no filtering is applied, are either stable or decrease when the other zones filter. This is probably linked to the case study, but it shows that the free-rider strategy can be profitable.

4.1.3 Impact of coordination

TSO costs depend on other TSOs' strategies: because RL agents are trained together, they learn each other's strategies and adapt their actions accordingly. This probably represents what would happen under actual conditions: TSOs would adapt to each other's actions even if they did not communicate in real time.

If this coordination is not properly regulated, its impact can be negative. For example, when the zone 2 TSO is Conservative and the zone 3 TSO is Benevolent, the Conservative TSO transfers its costs to the Benevolent TSO: zone 3 TSO costs increase to 44M€ compared to 28M€ in the Two Benevolent case, and zone 2 TSO costs decrease to 2M€ compared to 14M€ in the Two Benevolent case.

On the other hand, when both TSOs have a Benevolent strategy, coordination gives them an extra degree of freedom to maximize net Social Welfare. In Appendix B, filtering actions of the Two Benevolent and Local + Benevolent variants are presented on one time step, identified by the arrow in Figure 5. In this example, when both TSOs had a Benevolent strategy, the efficient solution they applied to minimize total costs would not have been possible without coordination. In the Benevolent + Local case, coordination gave the Local TSO an opportunity to profit from the Benevolent TSO.

The importance of inter-TSO coordination in congestion management has already been identified by (Kunz and Zerrahn, 2015; Glachant and Pignon, 2005; Glachant et al., 2017; Bertsch et al., 2016). (Kunz and Zerrahn, 2015) run a General Nash Equilibrium model between four TSOs and show that coordination in redispatch leads to lower congestion management costs (although volumes are similar, coordination leads to the use of cheaper redispatch resources), gains in supply security with more remaining margins on lines and fewer TSO interventions, and a decreased need for network expansion. (Bertsch et al., 2016) analyze different congestion management methods and find that one of the two main sources of inefficiencies compared to nodal pricing is that TSO actions are restricted to zones. (Glachant et al., 2017) identify the lack of coordination in redispatch as one of the roadblocks to the integration and decarbonisation of the European electricity sector. One of the proposed solutions is to share costs and benefits of redispatch among European TSOs with a mechanism that should be fair and provide sound incentives to TSOs. This approach could also be implemented for filtering with an ex-post distribution of compensation and security costs. (Glachant and Pignon, 2005) put forward several solutions to enhance coordination between TSOs in congestion management. They recommend exchanging information, increasing transparency of decisions, improving procedures so that TSO operations and markets are seamlessly connected, merging some TSOs and creating a multi-country regulating body (the article dates from 2005, before the creation of ACER in 2009).

Among these measures, information exchange and improved decision transparency could be applied to filtering. For instance, this could involve sharing filtering results or algorithms, thereby also making the TSO's objective function publicly accessible. Furthermore, TSOs from different zones could train their filtering agent on the same dataset to create coordination. It would be interesting to better quantify the level of coordination achieved in this paper. Future studies could look at a case with no coordination - agents trained separately - and a case with perfect coordination - one agent that filters for all zones.

4.1.4 Result coherence

The Benevolent and Conservative TSOs are coherent with their objective functions. Total costs are the lowest when both TSOs are Benevolent. When the zone 2 TSO is Conservative and zone 3 TSO is Benevolent, zone 2 TSO costs decrease compared to the Benevolent variant, and total costs are lower than when the zone 3 TSO is also Conservative. When both TSOs are Conservative, total costs increase as nobody aims to minimize them, and TSOs find a balance where both their individual costs are low. It seems that in the Conservative variant, the TSO strategies are close to Do Nothing. This point will be further analyzed in section 4.3.

Although the results for the Two Local TSOs are coherent, it seems that the models have converged in a local minimum. When only the zone 2 TSO is Local, its local costs are low, but they are higher than in the Conservative + Benevolent case so the model has not reached the optimum. In the same way, when both TSOs are Local, their local costs are low, but slightly higher than the Do Nothing case. This is likely due to the TSO demand price, which is counted in local costs and not in total costs and has a major impact on results: for Do Nothing, the TSO demand welfare amounts to 66M€ over the year, which is around the double total costs.

The arbitrary value of TSO demand price is driving results for local costs, which has two implications. First, it is likely that the difference in local costs between variants is not significant. Second, it was probably difficult for Local RL agents to train properly: the effect of their actions on their reward were buried by TSO demand welfare and they must have focused on filling TSO demand in the market as it had the most value, even though it was not the desired objective. In order to solve the issue, TSO balancing costs instead of TSO demand welfare could be counted in local costs as they do not take into account TSO demand price.

4.2 Hourly costs

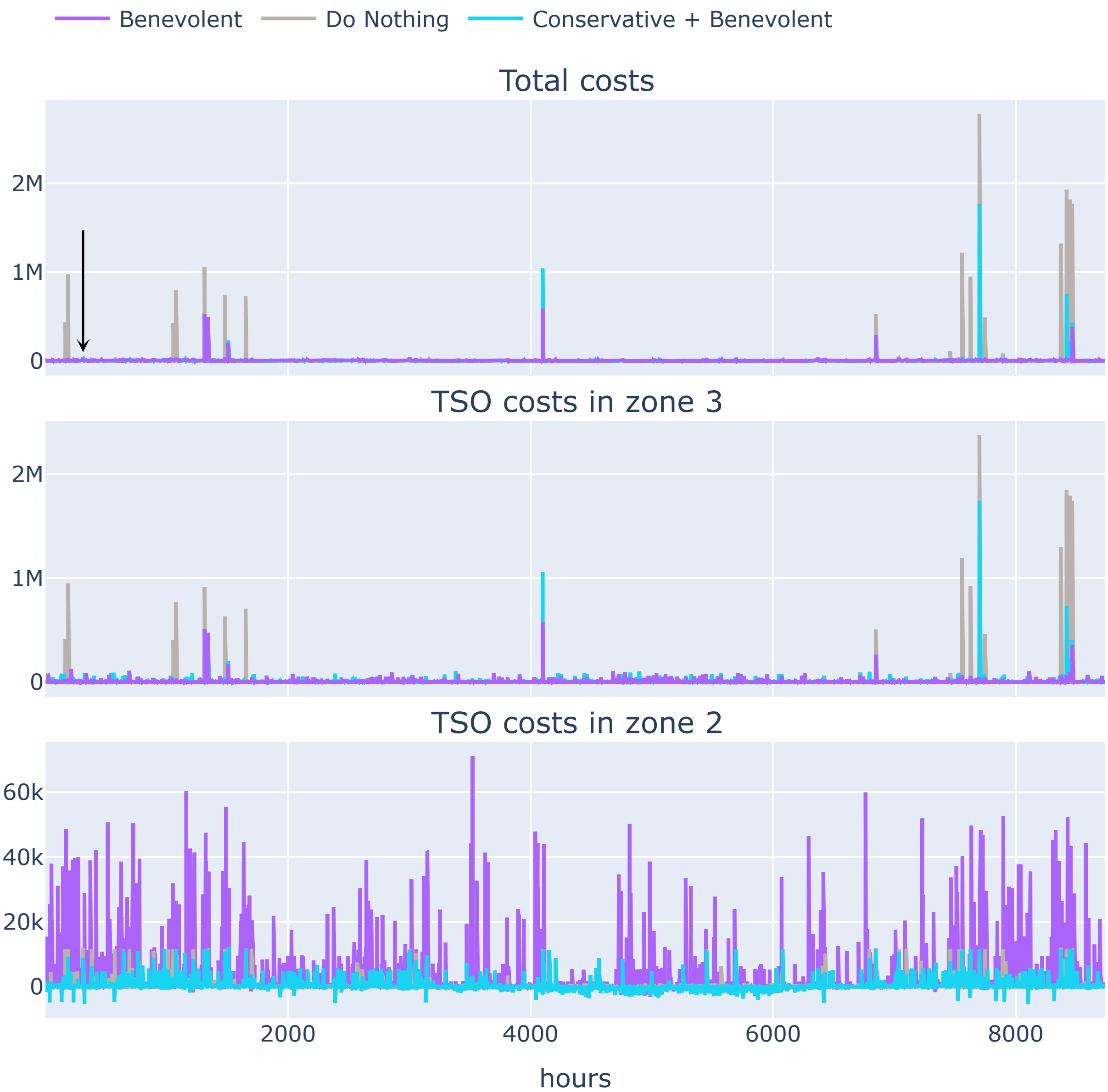


Figure 5: Hourly costs for Do Nothing, Two Benevolent TSOs and Conservative + Benevolent TSOs

Figure 5 presents hourly costs for three variants: Do Nothing, Conservative + Benevolent TSOs and Two Benevolent TSOs. The top graph shows total costs. The hours with peaks in costs are the hours where load shedding occurs. The Two Benevolent TSOs greatly reduce both the number of hours with load shedding and its cost. The Conservative + Benevolent TSOs do help reduce load shedding, but not as often or as efficiently. Appendix C presents total costs broken down by balancing generation costs and congestion management costs. It shows that congestion management costs always decrease with filtering, even if total costs increase.

The middle graph presents TSO costs in zone 3 for the corresponding hours. Over the year, zone 3 TSO costs are 33M€ with Do Nothing, 28M€ with Benevolent TSOs and 44M€ with Conservative + Benevolent TSOs. The TSO costs in hours with load shedding are similar to total costs, which means that in those three variants, the TSO in zone 3 bears most of the congestion management costs during high-stake hours.

TSO costs in zone 3 decrease with filtering in hours with load shedding, but increase in low-stake hours (hours where Do Nothing total costs are smaller than 10k€). Zone 3 TSO costs in low-stake hours amount to 23% of all zone 3 TSO costs in the Do Nothing case, 91% in the Two Benevolent case and 97% in the Conservative + Benevolent case. As discussed in (Girod et al., 2024), filtering could be applied only when the risk of load shedding is higher than a threshold, i.e. in high-stake situations, and most of filtering's value would be retained. Zone 3 TSO costs would be further reduced in the Two Benevolent TSOs variant as low-stake hour costs would decrease to the Do Nothing level. The zone 3 TSO would have a higher incentive to filter Benevolently with this high-stake/low-stake separation. Further refining the RL algorithm used and its parameters in future work could also help reduce unnecessary filtering.

The bottom graph presents TSO costs in zone 2. In zone 2, TSO costs are 2M€ with Do Nothing, 14M€ with Benevolent TSOs and 2M€ with Conservative + Benevolent TSOs. In the Do Nothing and Conservative + Benevolent cases, TSO costs seem very close and regularly distributed across hours. In the Two Benevolent TSOs case, the cost increase is spread out and is not particularly linked to hours with load shedding. Even if the risk threshold was added, the zone 2 TSO costs would increase with the Benevolent strategy. In any case, the zone 2 TSO would not filter Benevolently if it only considers its costs.

4.3 TSO incentives

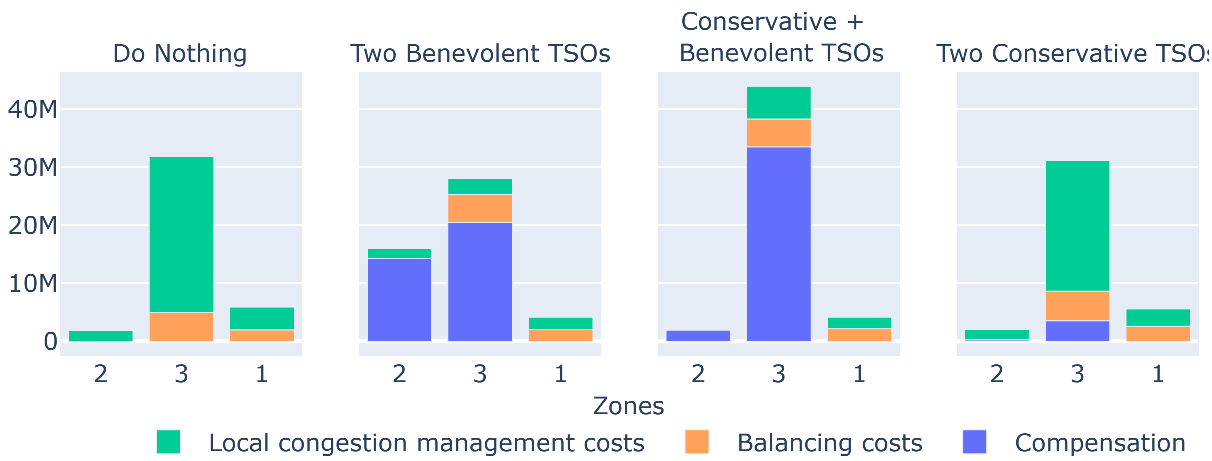


Figure 6: Breakdown of TSO costs for every zone in Do Nothing, Two Benevolent, Conservative + Benevolent and Two Conservative variants

Figure 6 shows a breakdown of TSO costs by local congestion management costs, balancing costs and compensation costs in each zone for four different variants. In the Two Benevolent TSOs variant, TSOs pay large compensation costs to reduce congestion management costs. It may seem counterintuitive that TSO costs increase compared to the Do Nothing variant, even though total costs decrease, as seen in Figure 4. This is because compensation costs, that are a financial transaction between the TSO and producers, are not included in the total costs computation. What does affect total costs, however, are changes in the generation dispatch to take into account network constraints. These changes are due either to Proposed filtering that modifies the merit order, or to congestion management. A solution where a preventive change in the merit order affects the final dispatch less than ex-post congestion management is therefore more efficient, whatever the TSO compensation costs. The TSO compensation costs could be redistributed among market agents later.

In the Conservative + Benevolent variant, the Conservative zone 2 TSO transfers nearly all its compensation and congestion management costs to the Benevolent zone 3 TSO. In the Two Conservative TSOs variant, both TSOs have nearly eliminated their compensation costs and pay congestion management costs instead. TSO costs, as well as total costs in Figure 4 are very close to the Do Nothing variant. Thus, it seems that if both TSOs try to minimize their costs, it is overall more advantageous for them to pay congestion management costs than compensate, even if it is not efficient for the global system.

Given the impact on TSO costs, it seems unlikely that, without regulation, all TSOs will choose the Benevolent strategy. If even only one TSO chooses another strategy, the other TSOs will likely adapt: it is not sustainable for TSOs to stay Benevolent if others are not, as others would transfer their costs to the Benevolent TSOs. As mentioned in section 4, not filtering can also be beneficial for the TSO: in the studied variants, TSO costs either decreased or remained stable in the zone with no filtering.

Furthermore, the TSO has fewer incentives to apply Proposed filtering than other congestion methods that are not necessarily more efficient in terms of Social Welfare. For example, the Baseline filtering method, described in (Girod et al., 2024), removes bids from the market instead of advantaging/disadvantaging them. Hence the TSO does not have any compensation costs. It is likely that some TSOs would have lower costs with Baseline filtering than with Proposed filtering, although (Girod et al., 2024) showed that Proposed filtering was significantly more efficient in terms of net Social Welfare. Similarly, with cross-border capacity, the TSO does not incur any costs and receives part of the congestion rent. Although filtering is not designed to replace cross-border capacity, in some cases applying filtering instead of reducing the cross-border capacity may be more efficient in terms of net Social Welfare but more costly for the TSO.

5 Conclusion and Policy Implications

In this paper, we studied the interaction between Transmission System Operators (TSOs) in the context of filtering for the upcoming common European balancing markets. Filtering enables TSOs to remove bids emitted in their zone if they are expected to create congestion. It is performed by each TSO separately but bids are shared in a common market, which means that one TSO's filtering affects the welfare of other zones. We simulated the effect of several TSOs filtering on a market clearing to determine market welfare, and a subsequent network security analysis, that computes congestion management costs. Social Welfare is considered here as the difference between market welfare and congestion management costs. Three different TSO strategies for filtering were studied: Benevolent, Local and Conservative, which respectively aim to maximize net Social Welfare, maximize local welfare and minimize TSO costs.

Several observations were drawn from the simulation. First, the most efficient variant in terms of net Social Welfare was when both TSOs were Benevolent. Second, two TSOs filtering separately but in a coordinated manner can help increase net Social Welfare. When both TSOs are Benevolent, coordination gives them an extra lever; otherwise it can give a non-Benevolent TSO the opportunity to take advantage of a Benevolent TSO. Third, Benevolent filtering can be expensive for individual TSOs even if it reduces system costs over all zones. In particular, if one TSO is Benevolent and the other is not, the latter TSO may take advantage of the situation and transfer local or TSO costs to the Benevolent TSO. If this happened in real conditions, the Benevolent TSO would probably adapt and turn to another strategy. It is therefore unlikely that all TSOs would adopt a Benevolent strategy without an adaptation of regulation.

These results lead to two policy recommendations. First, coordination should be fostered by appropriate regulation. In the design of the common European balancing markets, it has been decided that filtering would be applied by each TSO independently, without knowledge of bids emitted in other zones. No regulation currently encourages inter-TSO coordination : for example, there is no obligation to publish filtering methods used. It has been underlined previously that inter-TSO coordination is essential for efficient congestion management and the findings of this study further demonstrate its importance in the context of filtering. Several measures could help enhance coordination: sharing filtering results or algorithms among zones, or training filtering algorithms on the same dataset.

Second, regulation should be adapted to ensure Social Welfare-maximizing filtering. Results showed that the current market design does not send the correct signals. Unsuitable signals can be reinforced by national regulation. In France for example, the TSO is incentivized to reduce congestion management costs (Commission de regulation de l'Energie, 2021) - it is heavily penalized if it exceeds the reference trajectory. This may encourage the TSO to put congestion management cost reduction ahead of market welfare, whereas one of the keys to efficient filtering is finding the balance between both. If this TSO wished to implement Social Welfare-maximizing filtering, it would first need to request a change in tariff - which occurs every 4/5 years - before being able to implement it. Furthermore, the French TSO is state-owned; the incentive for a private TSO to reduce congestion management costs may be even stronger. As a solution, the national regulator can monitor the TSO's filtering decisions to check that it does not adopt a strategy adverse to market players and/or other TSOs. An alternative would be the development of an incentive regulation on congestion, taking into account a criterion of overall Social Welfare rather than congestion costs.

Funding sources — The research presented in this paper was supported by the French Electricity Transmission System Operator, RTE, and the French National Association for Technological Research under ANRT contract n2020/0857.

References

Barrows, C., Preston, E., Staid, A., Stephen, G., Watson, J.-P., Bloom, A., Ehlen, A., Ikaheimo, J., Jorgenson, J., Krishnamurthy, D., Lau, J., McBennett, B., and O'Connell, M. (2020). The IEEE reliability test system: A proposed 2019 update. *IEEE Transactions on Power Systems*, 35(1):119–127.

Bertsch, J., Hagspiel, S., and Just, L. (2016). Congestion management in power systems: Long-term modeling framework and large-scale application. *Journal of Regulatory Economics*, 50(3):290–327.

Cogen, F., Little, E., Dussartre, V., and Bustarret, Q. (2024). ATLAS: A model of short-term european electricity market processes under uncertainty – balancing modules.

Commission de regulation de l'Energie (2021). Deliberation of the french energy regulatory commission of 21 january 2021 deciding on the tariffs for the use of public transmission electricity grids (turpe 6 htb).

Commission de regulation de l'Energie (2022). Deliberation n2022-152.

Creos, Tennet, Amprion, RTE, Transnet BW, Elia, 50Hertz, APG (2020). CWE TSOs methodology for capacity calculation for the intraday timeframe.

Daneshfar, F. and Bevrani, H. (2010). Load–frequency control: a ga-based multi-agent reinforcement learning. *IET generation, transmission & distribution*, 4(1):13–26.

Doorman, G., Haberg, M., Overjordet, A. S., Warland, L., Maeland, H., Tveita, ., and Gronaas, H. (2022). Handling intra-zonal constraints in the upcoming european balancing markets. *CIGRE*.

Du, Y., Li, F., Zandi, H., and Xue, Y. (2021). Approximating nash equilibrium in day-ahead electricity market bidding with multi-agent deep reinforcement learning. *Journal of modern power systems and clean energy*, 9(3):534–544.

ENTSO-E (2020). MARI accession roadmap.

ENTSO-E (2023). Terre.

Girod, M., Donnot, B., Dussartre, V., Terrier, V., Bourmaud, J.-Y., and Perez, Y. (2024). Bid filtering for congestion management in european balancing markets–a reinforcement learning approach. *Applied Energy*, 361:122892.

Girod, M., Karangelos, E., Little, E., Terrier, V., Bourmaud, J.-Y., Dussartre, V., Jouini, O., and Perez, Y. (2022). Improving cross-border capacity for near real-time balancing. In *2022 18th International Conference on the European Energy Market (EEM)*, pages 1–6.

Glachant, J.-M. and Pignon, V. (2005). Nordic congestion's arrangement as a model for Europe? Physical constraints vs. economic incentives. *Utilities policy*, 13(2):153–162.

Glachant, J.-M., Rossetto, N., and Vasconcelos, J. (2017). Moving the electricity transmission system towards a decarbonised and integrated Europe: Missing pillars and roadblocks. *Policy Briefs, Florence School of Regulation, Energy, Climate.*

Guntermann, C., Wahl Gunderson, N., Lindeberg, E., and Hberg, M. (2018). Detecting unavailable balancing energy bids due to risk of internal congestions. In *2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I CPS Europe)*, pages 1–4.

Haarnoja, T., Zhou, A., Hartikainen, K., Tucker, G., Ha, S., Tan, J., Kumar, V., Zhu, H., Gupta, A., Abbeel, P., and Levine, S. (2019). Soft actor-critic algorithms and applications.

Harder, N., Qussous, R., and Weidlich, A. (2023). Fit for purpose: Modeling wholesale electricity markets realistically with multi-agent deep reinforcement learning. *Energy and AI*, 14:100295.

International Energy Agency (2024). World energy outlook. Technical report, IEA.

Kunz, F. and Zerrahn, A. (2015). Benefits of coordinating congestion management in electricity transmission networks: Theory and application to germany. *Utilities Policy*, 37:34–45.

Little, E., Cogen, F., Bustarret, Q., Dussartre, V., Lasri, M., Kasmi, G., Girod, M., Bienvenu, F., Fortin, M., and Bourmaud, J.-Y. (2024). ATLAS: A model of short-term european electricity market processes under uncertainty.

Marot, A., Donnot, B., Dulac-Arnold, G., Kelly, A., O'Sullivan, A., Viebahn, J., Awad, M., Guyon, I., Panciatici, P., and Romero, C. (2021). Learning to run a power network challenge: a retrospective analysis. In *Proceedings of the NeurIPS 2020 Competition and Demonstration Track*, volume 133 of *Proceedings of Machine Learning Research*, pages 112–132. PMLR.

May, R. and Huang, P. (2023). A multi-agent reinforcement learning approach for investigating and optimising peer-to-peer prosumer energy markets. *Applied Energy*, 334:120705.

Papavasiliou, A., Bjorndal, M., Doorman, G., and Stevens, N. (2020). Hierarchical balancing in zonal markets. In *2020 17th International Conference on the European Energy Market (EEM)*, pages 1–6. IEEE.

Appendices

A Acronyms and nomenclature

Table A.1: Acronyms and nomenclature

Acronyms	
ATC	Available Transfer Capacity
mFRR	manual Frequency Restoration Reserve
RL	Reinforcement learning
RR	Replacement Reserve
TSO	Transmission System Operators

Nomenclature	
s_i	Network state at timestep i
nSW_i	Net Social Welfare at timestep i
$a_i = (a_0, \dots, a_b)$	RL Filtering action at timestep i
$Q_i = (q_0, \dots, q_b)$	Bid volumes offered at timestep i
$P_i = (p_0, \dots, p_b)$	Bid prices offered at timestep i
$L_i = (l_0, \dots, l_n)$	Load at n nodes of the network at timestep i
ATC_i	Cross-border capacities at timestep i
$MC = (mc_0, \dots, mc_g)$	Marginal costs for all generators
$M(t)$	Market outcome
$SA(t)$	Security analysis outcome
B	Set of bids emitted in all zones
B_z	Set of bids emitted in zone z
q_b^*	Accepted volume of bid b
σ_b	Direction of bid b
C_z^*	Market clearing price in zone z
q_{TSO}^*	Accepted volume of TSO demand
σ_{TSO}	Direction of TSO demand
p_{TSO}	Price of TSO demand
$r_i = (r_{1,t}, \dots, r_{n,t})$	RL reward at timestep i
$o_i = (o_{1,t}, \dots, o_{n,t})$	RL observation at timestep i

B Analysis of filtering actions on a time step

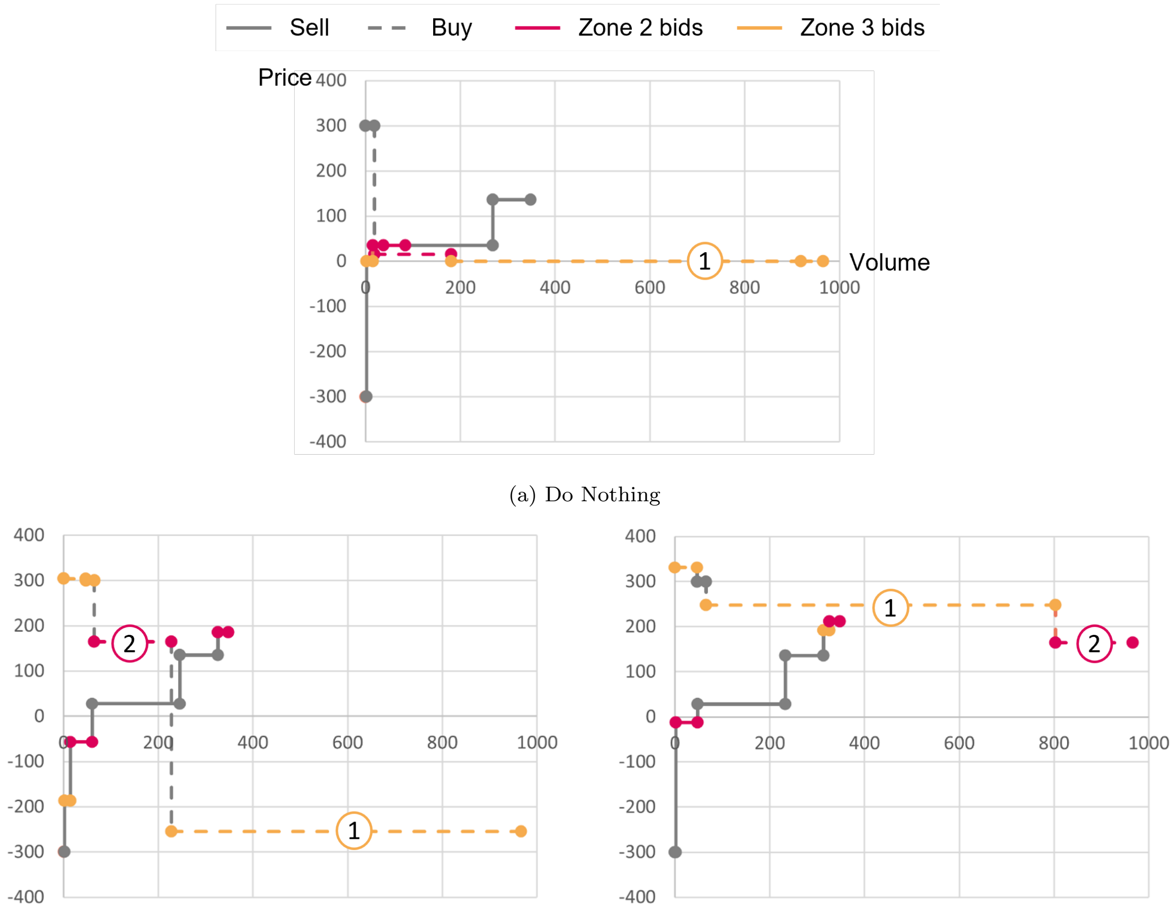
Figure 4 showed that Two Benevolent and Local + Benevolent variants perform best in terms of net Social Welfare, but lead to different situations in terms of local Welfare and TSO costs. This section focuses on the action of the corresponding RL agents on one hour. The chosen time step is January 13$^{\text{th}}$ at 7PM, marked on Figure 5 by an arrow, as Do Nothing, Two Benevolent TSOs and Local + Benevolent TSOs lead to contrasted results. Table B.1 summarizes the results for each variant.

Table B.1: Outcome of Do Nothing, Benevolent TSOs and Local + Benevolent TSOs on January 13th at 7PM

	Do Nothing	**Benevolent**	**Local + Benevolent**
Total cost (€)	28k	1.4k	12k
Local cost zone 2 (€)	11k	4.5k	-16k
Local cost zone 3 (€)	8k	0.3k	78k
Market price (€)	26	28	248

In that hour, Two Benevolent TSOs lead to the lowest total costs and zone 2 bears most of the costs. Local + Benevolent TSOs lead to higher total costs, but still lower than Do Nothing. Zone 2, where the TSO is Local, has negative local costs in this case and costs in zone 3 blow up. In the Local+ Benevolent variant, the market price skyrockets (the market price is equal across all zones in all variants).

Figure B.1 presents the effects of each variant on the balancing market at that hour. The dashed line is the 'Buy' curve - or downward bids -, the other is the 'Sell' curve - or upward bids. The bids from all zones are represented; the cross-border capacity is not limiting at that hour so all exchanges were possible. Bids that can be filtered by the zone 2 TSO (all bids emitted in the zone except TSO demand) are colored in magenta and those that can be filtered by the zone 3 TSO are colored in yellow.



(a) Do Nothing

(b) Benevolent TSOs

(c) Local+Benevolent TSOs

Figure B.1: Merit order curves for Do Nothing, Benevolent and Local + Benevolent strategies on January 13th at 7PM

In the Do Nothing case, the downward bid marked as 1 on Figure B.1, which is emitted by a wind power plant, is not accepted by the market. In the security analysis, this power plant is downward redispatched and expensive combustion turbine thermal generators are activated in all three zones to compensate.

In the Two Benevolent TSOs case, the downward bid marked as 1 is disadvantaged and the downward bid marked as 2 is advantaged and now accepted by the market. The security analysis leads to the same volume of downward redispatch on bid 1 power plant. This time however, the corresponding upward redispatch is performed by the bid 2 power plant, which is a gas combined cycle plant and is much cheaper than the combustion turbine plants. Advantaging bid 2 in the balancing market made it possible for the bid to be upward redispatched in the security analysis.

In the Local + Benevolent case, the zone 3 TSO advantages bid 1 more than zone 2 advantages bid 2. Bid 1 is accepted by the market at a volume higher than it was previously redispatched. The network is secure and no redispatch occurs in the security analysis.

Although the Local + Benevolent action may seem like the most straightforward solution, it actually leads to higher total costs than the Two Benevolent TSOs solution. Because the volume offered in bid 1 is so large, and in order to advantage it, the TSO has set its price higher than all upward bids, the market leads to a high volume being accepted. Advantaged bids create negative market welfare, as they disrupt the merit order. It is more beneficial to accept the smaller volume of bid 2 and redispatch than accept a large volume of bid 1 outside of the merit order.

Furthermore, in the Local + Benevolent variant, the zone 2 TSO transfers all the local costs to zone 3. Not only does the zone 2 TSO incur zero congestion management costs, it increases its local market welfare. It sets the price of an upward bid just below the price of bid 1: the upward bid sets the market price at 248€. Only upward bids are accepted for zone 2 and the high market price is favourable for them. TSO demand in zone 2 aims to decrease production i.e. it is a "Buy" bid, whereas in the other two zones they are "Sell" bids. TSO demand welfare is very high in zone 2 with the high market price, and low in the other zones.

Because the RL agents in zone 2 and zone 3 are trained together, there is a high level of coordination between TSOs: they learn the other's strategy and adapt accordingly. In the Benevolent case, this coordination gives an extra lever to meet the objective: the zone 2 and 3 TSO strategies are complementary and help reduce total costs. In the Local + Benevolent case, this coordination helps the Local zone 2 TSO profit from the Benevolent zone 3 TSO. If the Benevolent TSO does not advantage bid 1, the Local TSO will not either advantage bid 2, which would increase zone 2 local costs (Benevolent variant). The resulting situation would lead to higher total costs (Do Nothing variant). Not only does zone 2 TSO not push for bid 2 to be accepted, but it sets the price of its upward bids in a way that leads to very low costs in zone 2 and very high costs in zone 3.

C Total costs breakdown

Total costs are defined as the sum of balancing generation costs (the opposite of market welfare) and congestion management costs. Figure C.1 presents total costs for each variant, broken down by balancing generation costs and congestion management costs. Congestion management costs systematically decrease with filtering. Balancing generation costs increase with filtering, as Proposed filtering leads to deviating from the merit order. Compensation is counted in Balancing generation costs.

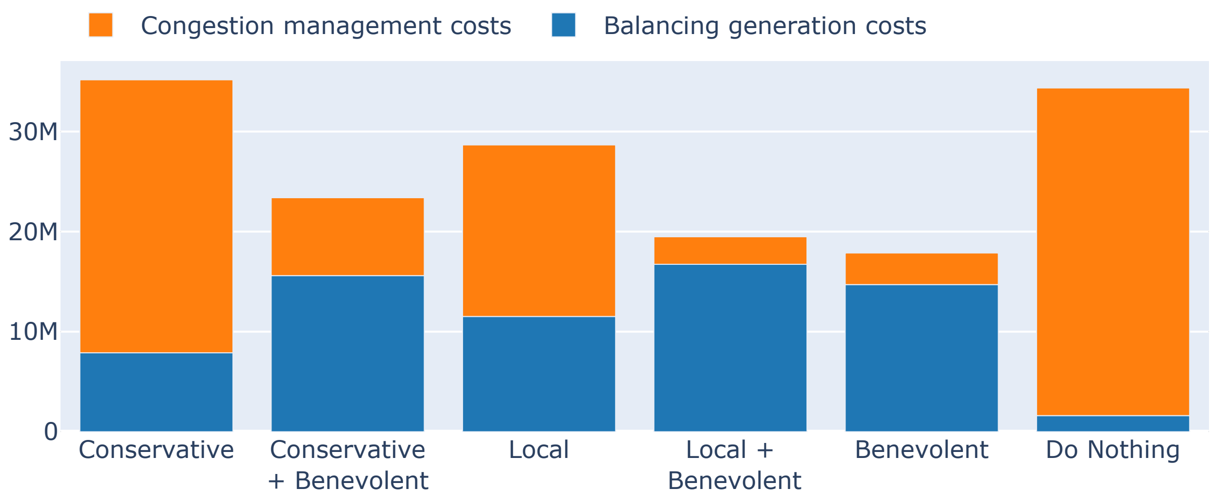


Figure C.1: Total costs broken down for each variant