

Affine Decision Rule Approximation to Address Demand Response Uncertainty in Smart Grids' Capacity Planning

Sajad Aliakbari Sani*, Olivier Bahn, and Erick Delage

GERAD and Department of Decision Sciences, HEC Montréal, Montréal, Quebec, H3T 2A7, Canada

February 23, 2022

Abstract

Generation expansion planning is a classical problem that determines an optimal investment plan for the expansion of electricity network. With the advent of demand response as a reserved capacity in smart power systems, recent versions of this class of problems model demand response as an alternative for the expansion of the network. This adds uncertainties, since the availability of this resource is not known at the planning phase. In this paper, we model demand response uncertainty in a multi-commodity energy model, called ETEM, to address the generation expansion planning problem. The resulting model takes the form of an intractable multi-period adjustable robust problem which can be conservatively approximated using affine decision rules. To tackle instances of realistic size, we propose a Benders decomposition that exploits valid inequalities and favors Pareto robustly optimal solutions at each iteration. The performance of our new robust ETEM is evaluated in a realistic case study that surveys the energy system of the Swiss “Arc Lémanique” region. Results show that an adjustable robust strategy can potentially reduce the expected cost of the system by as much as 33% compared to a deterministic approach when accounting for electricity shortage penalties. Moreover, an adjustable procurement strategy can be responsible for a 9 billion Swiss francs cost reduction compared to a naive static robust strategy. The proposed decomposition scheme improves the run time of the solution algorithm by 40% compared to the traditional Benders decomposition. To conclude, we provide a discussion on other possible problem formulations and implementations.

Keywords : OR in energy, Multi-period adaptive robust optimization, Affine decision rule, Bottom-up energy model, Demand response

1 Introduction

Generation expansion planning (GEP) is a classical energy problem that aims at determining the required power capacity to satisfy demand over a long-term horizon at minimum cost, while satisfying economic, environmental and technical constraints (see Koltsaklis and Dagoumas 2018, for a recent review). For this, one must consider existing and future electricity generation technologies to determine an optimal investment and retirement plan

*Corresponding author

Email addresses: sajad.aliakbarisani@hec.ca, olivier.bahn@hec.ca, erick.delage@hec.ca

for the power sector. Computationally speaking, in a deterministic setting, the GEP problem is a large-scale optimization problem that typically accounts for more than hundreds of thousands of decision variables and constraints in order to identify realistic solutions. Motivated by advances in the structure of smart grids, recent versions of GEP problems have become even more complex as they attempt to model the notion of demand response (DR) (see e.g. Babonneau et al. 2017, Lohmann and Rebennack 2017). DR can be defined as the shifting of demand by consumers, in response to supply-side incentives offered at times of high wholesale market prices or when system reliability is jeopardized (US Department of Energy 2006). Development of sophisticated data-gathering devices such as smart meters and voltage sensors, as well as increasing penetration of flexible loads such as electric vehicles (EV) and heating pumps, provide the technical possibility of implementing DR programs through utilities. In the wholesale capacity market, the participants bid on providing DR resources as an alternative for expensive generation and transmission expansion strategies. DR resources can drastically reduce marginal production costs, as they are substitutes for technologies that help meet peak loads. However, including DR in the GEP problem is challenging. Indeed, DR is a price-sensitive resource, whose availability is not fully known, especially in the capacity expansion planning phase. As the planning horizon in a GEP problem is typically several decades, any planning or prediction of the availability of DR resources will be affected by two sources of uncertainties: i) errors in the prediction of total demand; and ii) changes in the response behavior. Failure to properly consider DR uncertainty (DRU) is likely to lead to situations in which there are insufficient capacity resources to satisfy realized demands.

Robust Optimization (RO) and Stochastic Programming (SP) are the two main methodologies to address uncertainty. The underlying assumption in SP is that the probability distribution of an uncertain parameter is known (Birge and Louveaux 2011). Conversely, RO does not assume information about the distribution of an uncertain parameter, but instead assumes that uncertain parameters lie in a user-defined uncertainty set. RO aims at finding solutions that are immunized to all perturbations of the uncertain parameters in the uncertainty set (Bertsimas et al. 2011). In multi-stage versions of SP and RO, it is assumed that decisions are taken at different points of time. The simplest form, known respectively as two-stage SP or Adjustable RO (ARO), considers *here-and-now* decisions that need to be made before having any information about the uncertain parameters and *wait-and-see* decisions that can be adjusted according to the realized uncertain parameters at a later time. It is generally known that ARO (Ben-Tal et al. 2004) provides less conservative optimal solutions than RO, since it has more flexibility to adjust the decisions with respect to the uncertain parameters. However, the better performance comes at the price of problem intractability. To circumvent this issue, Ben-Tal et al. (2004) propose using affine decision rules for the delayed decisions, a technique also referred as the affinely adjustable robust counterpart (AARC). Following Ben-Tal et al. (2004), other types of decision rules have been proposed for ARO problems (Bertsimas and Georghiou 2018, Georghiou et al. 2020). For a review of the ARO approach, we refer the reader to Yanikoğlu et al. (2019).

A number of different multi-stage robust or stochastic approaches have been previously used in the literature to address uncertainties in GEP problems (Dehghan et al. 2014, Mejía-Giraldo and McCalley 2014, Domínguez et al. 2016, Amjady et al. 2018, Baringo and Baringo 2018, Han et al. 2018, Zou et al. 2018). Because of the inherent computational challenges that emerge with these models, there should be a compromise between the level of details and the size of the problem. For example, Han et al. (2018) propose a two-stage stochastic program that only considers a single-period planning horizon, while the multi-stage model proposed in Domínguez et al. (2016) is only implemented in a case study involving four periods. It therefore appears that tackling uncertainty in long-term industrial-size GEP problems is out of reach for currently available methods. Improving the computational efficiency of algorithms for solving such problems is also crucial in helping to understand better the interactions between demand and supply side, as such studies can require iterating through the resolution of a number of

GEP problems to identify market equilibria (see Babonneau et al. 2016, 2020, for examples of this approach).

In this paper, we propose a numerical method for addressing DRU on a realistic large-scale GEP problem. To do so, we use the Energy-Technology-Environment Model (E_{TEM}) proposed in Babonneau et al. (2017) to model the market mechanism that matches flexible load and supply, where we introduce for the first time DRU as an implementation error of the DR decision variable. We formulate the problem as a robust multi-period linear program. In a first stage, the policy-maker decides how to invest in the capacity of each generation technology and plans for a target DR for each future time period. Then, periodically, depending on the actual level of contribution in DR programs for this period, the planner decides on optimal energy procurement based on available resources. We evaluate the performance of the proposed *robust E_{TEM}* model on a real-world case study based on the energy system of the “Arc Lémanique” region in Switzerland. To solve the problem, we derive a Robust Multi-Period Conservative Approximation (RMPCA) of the problem, and develop a Benders decomposition algorithm (inspired from Ardestani-Jaafari and Delage 2018) to solve it. In order to focus on the effects of DRU on the design of smart electricity networks, we have limited the model in this paper to only consider the uncertainty of DR. However, the proposed methodology could be extended to accommodate supply-side uncertainties, such as availability of intermittent resources, or investment costs. We however leave this as the subject of future work as discussed in Section 7.

Overall, the contributions of this paper can be summarized as follows:

1. We introduce DRU in a GEP problem, addressed using the E_{TEM} energy model, and propose a robust multi-period linear program in which the procurement decisions are adjustable to the actual DR. The approach that we propose is flexible in the sense that, with minor modifications, it could accommodate other sources of uncertainties in GEP problems such as uncertainty of intermittent energy resources and investment costs.
2. We propose, for the first time, to seek Pareto Robustly Optimal (PRO) solutions (see Iancu and Trichakis 2014) of the master problem in a Benders decomposition algorithm in order to accelerate convergence. When combined with the valid inequalities proposed in Ardestani-Jaafari and Delage (2020), solution time is improved by 40% on average in our randomly generated instances. Note that this idea differs from the idea of Pareto optimal cut, introduced by Magnanti and Wong (1981), in the way that the latter seeks a Pareto optimal solution among multiple optimal solutions of the sub-problems. In contrast, we seek a PRO solution in the master problem, i.e., a solution that is guaranteed not to be dominated by any other optimal solutions in terms of its constraint margin profile.
3. Compared to previous work in the literature, we solve an adjustable robust multi-commodity GEP problem whose size, in deterministic form, is two orders of magnitudes larger than the largest instance addressed in previous case studies. Moreover, our case study confirms in a simulation that adjustable robust policies can significantly improve (i.e., around 33%) the expected total cost of the energy system compared to the solution of the deterministic version when accounting for electricity shortage penalties. Furthermore, it shows that the flexibility of energy procurement decisions can be responsible for a reduction of 9 billion Swiss francs (CHF) in expected total expansion planning costs.

The remainder of the paper is organized as follows. In the next section, we present an overview of different approaches to model DRU, and summarize studies that use a multi-stage approach to model uncertainties in GEP problems. In Section 3, we introduce briefly the E_{TEM} model. In Section 4, we present how we have modeled DRU in E_{TEM} and derive the multi-period robust optimization problem as well as its conservative approximation. In Section 5, the Benders decomposition is detailed. Section 6 presents the case study and provides numerical results. Finally, Section 7 provides concluding remarks.

2 Literature Review

In this section, we first review papers that use a multi-period approach to model a robust or stochastic GEP problem. We specifically summarize their solution method and the size of the instances they solve. Usually, the size of a GEP problem is dependent on the number of time periods, load duration curve (LDC) steps, technologies, and commodities. Hence, we will summarize the size of studied instances with the product of these quantities, referred as the Deterministic Size Indicator (DSI) index. In a second part, we review recently proposed strategies to model DRU in different energy-related problems. Finally, we review the literature on the ETEM model.

Bloom (1983) and Bloom et al. (1984) are among the first papers that propose a two-stage stochastic programming formulation for GEP problems with uncertain demand and supply. They propose a general Benders decomposition, where a master problem accounts for investment decisions, and a set of subproblems represent the annual operation cost and reliability level of the installed capacities while enforcing probabilistic reliability constraints. Han et al. (2018) develop a two-stage stochastic program to model the uncertainty of load demand and wind output in a GEP problem. Using affine decision rules, they reformulate the problem as a deterministic second-order cone optimization problem. They solve an instance with 1 period, 5 LDC steps, and 32 technologies to test their algorithm (DSI=160). Mejía-Giraldo and McCalley (2014) propose an adjustable robust optimization framework to address the uncertainty of fuel price, demand, and transmission capacities in a GEP problem. In this setting, investment decisions are set as an affine function of fuel price, whereas voltage angle (decision) is parameterized as an affine function of demand and transmission capacities. They reformulate the problem as an LP and test it on a simplified version of the US power system with 20 periods, 3 LDC steps, and 13 technologies (DSI=780). Domínguez et al. (2016) use linear decision rules to reformulate a multi-stage GEP problem and cast it as a tractable LP. They solve a GEP problem with 4 decision periods, 18 technologies and 100 LDC steps (DSI=7200). Zou et al. (2018) propose a partially adaptive multi-stage stochastic GEP problem. In this setting, the capacity expansion plan is adaptive to the uncertain parameter up to a certain point. Fuel price and demand are two sources of uncertainties in this study. They test their algorithm on a case study with 10 periods, 3 LDC steps, and 6 technologies (DSI=180). Dehghan et al. (2014) develop a two-stage robust optimization problem in which the load demand and investment cost are uncertain. They solve their problem, using a cutting plane method. The case study is a 10-period planning of a network with 120 technologies and 3 LDC steps (DSI=3600). None of the above studies model neither DR nor DRU. In addition, all of them are a single-commodity GEP problem, meaning that they only model the electricity network and account for an aggregated electricity demand over the planning horizon. By contrast, in a multi-commodity energy model such as ETEM, one also models the interactions of the electricity sector with other sectors of the energy system (e.g., heat, transportation sectors), as well as the demands for final energy services (such as heat and lighting). This increases the size of the deterministic problem. In this paper, we solve a case study with 10 decision-making periods, 12 LDC steps, 142 technologies, and 57 energy commodities (DSI=971,280).

DR is also modeled in other power system planning problems such as unit commitment, transmission or distribution expansion planning, and wholesale electricity market problems. In general, there are two strategies one can follow to implement DR programs: dispatchable and non-dispatchable. In the former, the utility directly controls the timing of some loads of the customers who voluntarily participate in the program. More precisely, the utility cuts down the consumption during peak periods to ensure system reliability, and in return, remunerates participants with annual payments. In non-dispatchable strategies, the utility sends to customers a price signal with the purpose of flattening the demand curve. Automatic price-responsive devices adjust the timing of the consumption. While developing an optimal consumption schedule is the focus of some research (Liu et al. 2019), another stream studies the conditions under which the dynamic price signals improve social welfare (see, e.g.,

Adelman and Uçkun 2019, Venizelou et al. 2018). Elaboration on the different strategies to implement DR and an overview on the success of these programs in the US electricity market is presented in Shariatzadeh et al. (2015).

The effect of DR on the wholesale and capacity market is another stream of research. Vatani et al. (2017) introduce DR to a capacity market problem in which the cost-minimizer capacity planner trades off between new generation and transmission capacity expansion and DR expenses. Arasteh et al. (2015) model a trade-off between DR expenses and distribution expansion costs. Zhang and Zhang (2019) model DR as an alternative procurement strategy for an electricity retailer. Namely, the electricity retailer must procure energy in a day-ahead wholesale market. Given that the prices in this market are highly uncertain, DR is introduced as a leverage to manage the retailer's risk. In this approach, the retailer solves its problem to determine a procurement strategy as well as DR incentive prices. Then, consumers respond to the signaled price in order to minimize their consumption cost. Babonneau et al. (2016) model DR in a capacity expansion planning problem considering distribution constraints. They develop a game framework between the retailer and many small consumers. The retailer solves a minimization model to determine his desired DR and capacity expansion plan. Then, using marginal cost of production, which corresponds to dual variables, he signals consumers to adjust their consumption. They argue that, if the price function, with which the retailer signals consumers, is convex with respect to both demand and required level of reserved capacity, their approach can be cast as a linear programming problem. Finally, Lohmann and Rebennack (2017) model DR in a long-term capacity expansion planning problem.

In all the above research, DR was mostly modeled in a deterministic way. In other words, when DR was modeled using a price elastic demand curve, no uncertainty affected this curve. Alternatively, when it was modeled as a decision variable, no implementation error was affecting optimal DR decisions. However, due to a variety of reasons, including demand prediction errors and changes in response behavior, in practice DR necessarily ends up different from what the supply side expected. According to Chatterjee et al. (2018), a lack of DR forecasting and estimation tools is one of the most important barriers to wider adoption of DR. On this regard, an active stream of literature has proposed a number of statistical learning methods for predicting demand response associated to storage-like flexible loads such as residential and industrial air conditioning (see, e.g., Dyson et al. 2014, Qi et al. 2020), and a stand-alone micro-grid with the presence of photovoltaic and wind generation units (see Amrollahi and Bathaee 2017). Another stream of research investigates the influence of DRU on its integration to the network. Li et al. (2015) model incentive-based DR as an alternative for transmission upgrade investments and consider consumer's bid for load reduction uncertainty. They develop a stochastic programming model and use Monte Carlo simulation and Benders decomposition for resolution. Ströhle and Flath (2016) model DRU in a local online market to match flexible load and uncertain electricity supply. Gärttner et al. (2018) formulate DRU in a demand aggregator's problem where the scheduling of demand and the optimal dispatch are optimized at the same time. He et al. (2019) develop a two-stage distributionally robust problem to model DRU in a distribution network expansion problem. Prices are modeled as a decision variable to optimally shape the desired demand response. However, the price elasticity is considered to be a random parameter. Moreover, distributed generation resources are also affected by uncertainty. They propose a column and constraint generation approach to solve the resulting problem. Zhao et al. (2013) model the uncertainty of DR and wind generation in a unit commitment problem. DR is modeled through a price-elastic demand curve with price uncertainty. They propose a three-stage robust optimization formulation. First stage decisions consider unit commitment planning, i.e., turning on and off the generators. Second stage decisions include the dispatch of electricity and are implemented after the actual wind power output is realized. Finally, in a third stage, the demand response is observed. Moreover, they propose a Benders decomposition to solve the three-stage problem. An overview of the different techniques to model ancillary services in unit commitment problems is

presented in Knueven et al. (2020). Asensio et al. (2018) model the DRU through a price-sensitive demand curve in a distribution network expansion problem and use a scenario-tree approach to solve the resulting stochastic programming model. Roveto et al. (2020) develop a data-driven distributionally robust optimization model of the demand response auction, where the risk of the retailer is controlled using value-at-risk and conditional value-at-risk measures. They model DRU by considering DR bid uncertainty. In other words, the aggregator receives bids from consumers to reduce their consumption at a specific time. However, consumers might not actually reduce their consumption to this full amount. The aggregator aims to ensure a certain level of demand response procurement at minimum cost. Huang et al. (2020) develop a multi-stage stochastic programming model to optimally schedule power generation assets when the system operator has access to the ancillary service market. In this paper, we model DRU in a multi-commodity energy planning model, called ETEM (Babonneau et al. 2017), to address a GEP problem. In ETEM, DR is modeled through a decision variable which is optimized by the supply-side. It is assumed that the supply-side can shape consumer’s response behavior through different long-term DR programs. For the first time, we model DRU as an implementation error of the DR decision in ETEM, and develop an adaptive robust multi-period conservative approximation formulation of the problem.

ETEM first was formulated by Babonneau et al. (2012). Although the developed formulation is deterministic, they also provide a stochastic programming model with different scenarios for the electricity import cost. As an extension to the primary ETEM model, Babonneau et al. (2016) propose a linear approximation of distribution system and power flow constraints in ETEM. The new formulation, called ETEM-SG, accounts for provision of reserved capacities provided by demand response in smart grids. Babonneau et al. (2017) and Babonneau and Haurie (2019) propose robust versions of ETEM to address the uncertainty of investment costs and of availability of technologies and transmission capacity of electricity transmission lines. In both papers, a static robust version of ETEM is developed and a relaxed version of the model is solved to reduce conservatism of the solution. Babonneau et al. (2020) is the first version of ETEM that employs robust optimization to handle uncertain demand response. Specifically, it proposes a robust ETEM model coupled with a mean-field game equilibrium model that represents the charging behavior of electric vehicle owners. The paper models the uncertainty of DR by introducing in ETEM two upper and lower bounding constraints on the sum of DR over all time periods. The bounding constraints are calibrated based on the confidence intervals of the random demand which is simulated using a mean-field game model. We note that Babonneau et al.’s way of accounting for DRU in ETEM is fundamentally different from ours. First, they again only involve static decisions and focus on a single-period investment problem. More importantly, a careful analysis of their approach reveals that it effectively offers no protection against the DR deviations in our setting. A more detailed demonstration of this weakness is presented in Appendix A. Our paper should therefore be seen as presenting a significant extension of ETEM that identifies adaptive long term strategies to immunizes for the first time the energy system against DR deviations.

3 The ETEM Energy Model

As proposed in Babonneau et al. (2012), ETEM is a long-term, bottom-up energy model cast as a linear programming problem. It is a member of the *MARKAL-TIMES* family of models, and represents the entire energy sector from primary resources to useful energy demands. ETEM is a demand-driven model in which the objective is to provide energy services at minimum cost. The planning horizon is typically more than 50 years and the model gives insights on both long-term strategic and short-term operational decisions. While the long-term decisions include capacity expansion and strategic targeting for a desired demand response, short-term decisions consist of energy procurement planning, which includes optimal generation, import, export and regional transmissions. ETEM can also account for technical, economic and market constraints (see Babonneau et al.

2017).

ETEM is based on the concept of a reference energy system (see Figure 1 for the depiction of the energy system of the Arc Lémanique region in Switzerland considered in our case study). The model considers different energy commodities (captured by vertical lines in the figure) namely “resources” (i.e., a mix of primary energy resources and some imported secondary energy forms), selected secondary forms (i.e., electricity and heat), and useful energy demands. More precisely, resources correspond to coal, oil products, gas, hydro, intermittent renewables (wind and solar), as well as other sources such as municipal solid waste and wood. Useful energy demands belong to four categories: industry, residential electricity, residential heat, and transportation. The boxes correspond to the different energy technologies, from generators to technologies providing energy services to end-users. In our implementation, we also add an energy resource (denoted “ExR”) with a high cost that represents the cost of shortage in the system. A dummy technology (denoted “ExT”) links this expensive resource to the useful demands (see Assumption 1 and following discussion for more details).

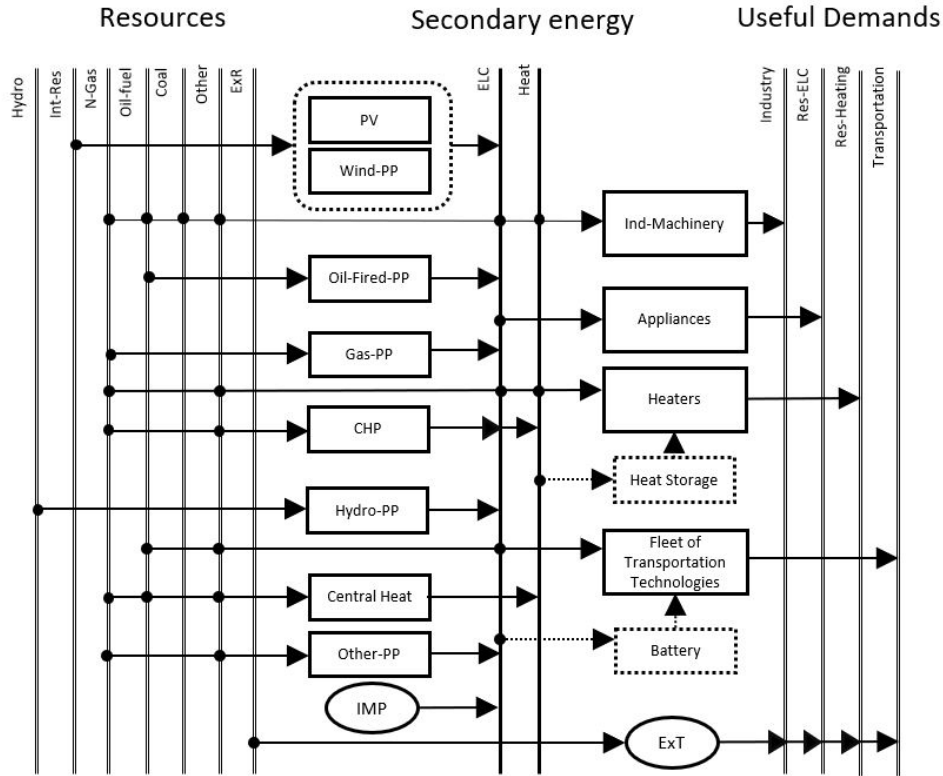


Figure 1: Arc Lémanique reference energy system, where Int-Res stands for intermittent resources (solar and wind), N-Gas for natural gas, Oil-fuel for processed oil products, Other for hydrogen and additional sources of energy (such as solid waste and wood and geothermal), ELC for electricity, RES for residential, PP for power plant, CHP for combined heat and power plant, IMP for electricity import, Ind-Machinery for industrial machinery, ExR for a dummy expensive resource added to avoid in-feasibility, ExT for a dummy expensive technology used to avoid infeasibility. Other-PP includes geothermal, fuel-cell and municipal waste power plants.

The planning horizon in ETEM spans over several decades to simulate the long investment cycles typical of the energy sector. We denote by $t \in \mathbb{T} := \{1, \dots, T\}$ the index representing the decision-making periods. In addition, to capture short-term operational patterns of the demand load, each period is divided into smaller time-slices $s \in \mathbb{S}$, representing load duration curve steps. In ETEM, 12 time-slices are assumed ($S = \{1, \dots, 12\}$),

which are partitioned among a set J of 3 typical seasons (winter, summer, and intermediate), and each day of these seasons is divided into 4 time slots (morning, midday, midnight, and night), as shown in Figure 2.

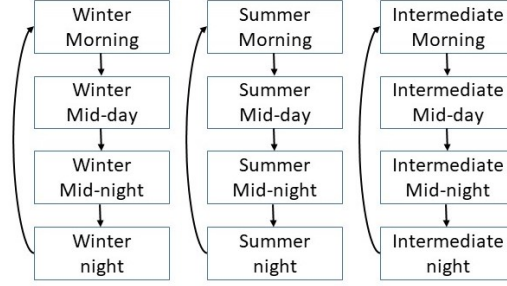


Figure 2: Sequence of time-slices

Before describing ETEM's mathematical formulation, it is worth introducing our nomenclature. Let C be the set of energy commodities and let P be the set of energy technologies. The set of indices F identifies the different input-output energy flows associated to each technology. For instance, in Figure 1, Combined Heat and Power (CHP) uses natural gas, solid waste and wood as input to generate both electricity and heat. Finally, let L be a set of buses in different geographical zones. A full nomenclature is presented below.

Nomenclature

$t \in T$ Index for time period

$s \in S$ Index for time-slices

$p \in P$ Index for technologies

$c \in C$ Index for energy commodities

$c_s \in CS$ Index for energy storage

$f \in F$ Index for energy flows

$l \in L$ Index for buses (geographical zones)

$j \in J$ Index for seasons

$i \in I$ Index for period-seasons (t, j)

P_c^C P Set of technologies consuming c

P_c^P P Set of technologies producing c

P^R P Set of intermittent technologies

C^I C Set of imported commodities

C^D C Set of useful demands

C^{EX} C Set of exported commodities

C^{TR} C Set of transmitted commodities

C_f C Set of commodities linked to flow f

C^G C Set of commodities with margin reserve

S^j S Set of time-slices s in season j

S^s S Set of successive time-slices of s

S^G S Set of time-slices in peak period

$F|_p$ F Set of inputs to technology p

FO_p F Set of outputs from technology p

$\alpha_{t,p}$ Investment cost

$\beta_{t,s,p}$ Capacity factor

η_c Network efficiency

$\eta_{f,j}^t$ Technology efficiency

$\lambda_{t,s,l,l^0,c}^\infty$ Transmission cost

$\lambda_{t,s,c}^0$ Export cost

$\lambda_{t,s,c}$ Import cost

$\nu_{t,s,c}$ Maximum deviation from nominal demand response

$\nu_{t,p}$ Variable cost

$\Omega_{t,l,p}$ Available capacity of technology p

$\pi_{t,p}$ Fixed production cost

ρ Discount factor

θ_p^c	Proportion of output c from technology p that can be used in peak period	$\mathbf{C}_{t,l,p}$	Variable for new capacity addition
$\Theta_{t,l,d}$	Annual final demand	$\mathbf{E}_{t,s,l,c}$	Variable for export
$v_{t,s,c}$	Nominal demand response	$\mathbf{I}_{t,s,l,c}$	Variable for import
$\varrho_{t,s,c}$	Required reserve for commodity $c \in \mathcal{C}^G$	$\mathbf{P}_{t,s,l,p,c}$	Variable for activity of technology p
l_p	Life duration of technology p	$\mathbf{T}_{t,s,l,l^0,c}$	Variable for regional transmission
		$\mathbf{V}_{t,s,d}$	Variable for demand response.

Next, we present the main elements of ETEM. Interested readers can refer to Babonneau et al. (2017) for more details:

$$W := \min \sum_{t,l} \rho^t \left(\sum_p \alpha_{t,p} \mathbf{C}_{t,l,p} + \pi_{t,p} \left(\sum_{k=0}^{l_p-1} \mathbf{C}_{t,k,l,p} + \Omega_{t,l,p} \right) + \sum_{s,c} \nu_{t,p} \mathbf{P}_{t,s,l,p,c} + \sum_{s,c} (\lambda_{t,s,c} \mathbf{I}_{t,s,l,c} - \lambda_{t,s,c}^0 \mathbf{E}_{t,s,l,c}) + \sum_{s,l^0,c} \lambda_{t,s,l,l^0,c}^{\infty} \mathbf{T}_{t,s,l,l^0,c} \right), \quad (1a)$$

s.t.

$$\left(\sum_{p \in \mathcal{P}^c} \mathbf{P}_{t,s,l,p,c} + \mathbf{I}_{t,s,l,c} \right) \eta_c + \sum_{l^0 \in \mathcal{L}} (\eta_c \mathbf{T}_{t,s,l^0,l,c} - \mathbf{T}_{t,s,l,l^0,c}) - \sum_{p \in \mathcal{P}^c} \mathbf{P}_{t,s,l,p,c} + \mathbf{E}_{t,s,l,c} \leq \delta_{t,s,l,c} \leq \mathcal{C}/\mathcal{C}^D \quad (1b)$$

$$\left(\sum_{l,p \in \mathcal{P}^c} \theta_p^c \beta_{t,s,p} \left(\sum_{k=0}^{l_p-1} \mathbf{C}_{t,k,l,p} + \Omega_{t,l,p} \right) + \sum_{l,p \in \mathcal{P}^c/\mathcal{P}^c} \theta_p^c \mathbf{P}_{t,s,l,p,c} + \mathbf{I}_{t,s,l,c} \right) \varrho_{t,s,c} - \sum_{l,p \in \mathcal{P}^c} \mathbf{P}_{t,s,l,p,c} + \mathbf{E}_{t,s,l,c} \leq \delta_{t,s} \leq \mathcal{S}^G, c \in \mathcal{C}^G \quad (1c)$$

$$\eta_c \sum_{p \in \mathcal{P}^c} \mathbf{P}_{t,s,l,p,c} = \sum_{p \in \mathcal{P}^c, s^0 \in \mathcal{S}^s} \mathbf{P}_{t,s^0,l,p,c} \leq \delta_{t,s,l,c} \leq \mathcal{C}\mathcal{S} \quad (1d)$$

$$\sum_{p \in \mathcal{P}^c} \mathbf{P}_{t,s,l,p,c} - \Theta_{t,l,c} \mathbf{V}_{t,s,c} \leq \delta_{t,s,l,c} \leq \mathcal{C}^D \quad (1e)$$

$$v_{t,s,c} (1 - \nu_{t,s,c}) \mathbf{V}_{t,s,c} - v_{t,s,c} (1 + \nu_{t,s,c}) \leq \delta_{t,s,c} \leq \mathcal{C}^D \quad (1f)$$

$$\sum_{s \in \mathcal{S}_j} \mathbf{V}_{t,s,c} = \sum_{s \in \mathcal{S}_j} v_{t,s,c} \leq \delta_{t,c} \leq \mathcal{C}^D, j \in \mathcal{J} \quad (1g)$$

$$\sum_{c \in \mathcal{P}^c} \mathbf{P}_{t,s,l,p,c} - \beta_{t,s,p} \left(\sum_{k=0}^{l_p-1} \mathbf{C}_{t,k,l,p} + \Omega_{t,l,p} \right) \leq \delta_{t,s,l,p} \leq \mathcal{P}^R, \quad (1h)$$

$$\sum_{c \in \mathcal{C}_m^p} \mathbf{P}_{t,s,l,p,c} = \beta_{t,s,p} \left(\sum_{k=0}^{l_p-1} \mathbf{C}_{t,k,l,p} + \Omega_{t,l,p} \right) \leq \delta_{t,s,l,p} \leq \mathcal{P}^R \quad (1i)$$

$$\sum_{c \in \mathcal{C}_f^0} \mathbf{P}_{t,s,l,p,c} = \eta_{f,f^0}^t \sum_{c \in \mathcal{C}_f} \mathbf{P}_{t,s,l,p,c} \leq \delta_{t,s,l,p} \leq \mathcal{F}\mathcal{I}_p, f^0 \in \mathcal{F}\mathcal{O}_p, t, s, l, p \quad (1j)$$

$$LB_{t,j,c} \leq \sum_{s \in \mathcal{S}^j, l, p \in \mathcal{P}^c} \mathbf{P}_{t,s,l,p,c} \leq UB_{t,j,c} \leq \delta_{t,j,c} \leq \mathcal{C} \quad (1k)$$

$$LB_{t,j,c} \sum_{l,s \in \mathcal{S}^j} \mathbf{I}_{t,s,l,c} \quad UB_{t,j,c} \quad \delta_{t,j,c} \geq C^I \quad (1l)$$

$$\sum_{s \in \mathcal{S}^j, l} \left(\sum_{p \in \mathcal{P}_c^P} P_{t,s,l,p,c} + \mathbf{I}_{t,s,l,c} \quad \sum_{p \in \mathcal{P}_c^C} P_{t,s,l,p,c} \quad \mathbf{E}_{t,s,l,c} \right) \quad UB_{t,j,c} \quad \delta_{t,j,c} \geq C \quad (1m)$$

$$(\mathbf{P}, \mathbf{I}, \mathbf{E}, \mathbf{T}) \geq \mathbf{Y}, \quad \mathbf{C} \geq \mathbf{X}, \quad \mathbf{P} \geq 0, \mathbf{I} \geq 0, \mathbf{E} \geq 0, \mathbf{T} \geq 0, \mathbf{C} \geq 0, \mathbf{V} \geq 0. \quad (1n)$$

The objective function (1a) minimizes a discounted sum of all costs of the system over all regions ($l \in \mathcal{L}$) and time periods ($t \in \mathcal{T}$), where parameters $\alpha_{t,p}$, $\pi_{t,p}$, $\nu_{t,p}$, $\lambda_{t,s,c}$, $\lambda_{t,s,c}^\theta$, and $\lambda_{t,s,l,l^o,c}^{\theta\theta}$ are unit costs of investment, fixed, variable, import, export, and regional transmission respectively.

Constraint (1b) is a commodity balance constraint. It ensures that the regional procurement of each energy commodity c is greater, or equal than its total consumption at each period t and time-slice s . Specifically, the left-hand side of this constraint consists of i) total production of commodity c in region l by all technologies producing it (\mathcal{P}_c^P) ii) import of commodity c , and, iii) net transmission of commodity c into region l . It is worth mentioning that in ETEM, import (and similarly export) of energy refers to amount of input (output) energy to (from) the energy system from outside of it, but energy transmission refers to amount of energy which is produced in one region and transmitted to another region inside energy system. Parameter η_c is the network efficiency with respect to commodity c , e.g., the efficiency of electricity transmission lines. On the right-hand side, the consumption is equal to the total consumption by technologies consuming the commodity, i.e., \mathcal{P}_c^C , added to the amount of commodity that is exported.

Constraint (1c) introduces a safety margin in procurement of commodity $c \in \mathcal{C}^G$, mostly electricity, during peak time-slices $s \in \mathcal{S}^G$, to protect against random events not explicitly represented in the model. Parameter $\varrho_{t,s,c} \in [0, 1]$ represents the fraction of reserved capacity needed to ensure covering the peak load. The left-hand side of this constraint models the maximum amount of commodity $c \in \mathcal{C}^G$ that can be procured in period t and time-slice s . This amount is equal to the sum of i) the maximum production capacity of commodity c by technologies that produce it as their main output (\mathcal{P}_c), ii) the production of commodity c by technologies that produce c as their by-product of their main activity, and iii) the import of commodity c . The left-hand side is the total consumption similar to constraint (1b). Parameter θ_p^c is the proportion of technology production that can be used during the peak period. Constraint (1d) models the balance of energy storage between consecutive time-slices. Amount of storage at time-slice c can be consumed at subsequent time-slice s^θ . The notion of subsequency of time-slices is presented in Figure 2. Constraints (1e) - (1g) are the key constraints that model the use of demand response. Parameter $\Theta_{t,l,c}$ is the total demand of service $c \in \mathcal{C}^D$ in the period t and zone l . Variable $\mathbf{V}_{t,s,c}$ is the demand response which optimally distributes the total demand of period t into all time-slices s inside period t . Constraint (1f) limits the demand response to vary between an interval around the nominal value, i.e., $v_{t,s,c}$. In addition, the sum of the demand response must be equal to the sum of the nominal values in each season, according to constraint (1g).

Constraint (1h) limits the maximum production or consumption of each technology to the available capacity of that technology. The parameter $\beta_{t,s,p}$ is the capacity factor. The capacity factor of technology p is defined as the average power generated divided by the rated peak power. This simply represents the fraction of the total capacity which is available at each time-slice. In addition, since renewable generation is usually given priority in dispatch over conventional forms of generation (see for example Knueven et al. 2020, Lohmann and Rebennack 2017), their production is imposed to take its maximum production capacity in constraint (1i). Constraint (1j) models the efficiency of technology p . In ETEM, the efficiency of a technology is modeled through the annual parameter η_{f,f^o}^t that links each output flow $f^o \in \text{FO}_p$ to an input flow $f \in \text{FI}_p$. The constraint ensures that the sum of all output energy commodity c which is linked to output flow f^o , i.e., C_{f^o} , is equal to a fraction of the sum of all input energy commodities c which is linked to input flow f , i.e., C_f .

We note that constraints (1k) and (1l) impose upper and lower bounds for seasonal production of technology p in region l , and seasonal import of energy commodity c in region l . Moreover, in constraint (1m) the seasonal procurement of commodity c , over all regions in \mathbb{L} , is imposed an upper bound. This departs slightly from the original ETEM formulation, which only imposed annual bounds but is reasonable and enables the decomposition scheme presented in Section 5. We will discuss later that seasonally independent procurement decisions enable us to decompose the problem into several independent sub-problems where one decides on the optimal procurement given a capacity plan and demand response. This structure benefits the efficiency of the solution method. Finally, space X and Y represent other operational, technical, and economical constraints that define a desirable space for capacity and procurement decisions. Moreover, X and Y contain constraints that define the structure of the energy network. In particular, these constraints enforce energy productions, $\mathbf{P}_{t,s,l,p,c}$, that do not exist in the energy network to be zero. Since these constraints do not affect our analysis, we omit to report them and refer the reader to Babonneau et al. (2017) for a complete list of constraints.

4 A Demand Response Robust Version of ETEM

The ETEM model presented in Section 3 optimizes the future evolution of the energy system by assuming full information. However, in reality, there is always uncertainty. In particular, demand response is an important source of uncertainty. ETEM assumes that demand response is completely controlled by the supply side. From a decentralized electricity market point of view, this structure can be interpreted as a situation where the supply side targets an optimal level of DR, and encourages the actual demand load to be close to target DR using DR programs such as incentive real-time pricing. But, one must expect some realized deviations from the plans. Such deviations can either be caused by differences in the actual overall daily demand loads, as time evolves over a 50-year horizon, or by the level of participation and compliance of consumers in DR programs. In this paper, we model both types of uncertainties as implementation errors of the planned demand response \mathbf{V} and we immunize ETEM against such error (see Ben-Tal et al. 2015, for a seminal work on robust optimization with implementation error). Specifically, given a set $C^U \cup C^D$ of useful demand, such as residential electricity, with error-prone demand response, the demand $\Theta_{t,l,c} \mathbf{V}_{t,s,c}$ for time period $t \in \mathbb{T}$, time-slice $s \in \mathbb{S}$, region $l \in \mathbb{L}$, and commodity $c \in C^U$, is replaced with $\Theta_{t,l,c} (\mathbf{V}_{t,s,c} + \boldsymbol{\delta}_{t,s,c})$, where $\boldsymbol{\delta}_{t,s,c}$ captures the Relative Demand Response Deviations (RDRD) from the planned response. In a general presentation, we will consider the vector of relative deviations $\boldsymbol{\delta} \in \mathbb{R}^d$, to be the vector that is obtained from arranging the parameters $\boldsymbol{\delta}_{t,s,c}$ for all $t \in \mathbb{T}$, $s \in \mathbb{S}$ and $c \in C^U$ into a vector. For simplicity of exposure, we assume that the vector of each seasonal subset of RDRD lies in a scaled budgeted uncertainty set (as introduced in Bertsimas and Sim 2004) defined as:

$$\Delta = \left\{ \boldsymbol{\delta} \in \mathbb{R}^d \mid \boldsymbol{\zeta} \in [-1, 1]^d, \boldsymbol{\delta}_{t,s,c} = \beta_{t,s,c} \boldsymbol{\zeta}_{t,s,c}, \beta_{t,s,c} \in C^U, \sum_{s \in \mathbb{S}} \sum_{c \in C^U} j \boldsymbol{\zeta}_{t,s,c} \leq \Gamma_{t,j} \forall t, j \in \mathbb{J} \right\}, \quad (2)$$

where $d := \sum_{t \in \mathbb{T}} \sum_{s \in \mathbb{S}} \sum_{c \in C^U} j$, $\beta_{t,s,c}$ captures the maximum possible RDRD, while $\Gamma_{t,j}$ is a seasonal budget that represents the fact that we expect a maximum of $\Gamma_{t,j}$ RDRD to take on their extreme values in season j of time period t . The idea of decomposing the uncertainty information structure over each season has three important advantages. First, from a numerical perspective, it will help by enabling us to employ decomposition schemes based on Benders decomposition. Second, it is also attractive from a statistical perspective as it allows us to calibrate the size of each seasonal uncertainty set using seasonal data which is usually readily available, while data over joint realizations over many seasons can be scarce. Finally, imposing uncertainty budgets on subsets of perturbed parameters is known to lead to less conservative solutions in robust optimization as it enforces the worst-case scenario to distribute the damage over parameters of different subsets instead of focusing on a single

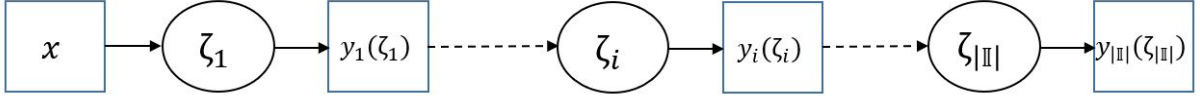


Figure 3: Sequence of decisions and uncertainty observations in our multi-period problem

one (see Bertsimas and Thiele 2006, where similar techniques are used).

In what follows, we present first, in Section 4.1, how we modify ETEM in order to identify capacity and demand response plans that are immunized against RDRD. The modified model is cast as a robust multi-period linear program. Unfortunately, this class of problems is known to be generally intractable. To address this issue, Section 4.2 presents a tractable conservative approximation that can be reformulated as a large-scale linear program.

4.1 A Robust Multi-Period Linear Programming Formulation

With the introduction of the RDRD in ETEM, the problem has the potential of turning into a multi-stage decision-making problem where different decisions are allowed to depend on the observed realization of RDRD. Indeed, we will assume that in the first stage the energy network planner decides on the capacity expansion of the system (variable \mathbf{C}) and the planned demand response (variable \mathbf{V}). Then, in later stages, on a season-by-season basis, the seasonal RDRD are revealed to the planner who responds by deciding on the optimal energy production, energy transfers between regions, and energy imports and exports. To simplify the analysis of such a multi-stage model, we will further assume that the planner’s response only depends on the current seasonal RDRD instead of reacting to all the observations made since the first season.

To be mathematically precise and help with presentation, decision variable $\mathbf{x} \in \mathbb{R}^m$, with $m = \sum_j \sum_l \sum_p j + \sum_j \sum_s j + \sum_j \sum_c j$, will capture the vector of all first stage decisions (\mathbf{C}, \mathbf{V}) in the demand response robust version of ETEM. Furthermore, for each time period t and season j , let $i \in \mathcal{I}$ denote the seasonal period (t, j) , and consider that $\mathbf{y}_i \in \mathbb{R}^{k_i}$, with $k_i = \sum_j \sum_l \sum_p j + \sum_j \sum_l \sum_c j + \sum_j \sum_l \sum_c j + \sum_j \sum_l \sum_c j + \sum_j \sum_l \sum_c j$, captures all adaptive “procurement” decisions $(\mathbf{P}_{t,s,l,p,c}, \mathbf{I}_{t,s,l,c}, \mathbf{E}_{t,s,l,c}, \mathbf{T}_{t,s,l,l^0,c})_{t,s \in \mathcal{S}_j, l \in \mathcal{L}, l^0 \in \mathcal{L}}$ implemented in season j and time period t . For simplicity, we will refer to seasonal periods as $i \in \mathcal{I}$ instead of $(t, j) \in \mathcal{T} \times \mathcal{J}$ and consider $A_i := \{(t, s) \in \mathcal{S}_j\}$ as the set of all (t, s) pairs associated to seasonal period i . Finally, we refer to $\boldsymbol{\zeta}_i \in [1, 1]^{d_i}$ as the vector of seasonal RDRD perturbances $f_{\boldsymbol{\zeta}_{t,s,c} \mathcal{G}(t,s) \mathcal{A}_i, c} \mathcal{U}$.

The chronology of our multi-period decision-making model is depicted in Figure 3. In particular, this model can be shown to take the following form:

$$\min_{\mathbf{x}, f, \mathbf{y}_i, \mathbf{g}_{i=1}^{j|j}} \max_{f, \mathbf{g}_{i=1}^{j|j}} f^T \mathbf{x} + \sum_{i=1}^{j|j} h_i^T \mathbf{y}_i(\boldsymbol{\zeta}_i) \quad (3a)$$

$$\text{s.t. } A_i \mathbf{x} + B_i \mathbf{y}_i(\boldsymbol{\zeta}_i) = \mathbf{b}_i + C_i \boldsymbol{\zeta}_i, \quad \boldsymbol{\zeta}_i \in Z_i(\Gamma_i), \quad \delta_i \in \{1, \dots, j|j\} \quad (3b)$$

$$D\mathbf{x} \leq \mathbf{e}, \quad (3c)$$

where constraint (3c) describes the polyhedral feasible set for \mathbf{C} and \mathbf{V} based on X and constraints (1f) and (1g), while constraint (3b) describes for each i the constraints imposed on $f_{\mathbf{P}_{t,s,l,p,c}, \mathbf{I}_{t,s,l,c}, \mathbf{E}_{t,s,l,c}, \mathbf{T}_{t,s,l,l^0,c} \mathcal{G}(t,s) \mathcal{A}_i, (l,l^0) \in \mathcal{L}^2}$. Namely, the latter includes the constraints in \mathcal{Y} , which in Babonneau et al. (2017) decompose over each season, and constraints (1b)-(1e), (1h)-(1m). Finally, $Z_i(\Gamma_i) := \{ \boldsymbol{\zeta}_i \in \mathbb{R}^{d_i} \mid k \boldsymbol{\zeta}_i \leq \mathbf{1}, \quad \mathbf{1} \leq k \boldsymbol{\zeta}_i \leq \mathbf{1} \}$ is the traditional budgeted uncertainty set while $C_i \in \mathbb{R}^{n_i \times d_i}$

models the linear effect of the standardized perturbances on each constraint. For completeness, we note that $f \in \mathbb{R}^m$, $h_i \in \mathbb{R}^{k_i}$, $A_i \in \mathbb{R}^{n_i \times m}$, $B_i \in \mathbb{R}^{n_i \times k_i}$, and $b_i \in \mathbb{R}^{n_i}$.

4.2 Affinely Adjustable Approximation

Given the known numerical intractability of problem (3) (see Ben-Tal et al. 2004), one can instead solve a conservative approximation model obtained by enforcing that adjustable variables $\mathbf{y}_i(\zeta_i)$ be affine functions of ζ_i ; i.e., $\mathbf{y}_i(\zeta_i) = \mathbf{y}_i + \mathbf{Y}_i \bar{\zeta}_i$, where $\mathbf{y}_i \in \mathbb{R}^{k_i}$ and $\mathbf{Y}_i \in \mathbb{R}^{k_i \times r_i}$ become the decision variables. We follow this approach after employing a commonly used lifting of the uncertainty set:

$$Z_i = P_i \bar{Z}_i,$$

where

$$\bar{Z}_i = \{ \bar{\zeta}_i \in \mathbb{R}^{r_i} \mid W_i \bar{\zeta}_i - \nu_i, \bar{\zeta}_i \geq 0 \},$$

with

$$W_i := \begin{bmatrix} I_{d_i} & I_{d_i} \\ \mathbf{1}_{1 \times d_i} & \mathbf{1}_{1 \times d_i} \end{bmatrix}, \quad \nu_i := \begin{bmatrix} \mathbf{1}_{d_i} \\ \Gamma_i \end{bmatrix}, \quad P_i := \begin{bmatrix} I_{d_i} & I_{d_i} \end{bmatrix},$$

and I_{d_i} is the $d_i \times d_i$ identity matrix, $r_i = 2d_i$. In other words, problem (3) is conservatively approximated by the following robust linear program (referred as the Robust Multi-Period Conservative Approximation model):

$$\text{(RMPCA)} \quad \min_{\mathbf{x}, f, \mathbf{Y}_i, \mathbf{y}_i, g_{i=1}^{j|j}} \max_{f \in Z_i, g_{i=1}^{j|j}} f^\top \mathbf{x} + \sum_{i=1}^{j|j} h_i^\top (\mathbf{y}_i + \mathbf{Y}_i \bar{\zeta}_i) \quad (4a)$$

$$\text{s.t.} \quad A_i \mathbf{x} + B_i (\mathbf{y}_i + \mathbf{Y}_i \bar{\zeta}_i) \leq b_i + C_i P_i \bar{\zeta}_i \quad \forall \bar{\zeta}_i \in \bar{Z}_i, \forall i \in \{1, \dots, j|j\} \quad (4b)$$

$$D \mathbf{x} \leq e. \quad (4c)$$

Applying standard robust reformulation techniques, we obtain the following tractable linear program:

$$\min_{\mathbf{x}, f, \Phi_i, \mathbf{Y}_i, \mathbf{y}_i, g_{i=1}^{j|j}} f^\top \mathbf{x} + \sum_{i=1}^{j|j} \nu_i^\top \psi_i + h_i^\top \mathbf{y}_i \quad (5a)$$

$$\text{s.t.} \quad D \mathbf{x} \leq e \quad (5b)$$

$$W_i^\top \psi_i - (h_i^\top \mathbf{Y}_i)^\top \leq \delta_i \quad \forall i \in \{1, \dots, j|j\} \quad (5c)$$

$$\Phi_i \nu_i \leq b_i - A_i \mathbf{x} - B_i \mathbf{y}_i \quad \forall i \in \{1, \dots, j|j\} \quad (5d)$$

$$\Phi_i W_i - B_i \mathbf{Y}_i - C_i P_i \leq \delta_i \quad \forall i \in \{1, \dots, j|j\} \quad (5e)$$

$$\Phi_i, \psi_i \geq 0, \quad (5f)$$

where $\psi_i \in \mathbb{R}^{(d_i+1) \times 1}$ and $\Phi_i \in \mathbb{R}^{n_i \times (d_i+1)}$ are new variables added to the model to guarantee the feasibility of the decision variables in the worst-case scenario. It is worth mentioning that problem (5) provides an upper-bound and feasible first-stage decision for problem (3) as the space of decision rule has been reduced to affine functions.

5 Improving Numerical Efficiency using Benders Decomposition

Because of the capacity expansion decisions with a horizon of several decades, alongside with the procurement decisions with daily resolution, ETEM is already in its deterministic form a large-scale LP model. The size of LP that needs to be solved is further exacerbated when considering the robust reformulation presented in (5). In this regard, Section 6 will consider instances where the number of decisions and constraints is increased twenty-fold and nine-fold respectively. For this reason, we propose here a decomposition scheme to reduce the computational burden.

Specifically, we start by reformulating problem (4) as follows:

$$\min_{\mathbf{x}, \mathbf{y}_i, g_{i=1}^{j|j}} f^> \mathbf{x} + \sum_{i=1}^{j|j} h_i^> \mathbf{y}_i + \rho_i \quad (6a)$$

$$\text{s.t. } D\mathbf{x} = \mathbf{e} \quad (6b)$$

$$\rho_i = g_i(\mathbf{x}, \mathbf{y}_i) \quad \delta_i \geq \mathbb{1} \dots j|jg \quad (6c)$$

$$A_i \mathbf{x} + B_i \mathbf{y}_i = \mathbf{b}_i \quad \delta_i \geq \mathbb{1} \dots j|jg, \quad (6d)$$

with

$$g_i(\mathbf{x}, \mathbf{y}_i) := \min_{\mathbf{Y}_i} \max_{\mathbf{Z}_i, g_{i=1}^{j|j}} h_i^> \mathbf{Y}_i \bar{\boldsymbol{\zeta}}_i \quad (7a)$$

$$\text{s.t. } (4b),$$

where constraint (6d) is a redundant constraint that describes the fact that (4b) must be satisfied for $\bar{\boldsymbol{\zeta}}_i = 0$ since $0 \geq \bar{Z}_i$. Given that $g_i(\mathbf{x}, \mathbf{y}_i)$ is jointly convex, we identify a supporting plane representation that will enable the use of a constraint generation approach, also commonly referred to as a Benders decomposition scheme, which relies on iteratively solving a master and set of subproblems until convergence. Note that constraint (6c) should be considered violated if \mathbf{x} and \mathbf{y}_i are such that problem (7) is infeasible.

In what follows, we start by describing the algorithm, then present two schemes that can be employed to improve its numerical efficiency by refining the quality of the upper and lower bounding process. In the latter case, the refinement will make use for the first time of the notion of Pareto robustly optimal solution (Iancu and Trichakis 2014) in decomposition schemes for robust optimization.

Remark 1. *It is worth noting that our scheme is strongly inspired by the one used in Ardestani-Jaafari and Delage (2018) and Ardestani-Jaafari and Delage (2020), yet we depart from that work in three ways. First, unlike this prior work we choose to keep in problem (6) the \mathbf{y}_i variables, which describe what the a ne strategy prescribes for the nominal scenario $\bar{\boldsymbol{\zeta}} = 0$. This allows us to more easily interpret the optimal solution. We also observed empirically that this led to a significant reduction in the number of iterations. Second, we derive conditions that are specific to this application under which we can ensure that the solution of the master's problem always leads to a set of feasible subproblems. Third, we identify a new scheme, which is based on the theory of Pareto robust optimality, to identify optimal solutions of the master problem that accelerate the convergence of the algorithm.*

5.1 The Benders Decomposition Algorithm

To simplify the presentation of this algorithm we make the following assumption.

Assumption 1. (Uncapacitated technology) ETEM includes a technology $p_x \in \mathcal{P}_c^P/\mathcal{P}^R$ with $c \in \mathcal{C}^U$ generated from a resource $c_x \in \mathcal{C}^I/\mathcal{C}^D$, with infinite capacity, i.e., $\Omega_{t,l,p_x} = 1$. Furthermore, X and Y do not impose additional constraints on $\mathbf{P}_{t,s,l,p_x,c}$ and \mathbf{I}_{t,s,l,c_x} .

The role of the above assumption is to ensure that when removing constraint (6c) from problem (6), we do not end up with solutions for $(\mathbf{x}, \bar{\mathbf{y}}_i \mathcal{G}_{i=1}^{l,j})$ that make problem (7) infeasible. Note also that in our description of ETEM, the EXT technology and ExR energy resource respectively play the role of p_x and c_x .

Lemma 1. Given that an uncapacitated technology exists (i.e., Assumption 1), problem (7) is feasible for any \mathbf{x} and $\bar{\mathbf{y}}_i \mathcal{G}_{i=1}^{l,j}$ that satisfy (6d).

Proof. One needs to show that for any given \mathbf{x} and \mathbf{y}_i , $i = 1, \dots, |J|$, that satisfy (6b) and (6d), there exists a \mathbf{Y}_i so that the constraint (4b) is feasible. To do so, given that we already know \mathbf{x} and \mathbf{y}_i are such that they can cover the nominal demand response, i.e., $\zeta^+ = \zeta^- = 0$, we will show how to construct a \mathbf{Y}_i that will simply cover any deviation of the demand coming from the RDRD using the uncapacitated technology. Specifically, we let all values of \mathbf{Y} equal to zero except for the terms that model the influence of δ^+ on $\mathbf{P}_{t,s,l,p_x,c}$ with $c \in \mathcal{C}^U$, $\mathbf{P}_{t,s,l,p_x,c_x}$, and \mathbf{I}_{t,s,l,c_x} . In particular, when setting all other terms to zero, most constraints reduce exactly to the constraints imposed in (6d) and are straightforwardly satisfied. The only constraints that need to be verified consist of a few members of (1b), (1e), (1h) and (1j), which can be described as follows:

$$(1b) \quad \left(\sum_{p \in \mathcal{P}^{c_x}} \bar{\mathbf{P}}_{t,s,l,p,c_x} + \bar{\mathbf{I}}_{t,s,l,c_x} + \sum_{c \in \mathcal{C}^U} \mathbf{I}_{t,s,l,c_x,c}^+ \delta_{t,s,c}^+ \right) \eta_{c_x} + \sum_{l^0 \in \mathcal{L}} (\eta_{c_x} \bar{\mathbf{T}}_{t,s,l^0,l,c_x} + \bar{\mathbf{T}}_{t,s,l,l^0,c_x}) \sum_{p \in \mathcal{P}_c^+} \bar{\mathbf{P}}_{t,s,l,p,c}^+ + \sum_{c \in \mathcal{C}^U} \mathbf{P}_{t,s,l,p_x,c_x}^+ \delta_{t,s,c}^+ + \bar{\mathbf{E}}_{t,s,l,c_x} \quad \begin{array}{l} \delta(\delta_i^+, \delta_i^-) \geq \bar{Z}_i, \forall l \in \mathcal{L} \\ \delta(t, s) \geq A_i, \forall i \in \mathcal{I} \end{array} \quad (8)$$

$$(1e) \quad \sum_{p \in \mathcal{P}_c^+} \bar{\mathbf{P}}_{t,s,l,p,c} + \mathbf{P}_{t,s,l,p_x,c}^+ \delta_{t,s,c}^+ \quad \Theta_{t,l,c}(\bar{\mathbf{V}}_{t,s,c} + \delta_{t,s,c}^+ \quad \delta_{t,s,c}) \quad \begin{array}{l} \delta(\delta_i^+, \delta_i^-) \geq \bar{Z}_i, \forall c \in \mathcal{C}^U \\ \delta(t, s) \geq A_i, \forall i \in \mathcal{I} \end{array} \quad (9)$$

$$(1h) \quad \sum_{c: p_x \in \mathcal{P}_c^+} \bar{\mathbf{P}}_{t,s,l,p_x,c} + \sum_{c \in \mathcal{C}^U} \mathbf{P}_{t,s,l,p_x,c}^+ \delta_{t,s,c}^+ \quad 1 \quad \begin{array}{l} \delta(\delta_i^+, \delta_i^-) \geq \bar{Z}_i, \forall l \in \mathcal{L} \\ \delta(t, s) \geq A_i, \forall i \in \mathcal{I} \end{array} \quad (10)$$

$$(1j) \quad \sum_{c \in \mathcal{C}_{f^0}} \bar{\mathbf{P}}_{t,s,l,p_x,c} + \sum_{c \in \mathcal{C}^U} \mathbf{P}_{t,s,l,p_x,c}^+ \delta_{t,s,c}^+ = \eta_{f^0}^t \left(\sum_{c \in \mathcal{C}_f} \bar{\mathbf{P}}_{t,s,l,p_x,c} + \sum_{c \in \mathcal{C}^U} \mathbf{P}_{t,s,l,p_x,c_x}^+ \delta_{t,s,c}^+ \right) \quad \begin{array}{l} \delta(\delta_i^+, \delta_i^-) \geq \bar{Z}_i, \forall l \in \mathcal{L} \\ \delta(t, s) \geq A_i, \forall i \in \mathcal{I} \end{array}, \quad (11)$$

where $\bar{\mathbf{P}}$, $\bar{\mathbf{V}}$, $\bar{\mathbf{C}}$, $\bar{\mathbf{I}}$, $\bar{\mathbf{T}}$, and $\bar{\mathbf{E}}$ refer to the assignments made through the fixed \mathbf{x} and \mathbf{y}_i 's, while $\mathbf{P}_{t,s,l,p_x,c}^+$, $\mathbf{P}_{t,s,l,p_x,c_x}^+$, and $\mathbf{I}_{t,s,l,c_x,c}^+$ model the influence of $\delta_{t,s,c}^+$ on $\mathbf{P}_{t,s,l,p_x,c}$, $\mathbf{P}_{t,s,l,p_x,c_x}$, and \mathbf{I}_{t,s,l,c_x} respectively, for $c \in \mathcal{C}^U$. In order to get a feasible assignment for \mathbf{Y} , for all $t, s, c \in \mathcal{C}^U$, we let $\mathbf{P}_{t,s,l,p_x,c}^+ := \Theta_{t,l,c}$, $\mathbf{P}_{t,s,l,p_x,c_x}^+ := \mathbf{P}_{t,s,l,p_x,c}^+ / \eta_{f^0}^t$, and $\mathbf{I}_{t,s,l,c_x,c}^+ = \mathbf{P}_{t,s,l,p_x,c_x}^+ / \eta_{c_x}$. One can readily verify that constraints (8), (10), and (11) reduce to the constraint already accounted for in (6d). On the other hand, constraint (9) reduces to:

$$\sum_{p \in \mathcal{P}_c^+} \bar{\mathbf{P}}_{t,s,l,p,c} \quad \Theta_{t,l,c}(\bar{\mathbf{V}}_{t,s,c} + \delta_{t,s,c}^+ \quad \delta_{t,s,c}) \quad \begin{array}{l} \delta(\delta_i^+, \delta_i^-) \geq \bar{Z}_i, \forall c \in \mathcal{C}^U \\ \delta(t, s) \geq A_i, \forall i \in \mathcal{I} \end{array},$$

which is necessarily satisfied since $\Theta_{t,l,c} = 0$ and $\delta_i = 0$. A similar argument can be used to confirm that all non-negative constraints on $\bar{P}_{t,s,l,p_x,c} + P_{t,s,l,p_x,c}^+ \delta_{t,s,c}^+$, $\bar{P}_{t,s,l,p_x,c_x} + \sum_{c \in \mathcal{C} \cup \mathcal{C}^+} P_{t,s,l,p_x,c_x}^+ \delta_{t,s,c}^+$, and $\bar{I}_{t,s,l,c_x} + \sum_{c \in \mathcal{C} \cup \mathcal{C}^+} I_{t,s,l,c_x}^+ \delta_{t,s,c}^+$ are also satisfied. \square

In practice, if Assumption 1 is not satisfied, a few remedies exist. First, one can create an artificial technology with either infinite or very large capacity, whose price is set to an arbitrarily large amount. One can interpret the price of this artificial technology as the marginal cost of shortages in the system. If this technology is not used in the optimal solution that is identified, then shortages do not occur and the solution is optimal for the problem without this artificial technology. Alternatively, one could also modify the algorithm that is presented below to have it identify ‘‘feasibility cuts’’ when the current candidate solution makes (7) infeasible (see Rahmaniani et al. 2017, for details).

We now focus on presenting a support plane representation of $g(\mathbf{x}, \mathbf{y}_i)$.

Theorem 2. *Given that Assumption 1 is satisfied,*

$$g_i(\mathbf{x}, \mathbf{y}_i) = \max_{\theta_i, \nu_i} (b_i + A_i \mathbf{x} + B_i \mathbf{y}_i)^\top \theta_i - \text{Tr}(C_i P_i \boldsymbol{\lambda}_i) \quad (12a)$$

$$\text{s.t. } \theta_i \nu_i \geq \boldsymbol{\lambda}_i W_i \geq 0 \quad (12b)$$

$$B_i^\top \boldsymbol{\lambda}_i = h_i \bar{\boldsymbol{\zeta}}_i \quad (12c)$$

$$W_i \bar{\boldsymbol{\zeta}}_i \leq \nu_i, \quad (12d)$$

where $\theta_i \in \mathbb{R}^{n_i - 1}$, $\boldsymbol{\lambda}_i \in \mathbb{R}^{n_i - r_i}$ and $\bar{\boldsymbol{\zeta}}_i \in \mathbb{R}^{r_i - 1}$.

Proof. In the first step, we identify the robust counterpart of constraint (4b). For each $i \in \mathcal{I}$, given that \bar{Z}_i is non-empty, LP duality applies on each robust constraint thus introducing auxiliary variables that we denote in matrix form in Φ_i . Namely, constraint (4b) can be said equivalent to

$$\forall \Phi_i \in \mathbb{R}^{n_i \times (d_i + 1)}, \quad \forall \delta_i \in \mathbb{F}_1 \quad \forall j \in \mathcal{J} \quad (13)$$

$$\Phi_i \nu_i \leq b_i + A_i \mathbf{x} + B_i \mathbf{y}_i \quad \forall \delta_i \in \mathbb{F}_1 \quad \forall j \in \mathcal{J} \quad (13)$$

$$\Phi_i W_i \leq B_i \mathbf{Y}_i + C_i P_i \quad \forall \delta_i \in \mathbb{F}_1 \quad \forall j \in \mathcal{J}. \quad (14)$$

Therefore, problem (7) can be reformulated as:

$$\min_{\mathbf{Y}_i, \Phi_i} \max_{\boldsymbol{\zeta}_i \in \bar{Z}_i} (h_i^\top \mathbf{Y}_i \bar{\boldsymbol{\zeta}}_i) \\ \text{s.t. } (13), (14),$$

which is a feasible minimization problem according to Lemma 1 since we assumed that Assumption 1 holds. In the next step, since \bar{Z}_i is a compact and convex set, Sion’s minmax theorem (Sion 1958) holds and one can reverse the order of minimization over \mathbf{Y}_i, Φ_i and maximization over $\bar{\boldsymbol{\zeta}}_i$ ’s. Finally, since the set defined by (13) and (14) is feasible, one can employ LP duality to replace the inner minimization problem with its dual maximization problem to obtain:

$$g(\mathbf{x}, \mathbf{y}_i) = \max_{\boldsymbol{\zeta}_i \in \bar{Z}_i, \theta_i, \nu_i} (b_i + A_i \mathbf{x} + B_i \mathbf{y}_i)^\top \theta_i - \text{Tr}(C_i P_i \boldsymbol{\lambda}_i) \\ \text{s.t. } \theta_i \nu_i \geq \boldsymbol{\lambda}_i W_i \geq 0 \\ B_i^\top \boldsymbol{\lambda}_i = h_i \bar{\boldsymbol{\zeta}}_i$$

$$\boldsymbol{\theta}_i \succeq 0, \boldsymbol{\lambda}_i \succeq 0,$$

where $\boldsymbol{\lambda}_i \in \mathbb{R}^{n_i - r_i}$, $\boldsymbol{\theta}_i \in \mathbb{R}^{n_i - 1}$ are dual variables associated with constraints (13) and (14) respectively. \square

Equipped with Theorem 2, we can define the following Benders decomposition algorithm. Intuitively, the procedure consists in solving a so-called master problem in which constraint (6c) is removed thus producing a sequence of lower bounds for problem (4). Given a current optimal solution of the master problem, the latter is refined by progressively reintroducing the constraints that are the most violated from the set:

$$\boldsymbol{\rho}_i = \arg \min_{\boldsymbol{\rho}_i} (b_i + A_i \mathbf{x} + B_i \mathbf{y}_i)^\top \boldsymbol{\theta}_i - \text{Tr}(C_i P_i \boldsymbol{\lambda}_i), \quad \delta(\boldsymbol{\lambda}_i, \boldsymbol{\theta}_i) \geq f(\boldsymbol{\lambda}_i, \boldsymbol{\theta}_i) \geq \mathbb{R}_+^{n_i - r_i} \quad \mathbb{R}_+^{n_i} \setminus \mathcal{Q}_{\bar{\boldsymbol{\zeta}}_i} \geq \bar{\boldsymbol{Z}}_i, \quad (12b), (12c)g, \quad \delta_i \geq 1.$$

For each i , the most violated constraint can be found by solving the LP presented in (12), which optimal value can be used to update an upper bound problem (4). We refer the reader to the pseudo-code presented in Algorithm 1.

Algorithm 1: Benders Decomposition (BD) algorithm

- 1 Set $UB = \infty, LB = -\infty, v = 0$;
 - 2 Let i be the index for each subproblem;
 - 3 Let $v = v + 1$ and solve the deterministic problem, i.e., problem (4) when $\bar{\boldsymbol{Z}}_i = f_0 g$, and store the value of \mathbf{x} as $\bar{\mathbf{x}}^{(v)}$ and \mathbf{y}_i as $\bar{\mathbf{y}}_i^{(v)}$;
 - 4 **while** $(UB - LB)/LB \geq \varepsilon$ **do**
 - 5 $\delta_i = f_1 - \text{Tr}(C_i P_i \bar{\boldsymbol{\lambda}}_i^{(v)}, \bar{\boldsymbol{\theta}}_i^{(v)})$ solve problem (12) with $(\bar{\mathbf{x}}^{(v)}, \bar{\mathbf{y}}_i^{(v)})$ and store the optimal value in $\boldsymbol{\rho}_i^*$ and decision variables in $\bar{\boldsymbol{\theta}}_i^v, \bar{\boldsymbol{\lambda}}_i^v, \bar{\boldsymbol{\zeta}}_i^v$;
 - 6 Update upper bound $UB = \min(UB, f^\top \bar{\mathbf{x}}^v + \sum_{i=1}^{|I|} \boldsymbol{\rho}_i^* + h_i^\top \bar{\mathbf{y}}_i^v)$;
 - 7 Solve the updated master problem :

$$\min_{\mathbf{x}, \{\boldsymbol{\rho}_i, \mathbf{y}_i\}_{i=1}^{|I|}} f^\top \mathbf{x} + \sum_{i=1}^{|I|} h_i^\top \mathbf{y}_i + \boldsymbol{\rho}_i \tag{15a}$$

s.t. (6b), (6d),

$$\boldsymbol{\rho}_i = \arg \min_{\boldsymbol{\rho}_i} (b_i + A_i \mathbf{x} + B_i \mathbf{y}_i)^\top \boldsymbol{\theta}_i - \text{Tr}(C_i P_i \bar{\boldsymbol{\lambda}}_i^{v'}), \quad \delta_{v'} \geq v, \delta_i \geq f_1 - \text{Tr}(C_i P_i \bar{\boldsymbol{\lambda}}_i^{v'}). \tag{15b}$$
 - 8 Let $v = v + 1$ and store the optimal value of the master problem variables in $\bar{\mathbf{x}}^v, \bar{\mathbf{y}}_i^v$ and $\bar{\boldsymbol{\rho}}^v$;
 - 8 Update the lower bound: $LB = f^\top \bar{\mathbf{x}}^v + \sum_i h_i^\top \bar{\mathbf{y}}_i^v + \bar{\boldsymbol{\rho}}^v$.
 - 9 **end**
-

5.2 Improving the Algorithm's Convergence

In order to improve the algorithm convergence, we apply two strategies. First, as proposed in Ardestani-Jaafari and Delage (2018) and Ardestani-Jaafari and Delage (2020), we add a set of valid inequalities to the master problem (15) based on violated scenarios from previous iterations in order to tighten the lower bounding problem. Specifically, for any $i \in I$ and finite set of scenarios $\mathcal{F}_{\bar{\boldsymbol{\zeta}}_i^l} \subseteq \bar{\boldsymbol{Z}}_i$, the following constraint necessarily holds in problem (6):

$$\delta_l \geq \Omega, \quad \mathcal{Q}_{\bar{\boldsymbol{\zeta}}_i^l} \subseteq \mathbb{R}^{d_i}, \quad \boldsymbol{\rho}_i = h_i^\top (\mathbf{y}_i^l - \mathbf{y}_i) \ \& \ A_i \mathbf{x} + B_i \mathbf{y}_i^l - b_i + C_i P_i \bar{\boldsymbol{\zeta}}_i^l,$$

given that for all $l \in \Omega$:

$$\begin{aligned} \rho_i \quad g(\mathbf{x}, \mathbf{y}_i) &= \min_{Y_i: (4b)} \max_{i \in \mathcal{Z}_i} h_i^{\triangleright} \mathbf{Y}_i \bar{\zeta}_i^l \\ &= \min_{Y_i^l, Y_i: Y_i^l = y_i + Y_i^l} \max_{i \in \mathcal{Z}_i} h_i^{\triangleright} (\mathbf{y}_i^l \quad \mathbf{y}_i) \end{aligned}$$

where \mathbf{y}_i^l plays the role of an adjustable plan specifically designed for scenario $\bar{\zeta}_i^l$ and which can depart from the affine decision rule. Finally, In our implementation, we choose to set $\bar{\zeta}_i^l \in \mathcal{G}_{l \geq 2}$ to be the violated scenarios in the last v iterations of the algorithm. For completeness, we include the new master problem below:

$$\min_{\mathbf{x}, \mathbf{f}} \quad f^{\triangleright} \mathbf{x} + \sum_{i=1}^{j|j} h_i^{\triangleright} \mathbf{y}_i + \rho_i \quad (16a)$$

$$\text{s.t.} \quad (6b), (6d), (15b),$$

$$\rho_i \quad h_i^{\triangleright} (\mathbf{y}_i^l \quad \mathbf{y}_i) \quad \forall i \in \Omega, \forall l \in \mathcal{L} \quad j|jg \quad (16b)$$

$$A_i \mathbf{x} + B_i \mathbf{y}_i^l \quad b_i + C_i P_i \bar{\zeta}_i^l \quad \forall i \in \Omega, \forall l \in \mathcal{L} \quad j|jg. \quad (16c)$$

The second proposed improvement has to do with how the upper bound is obtained. In particular, this bound comes from evaluating the true worst-case cost of the current best candidate based on the master problem (15), or its tighter version (16). Yet, since the master problem is itself a robust optimization problem, we can expect based on the findings of Iancu and Trichakis (2014) that at each iteration there exists a large set of optimal solutions for (15). In a traditional Benders decomposition approach (such as in Ardestani-Jaafari and Delage 2018), the candidate that is used to calculate the upper bound and generate an optimality cut is arbitrarily chosen by the LP solver used to solve (15). While such a solution is optimal for (15), it could have a much worst performance in problem (12) compared to other optimal solutions. We therefore recommend, after solving (15), to identify a Pareto robustly optimal solution of (15) by solving the following LP:

$$\min_{\mathbf{x}, \mathbf{f}} \quad f^{\triangleright} \mathbf{x} + \sum_{i=1}^{j|j} h_i^{\triangleright} \mathbf{y}_i + \rho_i^0 \quad (17a)$$

$$\text{s.t.} \quad (6b), (6d), (15b), (16b), (16c),$$

$$f^{\triangleright} \mathbf{x} + \sum_{i=1}^{j|j} h_i^{\triangleright} \mathbf{y}_i + \rho_i \quad (1 + \epsilon) \cdot \mathcal{M}_v \quad (17b)$$

$$\rho_i^0 \quad \sum_{v^0} \frac{1}{v} ((b_i + A_i \mathbf{x} + B_i \mathbf{y}_i)^{\triangleright} \bar{\theta}_i^{v^0} \quad Tr(C_i P \bar{\lambda}_i^{v^0})) \quad \forall i \in \mathcal{L} \quad j|jg, \quad (17c)$$

where \mathcal{M}_v is the optimal solution of the master problem (15) in iteration v of Algorithm 1, and $\epsilon > 0$ allows for some small ϵ sub-optimality in order to help numerically. The new candidate $(\bar{\mathbf{x}}^v, \bar{\mathbf{y}}_i^v)$ remains approximately optimal for (15) yet is guaranteed to not be Pareto dominated by other solutions of (15). This follows from the fact that constraint (17c) is equivalent to:

$$\rho_i^0 \quad (b_i + A_i \mathbf{x} + B_i \mathbf{y}_i)^{\triangleright} \left(\sum_{v^0} \frac{1}{v} \bar{\theta}_i^{v^0} \right) \quad Tr(C_i P \left(\sum_{v^0} \frac{1}{v} \bar{\lambda}_i^{v^0} \right)) \quad \forall i \in \mathcal{L} \quad j|jg,$$

and that $(\sum_{v^0} \frac{1}{v} \bar{\theta}_i^{v^0}, \sum_{v^0} \frac{1}{v} \bar{\lambda}_i^{v^0})$ is in the relative interior of the convex hull of $f(\bar{\theta}_i^{v^0}, \bar{\lambda}_i^{v^0})_{g_{v^0}}$. Hence,

Corollary 1 in Iancu and Trichakis (2014) can be applied to obtain the Pareto non-dominance guarantee. We refer the reader to Iancu and Trichakis (2014) for more details about how to identify Pareto robustly optimal solutions.

6 Computational Results

In this section, we evaluate the performance of the proposed robust ETEM approximation model (i.e., RMPCA) on a case study based on the energy system of the Arc Lémanique region in Switzerland. To do so, we compare the performance of the RMPCA formulation with a static robust model (SRM), which considers all decisions as *here-and-now* decisions, and a deterministic model (DET), which disregards demand response uncertainty. The purpose of these computational experiments is to empirically show that: i) considering DRU in a capacity expansion planning problem and solving it using robust optimization can decrease both worst-case and expected total costs of the system compared to a deterministic formulation of the problem; and ii) assuming flexibility of the procurement decisions with respect to observed actual DR decreases the level of conservatism of the model and consequently the expected total cost of the system. Recall that in RMPCA the planner first decides on the capacities and planned DR, then in seasonal periods, after observing the actual DR, he decides on the optimal procurement of energy. Theoretically, it is known that the policies proposed by SRM are more conservative compared to policies proposed by RMPCA. In addition, we present a detailed comparison of the structure of the RMPCA, SRM, and DET policies. The purpose is to understand how considering the uncertainty of the demand response affects the long-term capacity expansion strategies. Finally, we provide a computational analysis to compare the performance of the different decomposition algorithms presented in Section 5. All numerical studies are performed on a 64-bit computer with 128 GB of RAM and Intel(R) Xeon(R) 3.1 GHz (40 CPUs). The deterministic version of ETEM is originally written in AMPL. However, we extract the standard form matrices from AMPL and carry out robust optimization using the YALMIP toolbox and GUROBI solver in a MATLAB R2019a environment.

6.1 A Swiss Case Study

The ETEM model describing the Arc Lémanique region (Cantons of Geneva and Vaud, in Switzerland) encompasses 142 different technologies, including centralized electricity and heat production plants, decentralized electricity production, conventional and flexible loads, and end-use transportation technologies. In addition, 57 energy commodities, including 21 types of useful demand, are modeled. We choose to immunize ETEM against residential electricity demand response deviations (i.e., Res-ELC in Figure 1), which implies that the robust model is subject to 120 independent sources of perturbations. This choice is motivated by the fact that the response behavior of residential electricity is more prone to uncertainty in the long-run compared to industrial and large-scale electricity consumers, as the former corresponds to a larger and more diverse group of consumers. Table 1 presents techno-economic features of both presently available and expected future electricity generation technologies in the Arc Lémanique region. As we wish to avoid as much as possible energy shortages in the system, we set the cost of the additional energy source (ExR) high enough so that there is no capacity shortage in the optimal solution of RMPCA model when the largest amount of deviations are assumed. Finally, in order to better investigate the influence of robust optimization on the solutions, no upper bound limits the installation of renewable technologies (wind and photovoltaic power plants) in the model. For interested readers, we note that our deterministic version of ETEM presents 171,620 variables and 431,652 constraints while the LP reformulation of RMPCA, i.e., problem (5), presents 3,540,720 variables and 3,708,658 constraints.

Table 1: Characteristics of electricity generation technologies considered in the Arc Lémanique region model

Category ¹	Technology	Status	Input ²	Output ³	Investment cost (MCHF/GW)	Fixed cost (MCHF/GW)	Variable cost (MCHF/PJ)	Efficiency	Life (year)
Gas-PP	Gas Turbine	existing	NGA	ELC,CO2	300	25	0.22	0.3	4
	Gas CC	new	NGA	ELC,CO2	450	25	0.55	0.6	5
Hydro-PP	Veytaux (turbine)	existing	HYD	ELC	-	130	0	0.89	10
	Hydroelectric VAUD	existing	HYD	ELC	-	130	0	0.39	10
	Hydroelectric GENEVA	existing	HYD	ELC	-	130	0	0.39	10
PV	Photovoltaic	new	SOL	ELC	6000-2000	25	0	0.39	6
Wind-pp	Wind	new	WND	ELC	1500-1250	40	0	0.4	5
Oil-Fired-PP	Oil-Fired Steam-Cycle	new	DSL	ELC,CO2	-	92	0.52	0.39	6
	Oil-Fired Steam-Cycle (CPD)	new	NGA,DSL	ELC,CO2	1100	68	0.55	0.4	7
CHP	Gas CC (CHP)	new	NGA	ELC,CO2	1000	46	0.55	0.85	7
	Centrale de Cheneviers (heat and electricity)	existing	MSW	ELC,CO2	1200	30	0.0	0.38	10
	Enerbois (Vaud)	existing	WOR	ELC,CO2	-	0	0.0	1	10
	Pierre-de-Plan (heat+electricity)	existing	NGA	ELC,CO2	450	25	0.55	0.9	5
	Industrial Cogen. Gas Turbine (5 MW)	existing	NGA	ELC,CO2	800	78	0	0.28	5
	Industrial Cogen. STEAM Turbine (5 MW)	existing	NGA	ELC,CO2	1000	208	0	0.15	5
	Industrial Combined Cycle CC (5 MW)	existing	NGA	ELC,CO2	1200	68	0	0.37	5
	Cogeneration INDUstry (Motor)	existing	NGA	ELC,CO2	4160	104	0	0.35	3
	Cogeneration COMM (Motor)	existing	NGA	ELC,CO2	5434	65	0	0.53	3
	Cogeneration RESID (Motor)	existing	NGA	ELC,CO2	6136	92	0	0.53	3
	Chatillon Plant (heat)	existing	WOR	ELC,CO2	-	-	-	1	10
	CHP combined heat production	new	NGA	ELC,CO2	1800	15	0.71	0.9	4
	CHP combined heat production	new	NGA	ELC,CO2	1200	12	0.71	0.9	4
	Other-PP	Geothermie (Vaud)	existing	HTH	ELC	-	-	-	1
Tridel (Vaud) Electricity		existing	MSW	ELC,CO2	-	-	-	0.38	10
Gas fuel cell		new	NGA	ELC	2000-1250	78	4.78	0.42	6

¹ Definition of technology categories is presented in Fig 1.

² NGA: Natural gas, HYD: Hydro power, SOL: Solar, WND: Wind, DSL: Diesel, MSW: Municipal solid waste, WOR: Wood residential, HTH: High temperature heat.

³ ELC: Electricity, CO2: CO₂ emissions.

While we refer the reader to Babonneau et al. (2017) for a complete description of all input parameters and assumptions, it is worth mentioning that we consider here a case where CO₂ emissions are curbed to 2.5 Mt in 2050, a 45% reduction compared to 2010 levels (5.48 Mt).

6.2 Performance Analysis

In this section, we study the performance of strategies proposed by RMPCA, SRM, and DET in the average and worst-case scenarios. For a specific budget size Γ and box size β , $\mathbf{x}_{RMPCA}^{\beta}$ represents the RMPCA strategy obtained by solving problem (5). Similarly, \mathbf{x}_{SRM}^{β} is the static robust strategy obtained by solving problem (5) when forcing $\mathbf{Y}_i = 0 \ \forall i \in \{1, \dots, J\}$. It should be noted that, because problem (5) is computationally demanding and cannot be solved in a 72-hour time limit, we have used the PRO-BD-VI algorithm proposed in Section 5. Finally, the DET strategy \mathbf{x}_{DET} is obtained by solving the deterministic model, i.e., problem (4) when $\bar{Z}_i = \bar{r}_0 \ \forall i \in \{1, \dots, J\}$. After obtaining all the strategies from the different models, in a next step, we generate a set of 1000 DR random scenarios $\zeta_i^j \ (\forall i \in \{1, \dots, J\}, j \in \{1, \dots, 1000\})$ using a uniform distribution on the $[\beta, \beta]$ interval. For each scenario, we then re-optimize problem (4) with the value of \mathbf{x} and ζ_i fixed to the corresponding strategy and random scenario. In the first period (i.e., $t = 1$), we have no uncertainty given that a full information about the level of DR is available. If, for a specific policy and scenario, problem (4) becomes infeasible, this implies a capacity shortage. Table 2 reports the proportion of scenarios with capacity shortage for the different strategies and different budget sizes (expressed in percentage of β and denoted as $\Gamma\%$). When the budget size increases, the chances of capacity shortage of RMPCA and SRM reduces. This is because a larger budget size increases the conservatism of the model, and consequently, the model installs more capacity to avoid shortage, which obviously is more expensive. Nevertheless, for budget sizes in the range of [50%, 60%], the capacity shortage of RMPCA can be as low as 2%. To have a fair comparison between the expected total cost of different strategies, we report the average total cost both over all scenarios and over the scenarios without

capacity shortage.

Table 2: Proportion of random scenarios with capacity shortage when $\beta = 0.6$.

$\Gamma\%$ ¹	RMPCA	SRM	DET
10%	100%	100%	100%
20%	98%	96%	100%
30%	29%	0%	100%
40%	9%	0%	100%
50%	4%	0%	100%
60%	2%	0%	100%
70%	10%	0%	100%
80%	0%	0%	100%
90%	0%	0%	100%
100%	0%	0%	100%

¹ $\Gamma\%$ denotes that the units are in percent of $|d_i|$.

Figure 4a presents statistics of the total cost for the RMPCA and SRM strategies, over all scenarios, for different budget sizes and $\beta = 0.6$. Figure 4b gives similar results when the average costs are taken over only scenarios without capacity shortage. We can first note that the average cost for the SRM strategy remains fixed for budget sizes greater than 25%. This is because when $\Gamma\% = 25\%$, $|d_i| = 1$, the SRM already considers each constraint to be maximally perturbed due to the fact that each one only involves one demand deviation. Second, we note that Figures 4 (a) and (b) show that the average RMPCA cost is considerably lower than the SRM cost for all budget sizes. This already suggests that the flexibility of energy procurement decisions can significantly reduce the expected total cost. More precisely, the maximum difference occurs when $\Gamma\% = 30\%$ yet this difference comes at the price of 29% chances of capacity shortage (see again Table 2). A better trade-off is achieved when $\Gamma\% = 50\%$, where the average total cost of RMPCA is around 4% less than for SRM (i.e., a 9 billion CHF reduction) while the chances of capacity shortage are estimated at only 4%. In comparison, the average total cost (including shortage penalties) from implementing the DET strategy is around 256 billion CHF, which is about 33% more than the total cost of RMPCA. Overall, this evidence supports the importance of: i) accounting of uncertainty of the demand response in ETEM; and ii) the importance of considering that the procurement is adjustable with respect to the actual DR.

Figure 5 is similar to Figure 4, but we have fixed the budget size to 50% and let the box size β vary. For example, $\beta = 0.2$ means that the actual demand response at each time-slice can deviate from the planned demand response by up to 20% of the total demand. As expected, increasing the support of ζ leads to an increase of the average total cost of the system. Interestingly, the rate of increase of the DET cost is larger than for RMPCA and SRM. Indeed, as SRM and RMPCA are sensitive to worst-case values of DR, more capacity is built under these strategies. This gives more production flexibility with cheaper prices for the system when the actual DR is lower than planned DR. Finally, the figure also illustrates that the average total cost (192.5 billion CHF) of preventing 96% of shortages when $\beta = 0.6$ is only about 5% more expensive than the cost estimated by the deterministic model (i.e. 182.5 billion CHF).

Figure 6 corresponds to the Cumulative Distribution Function (CDF) diagram of the simulated total cost of RMPCA, SRM and DET strategies when $\Gamma\% = 50\%$ and $\beta = 0.6$. We can note that RMPCA stochastically dominates SRM, while SRM stochastically dominates DET. In fact, the maximum cost of the RMPCA strategy over all random scenarios is less than the minimum cost of the SRM strategy, and similarly for the DET strategy.

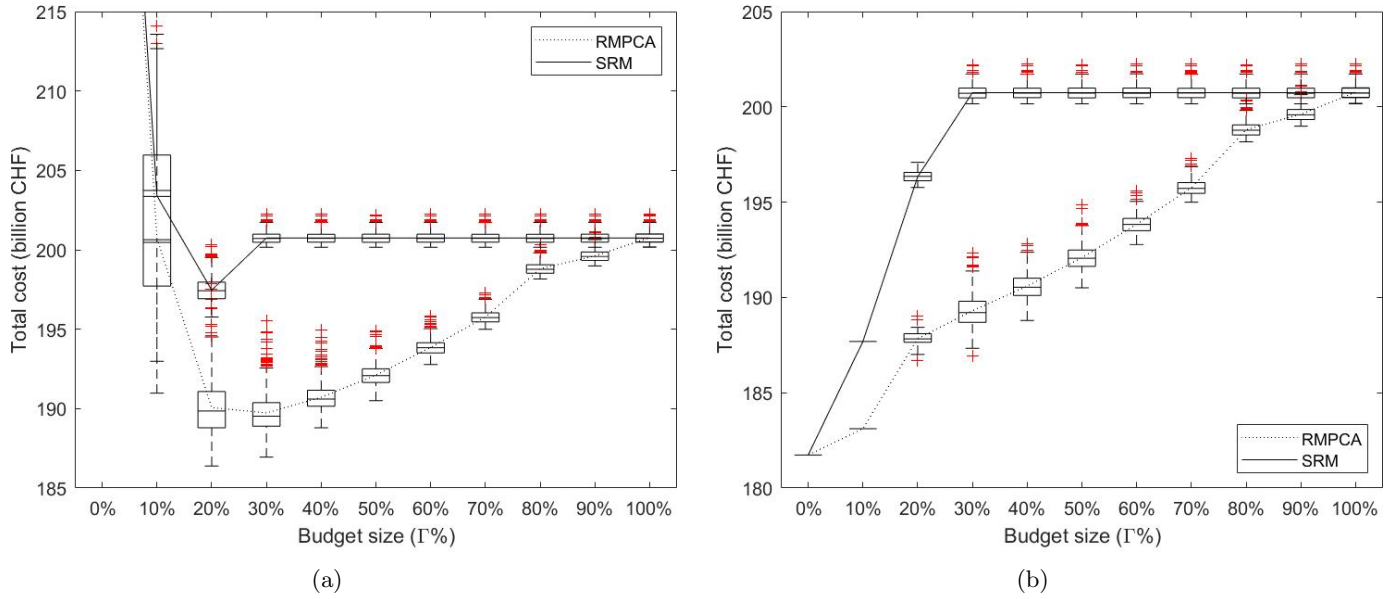


Figure 4: Statistics of the total cost of RMPCA and SRM strategies for different budget sizes and $\beta = 0.6$. (a) presents the distribution of the total cost over all 1000 random scenarios, whereas (b) presents the conditional distributions given that no shortages occurred. The curves present the mean values while each box presents the min, max, 25th and 75th percentile and some extreme realizations (+).

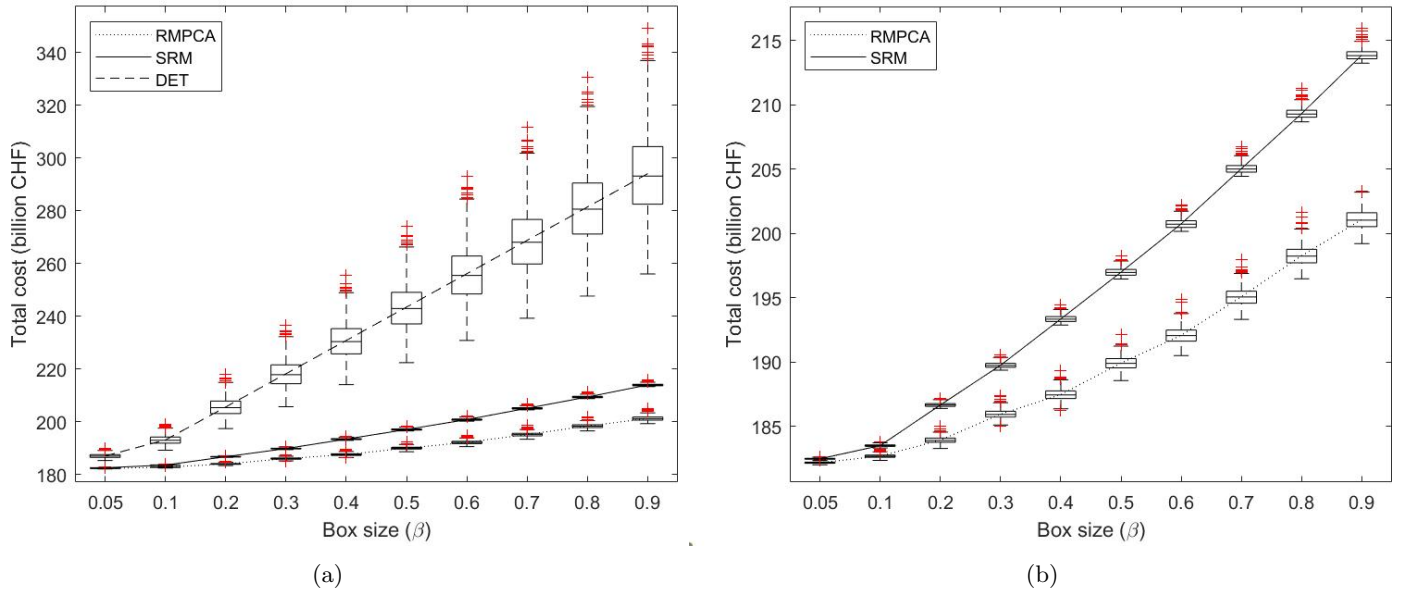


Figure 5: Statistics of the total cost of RMPCA, SRM and DET strategies for different box sizes and $\Gamma = 50\%$. (a) presents the distribution of the total cost over all 1000 random scenarios, whereas (b) presents conditional distributions given that no shortages occurred.

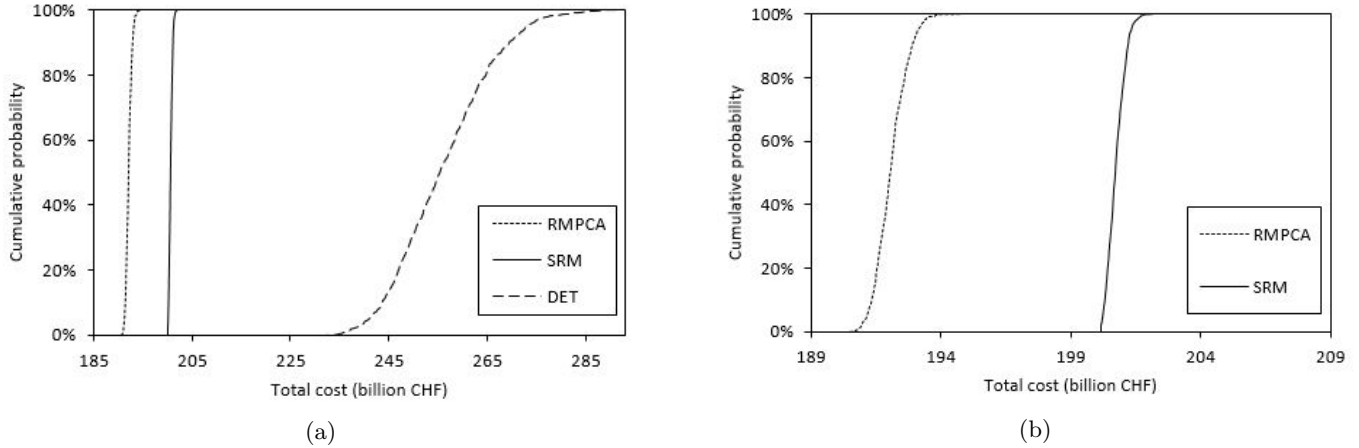


Figure 6: Comparison of the cumulative distribution functions of simulated total cost associated with RMPCA, SRM and DET policies when $\Gamma\% = 50\%$ and $\beta = 0.6$. (a) presents the distribution for all 1000 random scenarios, while (b) presents the distribution for the scenarios without capacity shortage.

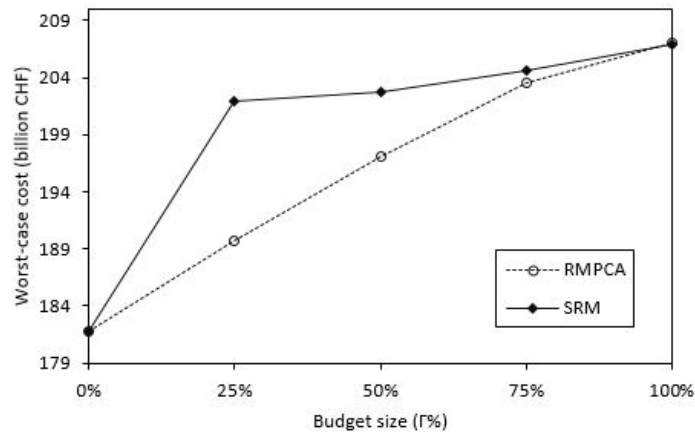


Figure 7: Worst-case performance of RMPCA and SRM as a function of budget size (for $\beta = 0.6$)

Again, this figure shows the importance of considering uncertainty of the DR in capacity expansion models.

Next, Figure 7 shows the worst-case total cost of the system when implementing the RMPCA and SRM strategies for different budget sizes. As this figure shows, RMPCA performs up to 6% better than SRM which is equivalent to approximately 12 billion CHF over the entire planning horizon.

Figure 8 provides finally an analysis on the trade-off between rate of capacity shortage and the average total cost for RMPCA and SRM strategies when the budget size varies from 10% to 100% and the box size $\beta = 0.6$. As it is also shown in Table 2, for budget sizes greater than 30%, the rate of capacity shortage for SRM is 0%. However, this zero-shortage comes at the cost of building too much capacity (see Figure 9, below). As Figure 8 suggests, allowing a certain degree of shortage can reduce the total cost drastically. For example, for $\Gamma\% = 60\%$ allowing a 2% chance of capacity shortage reduces the average total cost by around 7 billion CHF (equivalent to 4% of the total system cost).

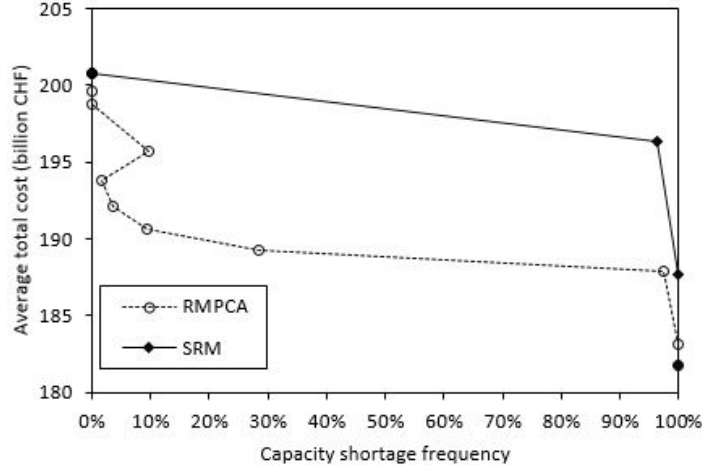


Figure 8: Trade-off between capacity shortage frequency and average total cost of RMPCA and SRM for different budget sizes and $\beta = 0.6$.

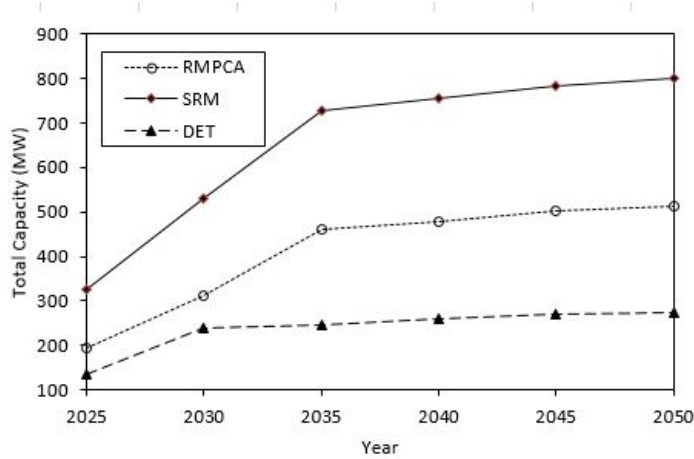


Figure 9: Total installed capacity for RMPCA, SRM and DET.

6.3 Electricity Generation Structure

In this section, we discuss how the different modeling schemes impact the evolution over time of electricity generation. Figure 9 presents first the total installed capacity for the different strategies. SRM is the most conservative approach, as it yields more capacity than RMPCA and DET (around 1.5 times more than RMPCA, and 3 times more than DET). It is indeed reasonable that RMPCA is less conservative than SRM, as it enjoys the flexibility of the energy procurement decisions to adjust for the actual (observed) demand response.

Figure 10 presents next the installed capacity by type of technology for the RMPCA, SRM and DET strategies. It highlights the importance of wind power to decarbonize the electricity generation sector. Specifically, wind power will form around 88% of the total installed capacity by 2050. It is followed by hydro power plant with 8%, and other technologies (including CHP, PV, geothermal, fuel-cell, and municipal-waste power plants) with less than 4% share in the total installed capacity. In addition, Figure 11 reports on the average annual electricity generation by different technology types for RMPCA, SRM, and DET policies. A more detailed description of the installed capacity and the average annual electricity generation is presented in Appendix B.

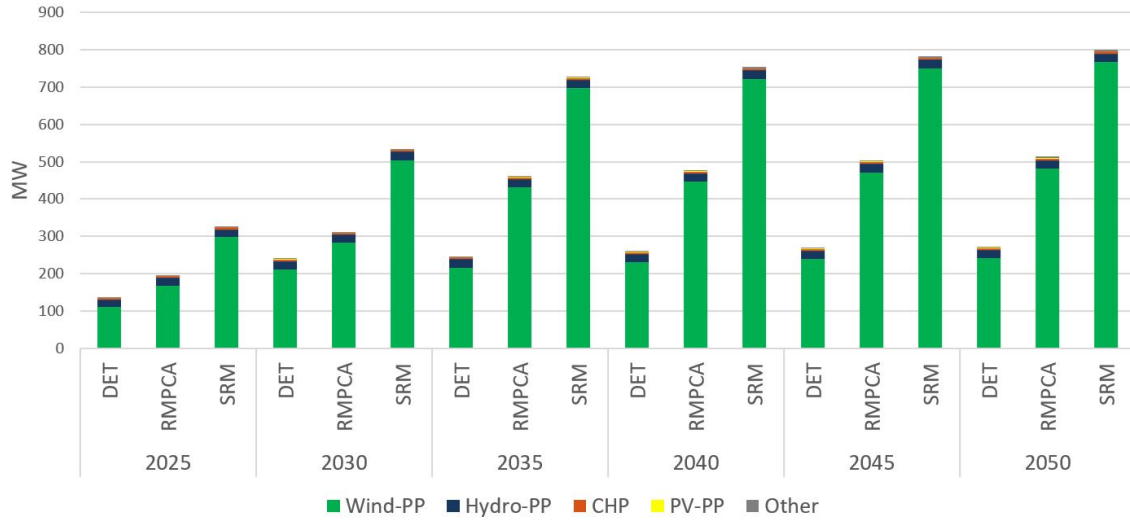


Figure 10: Total installed capacity (in MW) by type of technology for DET, RMPCA, and SRM models. Other technologies include geothermal, fuel-cell and municipal-waste power plant.

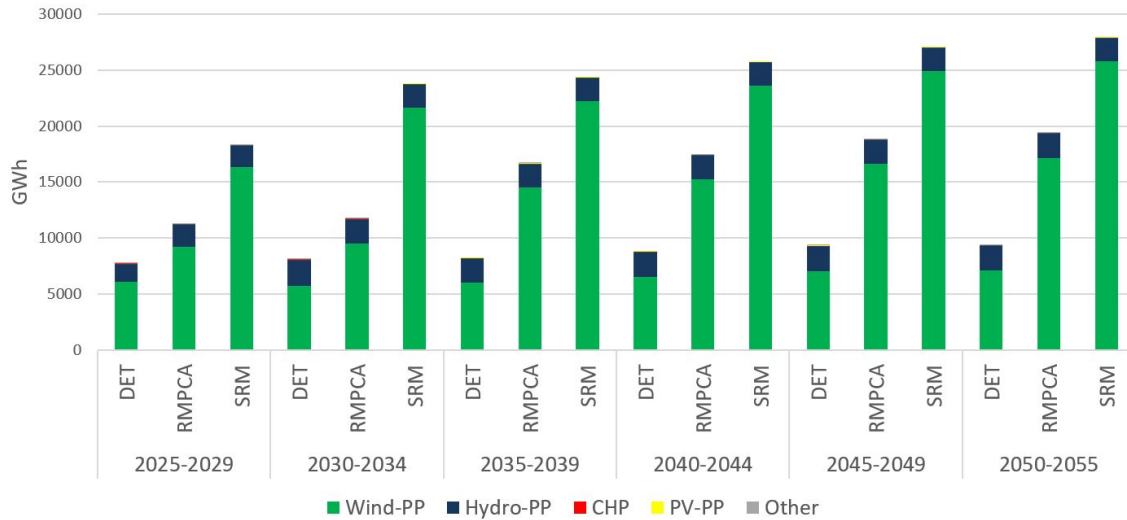


Figure 11: Average annual electricity generation (in GWh) by type of technology for DET, RMPCA, and SRM models over 1000 random scenarios. Other technologies include geothermal, fuel-cell and municipal-waste power plant.

Table 3: Solution time (ST) in seconds, number of iterations (It), percentage of improvement (% impr.) with respect to BD for different versions of the decomposition algorithms. Four problem sizes are considered, for which the number of variables (VAR) and constraints (CON) are reported together with the solution time of solving (5) with CPLEX.

T	DET (1) size		RMPCA (5) size		CPLEX	BD		PRO-BD			PRO-BD-VI		
	VAR	CON	VAR	CON	ST	ST	It	ST	It	% impr.	ST	It	% impr.
3	4,644	11,830	97,112	103,095	71.8	47.0	15	37.5	10	20%	26.0	4	45%
5	7,740	19,718	161,860	171,837	227.5	92.8	15	78.1	11	16%	68.9	5	26%
7	10,836	27,601	226,583	240,534	519.2	78.0	10	68.0	7	13%	48.0	3	39%
9	13,932	35,486	291,316	309,249	1242.8	97.0	9	82.1	6	15%	81.3	3	16%
11	17,028	43,380	356,094	378,045	2336.4	141.8	11	90.5	7	36%	58.6	3	59%
13	20,124	51,263	420,817	446,742	4350.4	154.4	10	105.6	7	32%	69.7	3	55%
Average					1458.0	101.8	12	77.0	8	22%	58.75	4	40%

6.4 Computational Analysis

In this section, we numerically evaluate the performance of the decomposition algorithms presented in Section 5. Specifically, we investigate how considering PRO and valid inequalities can improve the solution time and the number of iterations of Algorithm 1. In order for our results to be more general, we have designed smaller versions of *ETEM* with 16 technologies and 9 energy commodities, in which parameters are randomly assigned. Moreover, the size of the model changes with different problem horizons $T \in \{3, 5, 7, 9, 11, 13\}$. Then, for each problem size, we generate 20 random instances and solve these random instances for different budget sizes $\Gamma \in \{20\%, 40\%, 60\%, 80\%\}$. Table 3 reports, for different problem sizes, the average solution time (ST) of CPLEX for solving the RMPCA problem (5), together with the average solution time and the number of iterations (It) for different versions of our decomposition algorithm. Specifically, BD refers to the traditional Benders decomposition, i.e., Algorithm 1. PRO-BD refers to the algorithm with improved upper bound based on PRO solutions while PRO-BD-VI additionally exploits valid inequalities. It is worth mentioning that based on our numerical experiments, restricting Ω to the last 4 worst-case scenarios led to the best solution time. Results show that PRO-BD improves the solution time of the BD by up to 36%, whereas PRO-BD-VI is able to improve it by up to 59%. On average, PRO-BD improves the solution time by 22% and PRO-BD-VI improves by 40%. Finally, a considerable improvement in terms of the number of iterations is observed when comparing PRO-BD-VI and BD.

7 Concluding Remarks

In this paper, we model demand response uncertainty in an adjustable multi-period robust generation capacity expansion problem. In a first stage, the planner decides on capacity expansion, as well as, planned demand response. Then, in sequence of seasonal stages, the actual DR is revealed, and accordingly, the optimal procurement decisions are taken. As the resulting adjustable model is intractable, we use a conservative approximation scheme based on affine decision rules to solve the problem. We present a general formulation of the problem, where with slight modifications of the formulation one could accommodate other sources of uncertainty such as the level of production one can achieve from intermittent resources. We test our algorithm on a real-world case study based on the energy system of the Arc Lémanique region (Cantons of Geneva and Vaud, in Switzerland). Numerical results reinforce the importance of adjustable robust formulation. Namely, RMPCA can reduce the average total cost of the system by 33% compared to the solution of the deterministic problem, while keeping the chances of shortage near 4%. RMPCA also performs 4% better than a more naive static robust counterpart

(SRM). Finally, we develop a Benders type of decomposition to solve our large-scale RMPCA problem. In this algorithm, the speed of convergence is improved through i) adding valid inequalities to the master problem, and ii) identifying PRO solutions of the master problem. One alternative strategy to calculate the optimal investment decisions is to sequentially solve forward-shifting horizon RMPCA problems, and provide the investment plan over the entire horizon by assembling the capacity addition decisions made at the beginning of each forward-shifting horizon. However, this practical implementation of RMPCA problem is beyond the scope of this paper and shall be considered as future work. Moreover, it is worth mentioning that although this paper has focused on the uncertainty of demand response, one can easily introduce a group of supply-side uncertainties, such as availability of intermittent resources, or uncertainty of investment costs, to the general formulation of the problem (4). In this case, matrices A_i and vector f will be affected by these uncertainties. As our problem still preserves the “fixed recourse” condition, the approach presented in Sections 4.2 and 5 remains valid. Conversely, considering another group of supply-side uncertainties, such as fuel cost, or efficiency uncertainties, will affect matrices h_i and B_i , and thus consists of more significant changes to the structure of the decision model. In conclusion, here follows some interesting directions to expand on this paper. One idea is to consider investment decisions to be adaptive to the evolution of the uncertainty. Another extension is to consider other sources of uncertainties, especially the ones that do not preserve the fixed recourse characteristic of the problem. One could also control the loss of loads through reliability constraints. Finally, another possible extension would be to link the robust ETEM with a collective behavior demand model to more precisely adjust the size of the specific uncertainty sets of each period.

Acknowledgments. This research is supported by the Canadian IVADO programme (VORTEX Project). Olivier Bahn and Erick Delage also both acknowledge the financial support from the Natural Sciences and Engineering Research Council of Canada (respectively RGPIN-2016-04214 and RGPIN-2016-05208). The authors are thankful to Frédéric Babonneau and Alain Haurie for providing the ETEM-SG model used in Babonneau and Haurie (2019) and for their valuable discussions and comments to improve the paper.

References

- Adelman D, Uçkun C (2019) Dynamic electricity pricing to smart homes. Operations Research 67(6):1520–1542.
- Aliakbarisani S, Bahn O, Delage E, Foguen R (2021) Robust integration of electric vehicles charging load in smart grid’s capacity expansion planning. Les Cahiers du GERAD G-2021-44.
- Amjady N, Attarha A, Dehghan S, Conejo AJ (2018) Adaptive robust expansion planning for a distribution network with DERs. IEEE Transactions on Power Systems 33(2):1698–1715.
- Amrollahi MH, Bathaee SMT (2017) Techno-economic optimization of hybrid photovoltaic/wind generation together with energy storage system in a stand-alone micro-grid subjected to demand response. Applied Energy 202:66–77.
- Arasteh H, Sepasian M, Vahidinasab V (2015) Toward a smart distribution system expansion planning by considering demand response resources. Journal of Operation and Automation in Power Engineering 3(2):116–130.
- Ardestani-Jaafari A, Delage E (2018) The value of flexibility in robust location–transportation problems. Transportation Science 52(1):189–209.
- Ardestani-Jaafari A, Delage E (2020) Linearized robust counterparts of two-stage robust optimization problems with applications in operations management. INFORMS Journal on Computing. Forthcoming.
- Asensio M, Meneses de Quevedo P, Muñoz-Delgado G, Contreras J (2018) Joint distribution network and re-

- newable energy expansion planning considering demand response and energy storage—part i: Stochastic programming model. IEEE Transactions on Smart Grid 9(2):655–666.
- Babonneau F, Caramanis M, Haurie A (2016) A linear programming model for power distribution with demand response and variable renewable energy. Applied Energy 181:83 – 95.
- Babonneau F, Caramanis M, Haurie A (2017) ETEM-SG: Optimizing regional smart energy system with power distribution constraints and options. Environmental Modeling & Assessment 22(5):411–430.
- Babonneau F, Foguen R, Haurie A, Malhame R (2020) Coupling a power dispatch model with a wardrop or mean-field-game equilibrium model. Dynamic Games and Applications .
- Babonneau F, Haurie A (2019) Energy technology environment model with smart grid and robust nodal electricity prices. Annals of Operations Research 274.
- Babonneau F, Haurie A, Tarel G, Thénicié J (2012) Assessing the future of renewable and smart grid technologies in regional energy systems. Swiss J Economics Statistics 148:229–273.
- Baringo L, Baringo A (2018) A stochastic adaptive robust optimization approach for the generation and transmission expansion planning. IEEE Transactions on Power Systems 33(1):792–802.
- Ben-Tal A, den Hertog D, Vial J (2015) Deriving robust counterparts of nonlinear uncertain inequalities. Math. Program 149:265–299.
- Ben-Tal A, Goryashko A, Guslitzer E, Nemirovski A (2004) Adjustable robust solutions of uncertain linear programs. Mathematical Programming 99(2):351–376.
- Bertsimas D, Brown DB, Caramanis C (2011) Theory and applications of robust optimization. SIAM Review 53(3):464–501.
- Bertsimas D, Georghiou A (2018) Binary decision rules for multistage adaptive mixed-integer optimization. Math. Program. 167:395–433.
- Bertsimas D, Sim M (2004) The price of robustness. Operations Research 52(1):35–53.
- Bertsimas D, Thiele A (2006) A robust optimization approach to inventory theory. Operations Research 54(1):150–168.
- Birge JR, Louveaux F (2011) Introduction to Stochastic Programming, chapter Solution methods, 102–180 (New York, NY, USA: Springer-Verlag), second edition.
- Bloom JA (1983) Solving an electricity generating capacity expansion planning problem by generalized benders’ decomposition. Operations Research 31(1):84–100.
- Bloom JA, Caramanis M, Charny L (1984) Long-range generation planning using generalized benders’ decomposition: Implementation and experience. Operations Research 32(2):290–313.
- Chatterjee N, Glick R, McNamee B (2018) Assessment of demand response and advanced metering. Technical report, Federal Energy Regulatory Commission.
- Dehghan S, Amjady N, Kazemi A (2014) Two-stage robust generation expansion planning: A mixed integer linear programming model. IEEE Transactions on Power Systems 29(2):584–597.
- Domínguez R, Conejo AJ, Carrión M (2016) Investing in generation capacity: A multi-stage linear-decision-rule approach. IEEE Transactions on Power Systems 31(6):4784–4794.
- Dyson ME, Borgeson SD, Tabone MD, Callaway DS (2014) Using smart meter data to estimate demand response potential, with application to solar energy integration. Energy Policy 73:607–619.
- Georghiou A, Tsoukalas A, Wiesemann W (2020) A primal–dual lifting scheme for two-stage robust optimization. Operations Research 68(2):572–590.

- Gärttner J, Flath CM, Weinhardt C (2018) Portfolio and contract design for demand response resources. European Journal of Operational Research 266(1):340–353.
- Han D, Wu W, Sun W, Yan Z (2018) A two-stage robust stochastic programming approach for generation expansion planning of smart grids under uncertainties. 2018 IEEE Power Energy Society General Meeting (PESGM), 1–5.
- He S, Gao H, Liu J, Liu J (2019) Distributionally robust optimal DG allocation model considering flexible adjustment of demand response. 2019 IEEE Power Energy Society General Meeting (PESGM), 1–5.
- Huang J, Pan K, Guan Y (2020) Multistage stochastic power generation scheduling co-optimizing energy and ancillary services. INFORMS Journal on Computing. Advance online publication. .
- Iancu DA, Trichakis N (2014) Pareto efficiency in robust optimization. Management Science 60(1):130–147.
- Knueven B, Ostrowski J, Watson JP (2020) On mixed-integer programming formulations for the unit commitment problem. INFORMS Journal on Computing. Advance online publication. .
- Koltsaklis NE, Dagoumas AS (2018) State-of-the-art generation expansion planning: A review. Applied Energy 230:563–589.
- Li C, Dong Z, Chen G, Luo F, Liu J (2015) Flexible transmission expansion planning associated with large-scale wind farms integration considering demand response. IET Generation, Transmission Distribution 9(15):2276–2283.
- Liu C, Deng K, Wen G, Yu X (2019) Optimal scheduling of electric vehicle charging with energy storage facility in smart grid. IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society, volume 1, 6549–6554.
- Lohmann T, Rebennack S (2017) Tailored benders decomposition for a long-term power expansion model with short-term demand response. Management Science 63(6):2027–2048.
- Magnanti TL, Wong RT (1981) Accelerating benders decomposition: Algorithmic enhancement and model selection criteria. Operations Research 29(3):464–484.
- Mejía-Giraldo D, McCalley J (2014) Adjustable decisions for reducing the price of robustness of capacity expansion planning. IEEE Transactions on Power Systems 29(4):1573–1582.
- Qi N, Cheng L, Xu H, Wu K, Li X, Wang Y, Liu R (2020) Smart meter data-driven evaluation of operational demand response potential of residential air conditioning loads. Applied Energy 279:115708.
- Rahmaniani R, Crainic TG, Gendreau M, Rei W (2017) The benders decomposition algorithm: A literature review. European Journal of Operational Research 259(3):801–817.
- Roveto M, Mieth R, Dvorkin Y (2020) Co-optimization of VaR and CVaR for data-driven stochastic demand response auction. IEEE Control Systems Letters 4(4):940–945.
- Shariatzadeh F, Mandal P, Srivastava AK (2015) Demand response for sustainable energy systems: A review, application and implementation strategy. Renewable and Sustainable Energy Reviews 45:343–350.
- Sion M (1958) On general minimax theorems. Pacific Journal of Mathematics 8(1):171–176.
- Ströhle P, Flath CM (2016) Local matching of flexible load in smart grids. European Journal of Operational Research 253(3):811–824.
- US Department of Energy (2006) Benefits of demand response in electricity markets and recommendations for achieving them. Technical report, US Department of Energy, Washington DC.
- Vatani B, Chowdhury B, Lin J (2017) The role of demand response as an alternative transmission expansion solution in a capacity market. 2017 IEEE Industry Applications Society Annual Meeting, 1–7.

- Venizelou V, Philippou N, Hadjipanayi M, Makrides G, Efthymiou V, Georghiou GE (2018) Development of a novel time-of-use tariff algorithm for residential prosumer price-based demand side management. Energy 142:633–646.
- Yanikoğlu I, Gorissen B, den Hertog D (2019) A survey of adjustable robust optimization. European Journal of Operational Research 277(3):799–813.
- Zhang Q, Zhang S (2019) Robust decision-making model for electricity retailers considering incentive-based demand response. 2019 IEEE Sustainable Power and Energy Conference (iSPEC), 1157–1162.
- Zhao C, Wang J, Watson J, Guan Y (2013) Multi-stage robust unit commitment considering wind and demand response uncertainties. IEEE Transactions on Power Systems 28(3):2708–2717.
- Zou J, Ahmed S, Sun XA (2018) Partially adaptive stochastic optimization for electric power generation expansion planning. INFORMS Journal on Computing 30(2):388–401.

Appendices

A Comparison with robust ETEM from Babonneau et al. (2020)

In this section, we first summarize the technique proposed in Babonneau et al. (2020) to model the demand response uncertainty in ETEM. Then, we discuss issues that arise when using their proposed approach to account for demand response uncertainty that is present in our case study. Namely, we will show that Babonneau et al. (2020)’s approach ends up simply adding redundant constraints to the deterministic problem (1). We also encourage interested readers to have a look at Appendix B in Aliakbarisani et al. (2021), which additionally argues that the constraints proposed in Babonneau et al. (2020) to model demand response uncertainty fail to offer protection against observed demand response deviations.

A.1 A summary of the method used in Babonneau et al. (2020)

In Babonneau et al. (2020), the authors propose an iterative procedure to robustify the deterministic version of ETEM presented in (1) in context of uncertainty. They first recommend solving the deterministic version of ETEM to obtain an initial demand response plan $\Theta_{t,l,c} \bar{V}_{t,s,c}$. This plan can then be used to produce distributions of realized demand responses $\tilde{V}_{t,l,s,c}$ for each period, time slice, region, and commodity. Descriptive statistics of each $\tilde{V}_{t,l,s,c}$ are then extracted in the form of a mean value $\hat{\Delta}_{t,s}^{l,c}$ and a confidence region $[\Delta_{t,s}^{l,c,-}, \Delta_{t,s}^{l,c,+}]$. Finally, they use arguments based on moment inequalities to suggest “robustifying” the ETEM by adding the following constraints:

$$\Delta_{l,c}^{\min} \leq \sum_{t \in \mathcal{T}} \sum_{s \in \mathcal{S}} \Theta_{t,l,c} V_{t,s,c} \leq \Delta_{l,c}^{\max}, \quad \forall l \in \mathcal{L}, c \in \mathcal{C}^U, \quad (18)$$

with

$$\begin{aligned} \Delta_{l,c}^{\min} &:= \sup_{\xi \in [0,1]^{\mathcal{T} \times \mathcal{S}}} \sum_{(t,s) \in \mathcal{T} \times \mathcal{S}} \Delta_{t,s}^{l,c,-} + (\hat{\Delta}_{t,s}^{l,c} - \Delta_{t,s}^{l,c,-}) \xi_{t,s} \\ &\text{s.t.} \quad \sum_{(t,s) \in \mathcal{T} \times \mathcal{S}} \xi_{t,s} = \Gamma, \end{aligned}$$

and

$$\begin{aligned} \Delta_{l,c}^{\max} &:= \inf_{\xi \in [0,1]^{\mathcal{T} \times \mathcal{S}}} \sum_{(t,s) \in \mathcal{T} \times \mathcal{S}} \Delta_{t,s}^{l,c,+} + (\hat{\Delta}_{t,s}^{l,c} - \Delta_{t,s}^{l,c,+}) \xi_{t,s} \\ &\text{s.t.} \quad \sum_{(t,s) \in \mathcal{T} \times \mathcal{S}} \xi_{t,s} = \Gamma. \end{aligned}$$

This iterative process is repeated until the robustified model starts producing the same solutions in two consecutive iterations.

A.2 Implementing Babonneau et al. (2020)’s approach in our Swiss Case study

In the Swiss case study presented in Section 6, the random demand response is modeled as the summation of the planned demand and the relative demand response deviations (RDRDs), i.e., $\tilde{V}_{t,l,s,c} := \Theta_{t,l,c} (\bar{V}_{t,s,c} + \delta_{t,s,c})$, where each $\delta_{t,s,c}$ follows a uniform distribution over the interval $[-\beta_{t,s,c}, \beta_{t,s,c}]$. Therefore, in the first iteration

of Babonneau et al. (2020)'s, \bar{V} is the solution of the deterministic ETEM (i.e. problem (1)) and the following descriptive statistics are obtained from the observed realized demand responses:

$$\Delta_{t,s}^{l,c,-} := \Theta_{t,l,c} [\bar{V}_{t,s,c} - (1 - \varepsilon)\beta_{t,s,c}], \quad (19)$$

$$\Delta_{t,s}^{l,c,+} := \Theta_{t,l,c} [\bar{V}_{t,s,c} + (1 - \varepsilon)\beta_{t,s,c}], \quad (20)$$

$$\hat{\Delta}_{t,s}^{l,c} := \Theta_{t,l,c} \bar{V}_{t,s,c} \quad (21)$$

where $1 - \varepsilon$ is the level of confidence for the confidence interval $[\Delta_{t,s}^{l,c,-}, \Delta_{t,s}^{l,c,+}]$. Lemma 3 demonstrates that \bar{V} remains optimal after adding constraint (18) to problem (1).

Lemma 3. *The solution \bar{V} satisfies constraint (18).*

Proof. We will first show that the lower bounding part of constraint (18) is satisfied. Namely,

$$\begin{aligned} \Delta_{l,c}^{\min} &= \sup_{\xi \geq \hat{\xi}} \sum_{(t,s) \in \mathcal{I}} \Delta_{t,s}^{l,c,-} + (\hat{\Delta}_{t,s}^{l,c} - \Delta_{t,s}^{l,c,-}) \xi_{t,s} \\ &= \sup_{\xi \geq \hat{\xi}} \sum_{(t,s) \in \mathcal{I}} \Theta_{t,l,c} \bar{V}_{t,s,c} - \Theta_{t,l,c} (1 - \varepsilon) (1 - \xi_{t,s}) \beta_{t,s,c} \\ &= \sum_{(t,s) \in \mathcal{I}} \Theta_{t,l,c} \bar{V}_{t,s,c} - \inf_{\xi \geq \hat{\xi}} \sum_{t,s} \Theta_{t,l,c} (1 - \varepsilon) (1 - \xi_{t,s}) \beta_{t,s,c} \\ &= \sum_{(t,s) \in \mathcal{I}} \Theta_{t,l,c} \bar{V}_{t,s,c}, \end{aligned}$$

where $\hat{\xi} := \frac{\Delta_{l,c}^{\min}}{\sum_{(t,s) \in \mathcal{I}} \xi_{t,s}} \in [0, 1]$, and which would confirm that \bar{V} satisfies the lower bound. In details, the first two equalities follow from the definition of $\Delta_{l,c}^{\min}$ and $\Delta_{t,s}^{l,c,-}$ (see equation (19)), while the first inequality follows from the fact that all terms in $\Theta_{t,l,c} (1 - \varepsilon) (1 - \xi_{t,s}) \beta_{t,s,c}$ are non-negative. A similar analysis can be done to prove that \bar{V} also satisfies the upper bound in constraint (18). \square

Interestingly, Lemma 3 implies that the procedure proposed by Babonneau et al. (2020) will converge in the second iteration and prescribe the solution of the deterministic version of ETEM in this case study. As reported in Aliakbarisani et al. (2021), it therefore once again seems that Babonneau et al. (2020)'s approach is unable to immunize ETEM from demand response uncertainty.

B Details on electricity generation structure

In this section we present a table format of Figures 10 and 11. Table 4 reports on the total installed capacity by type of power plant for each of DET, RMPCA, and SRM models. While, DET is proposing to build around 270 MW of power plant, RMPCA and SRM are proposing to build 513 and 800 MW of power plants. According to the capacity expansion strategy of all models, the wind power plant will prevail. The second technology is hydro power plant with around 23 MW by 2050. Finally, the capacity of other technologies, including photovoltaic, CHP, geothermal, fuel-cell, and municipal-waste power plants will be around only 9 MW. Table 5 reports on the average annual electricity production by each technology for RMPCA, SRM, and DET models. According to RMPCA strategy, Wind power plants will produce up to 88% of the total electricity generated in the region in 2050. Hydro power will produce 11%, and other technologies will produce only 0.2% of the total electricity produced in this region in 2050.

Table 4: Total installed capacity (in MW) by type of technology for RMPCA, SRM and DET models.

	2025			2030			2035			2040			2045			2050		
	RMPCA	SRM	DET	RMPCA	SRM	DET	RMPCA	SRM	DET	RMPCA	SRM	DET	RMPCA	SRM	DET	RMPCA	SRM	DET
Hydro-PP	21.3	21.3	18.9	22.4	22.4	22.4	22.4	22.4	22.4	22.6	22.6	22.6	22.8	22.8	22.8	23	23	23
PV-PP	0.2	0.2	0.2	0.2	0.2	0.2	0.8	0.8	0.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8
Wind-PP	168.4	297.6	110.8	282.5	504.1	211.3	431.1	697.3	216.5	446.2	721.9	230.0	471.2	750.1	239.1	480.9	767.2	240.7
CHP	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7
Other ¹	0.5	0.5	0.5	1.2	0.5	0.5	1.2	3.1	0.5	1.2	3.1	0.5	1.2	3.1	0.5	3.1	3.1	3.1

¹ Other technologies include geothermal, fuel-cell and municipal-waste power plants.

Table 5: Average annual electricity generation (in GWh) by type of technology for RMPCA, SRM and DET models over 1000 random scenarios.

	2025-2029			2030-2034			2035-2039			2040-2044			2045-2049			2050-2055		
	RMPCA	SRM	DET	RMPCA	SRM	DET	RMPCA	SRM	DET	RMPCA	SRM	DET	RMPCA	SRM	DET	RMPCA	SRM	DET
Hydro-PP	2010	1943.6	1642.8	2210.6	2086.8	2306	2153.4	2088	2146.1	2173.3	2108.7	2202.4	2188.5	2129.7	2244.5	2214.1	2152	2210.8
PV-PP	4.9	4.9	4.9	3.7	3.7	3.7	16.7	16.7	16.7	38.3	38.3	38.3	38.3	38.3	38.3	38.3	38.3	38.3
Wind-PP	9238	16323	6076	9506	21654	5744	14495	22233	6033	15233	23583	6537	16602	24892	7038	17137	25749	7124
CHP	12.5	0.5	40.3	10.2	0.0	39.4	1.6	0	15.8	1.6	0	15.8	1.5	0	16.6	1.5	0	16.6
Other ¹	16.2	0.7	32.5	13.8	0.0	29	2.4	0	0	2.3	0	0	2.2	0	6.8	2.3	0	8.9

¹ Other technologies include geothermal, fuel-cell and municipal-waste power plants.