Modeling uncertainty processes for multi-stage optimization of strategic energy planning: An auto-regressive and Markov chain formulation

Esnil Guevara[1], Frédéric Babonneau[2], and Tito Homem-de-Mello[3]

[1]PhD Program in Industrial Engineering and Operations Research, Universidad Adolfo Ibáñez, Chile
[2]KEDGE Business School, France and ORDECSYS, Switzerland
[3]School of Business, Universidad Adolfo Ibáñez, Chile

March 28, 2022

Abstract

This paper deals with the modeling of stochastic processes in long-term multistage energy planning problems when little information is available on the degree of uncertainty of such processes. Starting from simple estimates of variation intervals for uncertain parameters, such as energy demands and costs, we model the temporal correlation of these parameters through autoregressive (AR) models. We introduce a coefficient for zigzag effects in the evolution of uncertain processes that controls the likelihood of extreme scenarios. To preserve the convexity of the stochastic problem, we discretize the AR models associated with the cost parameters involved in the objective function by Markov chains. The resulting formulation is then solved with an advanced SDDP algorithm available in the literature that handles finite-state Markov chains. Our numerical experiments, performed on the Swiss energy system, show a very desirable adaptation strategy of investment decisions to uncertainty scenarios, a behavior that is not observed when the temporal correlation is ignored. Moreover, the solutions lead to better out-of-sample cost performances, especially on extreme scenario realizations, than the non-correlated ones which usually yield overcapacities to protect against high, but unlikely, parameter variations over time.

Keywords: Strategic energy planning, Stochastic Dual Dynamic Programming, autoregressive process, Markov chain.

1 Introduction

Energy plays a fundamental role in our lives as it is an essential contribution not only for daily tasks but also for carrying out various productive activities. Due to the constant transformation that the energy sector is undergoing, it is essential to plan long-term strategies that will guarantee our energy supply in the future. The work of [Rojas-Zerpa and Yusta, 2014, p. 67] indicates that "*energy planning implies finding a set of sources and conversion equipment that optimally satisfy the energy demand of all activities*". In other words, a strategic energy plan establishes the resources and technologies that will be needed to meet our future energy needs. The planning horizon is generally long enough, i.e., 20 to 50 years, to identify profitable and robust developments according to technical, economic, social and environmental criteria.

Optimization models are generally used to find an optimal strategy to minimize the investment costs of new technologies and the operating and maintenance costs of the entire system while satisfying a set of underlying conditions. These models can be divided into two broad categories. In the first one, the system is optimized for a target year without looking at the transition. Models in this category, such as [Hilpert et al., 2018], Calliope [Pfenninger and Pickering, 2018] and EnergyScope [Limpens et al., 2019], therefore have only one investment and operation period. In the second category, we find multi-period planning models that optimize the transition of the energy system over a given horizon such as MARKAL/TIMES [Krzemień, 2013], OSeMOSYS [Howells et al., 2011], ETEM [Babonneau et al., 2017], MESSAGE [Sullivan et al., 2013] and SMART [Powell et al., 2012]. In general, energy planning models are large-scale optimization models and integrate all energy sectors that are mutually dependent on each other, such as electricity, heating or transportation. In addition, these models combine different time scales in order to model both short-term operations and long-term investment decisions. Energy planning solutions are often based on deterministic demand forecasts, fuel prices, technological efficiencies, resource availability, etc., assuming that all input parameters are known over the entire planning horizon. However, if some of the future forecasts deviate from the estimated values, the planned strategic investment decisions are likely to be sub-optimal. In addition, when formulating a deterministic problem, rare events and special cases are often excluded from optimization.

Given these limitations, stochastic formulations have been proposed to obtain robust solutions that protect the system against future uncertainties. For example, in [Powell et al., 2012], resource costs, renewable generation and demand growth are considered uncertain in the SMART model. In [Babonneau et al., 2017], the authors deal with uncertainty on resource costs, renewable generation and demand growth in the ETEM model. See [Moret et al., 2020] for a review of applications. There are currently many ways to deal with uncertainty in optimization problems, Stochastic Programming (SP) and Robust Optimization (RO) being the most widely used in energy models. Nevertheless, it is well-known that these techniques come with numerical and modeling complexities and, as a consequence, most of the studies in the literature are limited either to two-stage problems and/or simplistic uncertainty models. This paper aims to fill this gap by proposing a realistic modeling of uncertain parameters in a multi-stage optimization problem. It is a continuation of two recent papers [Moret et al., 2020, Guevara et al., 2020] that we discuss next.

A two-stage mixed-integer linear optimization model, EnergyScope, for the Swiss real energy system has been formulated under these two types of approaches. The deterministic model was first introduced by Moret et al. [2016] and formulated in a complete Robust Optimization framework in Moret et al. [2020]. The model considers a single investment period and a monthly time scale for operations, which allows for capturing the seasonal storage of hydroelectric dams. As described in [Moret et al., 2017], this model has a large number of uncertain parameters, which were grouped into 21 categories (see Moret et al. [2017], Table 1), such as resources costs, investment costs of technologies, maintenance costs of technologies, lifetimes of technologies, technology efficiencies and end-use energy demands to mention a few. All these uncertain parameters are thus taken into consideration in [Moret et al., 2020] under a unified robust but static formulation, i.e, no adaptation of the second-stage operation decisions to the revealed uncertainty is possible. More recently, the authors in [Guevara et al., 2020] have formulated EnergyScope as a Two-Stage Distributionally Robust Optimization (DRO) model to generate robust strategic investment decisions that are insensitive to the probability distribution functions of the uncertain parameters, which are considered as unknown. DRO is based on the design of a set of distributions –called an ambiguity set– and it aims at providing the model with protection against the worst distribution within that set; see, for instance, [Wiesemann et al., 2014]. However, in both models mentioned above the dynamics of data process and sequential investment decisions are not considered.

Multistage Stochastic Linear Programming (MSLP) is a modeling framework that allows decisions to be made sequentially under uncertainty (see, for instance, Birge and Louveaux [2011]). In its classic version, the uncertainty is represented by a stochastic process, with a known probability distribution, where the objective is to seek a sequence of decisions (or policy) for each stage that optimizes an objective function of an expected value form. More specifically, a policy is a sequence of functions—one for each

stage—that map the observed outcomes up to the current stage into a decision that satisfies the constraints of that stage. Usually, the random data of the processes are assumed to have continuous distribution, so when calculating the implied expected value (multivariate integrals), a standard approach is to discretize the random process by generating a finite number of possible realizations, called *scenarios*. These scenarios can be represented by the corresponding scenario tree that branches at each stage. Therefore, in order to use a dynamic model that provides a more realistic representation of the problem including multiple investment periods, it is necessary to define stochastic processes that represent adequately the evolution of the uncertain parameters of the model.

1.1 Modeling uncertainty

Modeling uncertain processes in energy models correctly is a complex task. In forecasting models, it is often necessary to have adequate knowledge of the past behavior of the data in order to be able to forecast what will happen in the future. Time series forecasting methods are well-designed to take into account a spatial-temporal relationship. There are many ways to model time series, such as the Autoregressive model (AR), the Autoregressive Moving Average model (ARIMA) and Moving-Average model (MA), in which a linear function of past observations is used to predict future values [Box et al., 2015]. Some works in energy planning incorporate more complex modeling techniques to have a better representation of uncertainty. For example, in Jin et al. [2011] it is assumed that annual electricity demand and natural gas prices can be represented as a GBM (Geometric Brownian Motion) since both are modeled with an annual growth rate relative to previous years, and the annual geometric growth rate is uncorrelated in different years. Such an approach however requires large volumes of data that are often not available, and a series of transformations and hypothesis tests to validate that it corresponds to the proposed model [Marathe and Ryan, 2005].

Another way to model uncertainty is through scenario trees. In a scenario tree, each node represents the state of the system in a given time period. The arcs connecting the nodes between successive stages represent the possible transitions from one state at time period t to another state at time period $t+1$, where each of the states has an associated probability of occurrence of the given uncertain parameter. A path from the root node to the leaf node represents a scenario. There are methods that construct scenario trees given a set of scenarios by approximating an appropriate metric over the probability distribution space [Heitsch and Roemisch, 2009]. Stochastic processes are often discretized to generate scenario trees, which are then used to solve multistage stochastic problems.

Regardless of the many possible ways of constructing a scenario tree, the problem of dimensionality and exponential growth of the problem size is unavoidable especially when the input data involves multiple stochastic processes, thereby making the optimization of such models very difficult even when the number of stages is small. This issue, called "Curse of dimensionality", is mainly due to the fact that each node must have a unique predecessor, i.e., each node uniquely mapped into the entire history of the process. Although this framework is quite flexible in that it allows for modeling arbitrary processes, in some cases it may not be necessary to consider the entire history of a scenario. For instance, a finite-state Markov chain may be used in such cases, allowing each node to have multiple predecessors in each stage. In a Markov chain it is assumed that the future state evolves probabilistically as a function of the current state. When a real-valued stochastic process satisfies this property, a finite state Markov chain can be obtained by discretization techniques [Tauchen, 1986].

1.2 Solutions methods for multi-stage stochastic programming

We now discuss the impact of uncertainty modeling on solution methods for MSLP problems and on their efficiency. As discussed earlier, multistage stochastic programs usually approximate the stochastic process $\{\xi_t\}_{t=1}^{T}$ with a finite number of scenarios exhibiting a tree structure. This usually leads to large-scale Linear programming (LP) or Mixed-integer linear programming (MILP) reformulations that can be solved by decomposition methods. In particular, energy planning formulations can yield significant

computational challenges due to the number of scenarios, the number of stages and/or the presence of integer decision variables. Singh et al. [2009] present a multistage stochastic MILP model for planning capacity expansion of an electricity distribution network in New Zealand. Only the future demand is considered uncertain and is modeled by a scenario tree. The deterministic-equivalent MILP problem is solved using Dantzig-Wolfe decomposition. Another multistage stochastic MILP model is formulated in [Park and Baldick, 2016] to solve the generation expansion problem, where the uncertain parameters of load and wind availability are considered as random variables. A scenario tree is generated (using the Gaussian copula method), and then reduced for improved computation performance when using a rolling-horizon method. Ioannou et al. [2019] formulate the power generation expansion planning as a multistage stochastic LP model in which ten power generation technologies are considered as alternatives for new power plants. The authors propose a hybrid uncertainty modeling approach, in which the setting of scenarios tree (e.g. uncertainty of energy demand and capital cost of renewable energy technologies) is combined with a Monte Carlo simulation method. However, fuel prices are modeled as time-independent random variables and the number of time periods is limited. Liu et al. [2018] propose a multistage stochastic investment planning for a centralized power system model with five generic generation technologies, where the long-term uncertainty (e.g. investment, fuel-cost and demand-growth rate) is explicitly modeled through a scenario tree and the short-term uncertainty (e.g. hourly demand and renewable generation) is modeled by means of time-series approaches. To solve the resulting problem, the authors decompose the model by scenarios through a progressive hedging algorithm to reduce computational time.

As the number of the uncertain parameters realizations and the number of periods in the planning horizon make exponential growth of the tree inevitable [Shapiro and Nemirovski, 2005], usually leading to numerical difficulties, the papers above are limited either to few uncertainties, few periods or do not consider the whole energy system. In other words, the tree approach is applicable when the number of stages is small, or when combined with scenario reduction techniques (such as backward reduction or forward selection). To deal with the rapidly growing scenario tree issue, Stochastic Dual Dynamic Programming (SDDP) was introduced first in the context of the hydrothermal scheduling problem by Pereira and Pinto [1991]. This algorithm assumes that the MSLP has relatively complete recourse (i.e., for any feasible solutions at stages $t = 0, ..., T-1$, there exists a feasible solution to any realized stage $t+1$ subproblem with probability one), and that the stochastic process is *stage-wise independent*, i.e., the information revealed at stage t of the random variable ξ_t does not depend on the past information of the variables $\xi_{t-1}, \xi_{t-2}, ..., \xi_2, \xi_1$ of the previous stages. Although these assumptions are helpful in easing the computational burden for the solution method, thereby making SDDP one of the few algorithms available to solve large-scale problems, they can be considered as very unrealistic. To circumvent such limitations, the SDDP algorithm has been extended to integrate correlation or non-continuous variables and applied in several publications [Lara et al., 2018, 2019, Rebennack, 2014, Zou et al., 2019, Thomé et al., 2019]. SDDP relies on the nested decomposition algorithm of Benders [Birge, 1985] and thus avoids solving all the possible scenarios of the multi-stage scenarios tree.

In the literature, it has been assumed that the uncertainties in generation expansion planning addressed by the SDDP are stagewise independent. Rebennack [2014] solves a generation planning problem for hydro-thermal power systems via Benders decomposition, where the master problem decides investment decisions "here and now" without knowledge of uncertainty, while operational decisions are made under a "wait and see" approach. The underlying uncertainty of hydro inflows is considered in the operational decisions and is solved using a standard SDDP algorithm. Similar problems where the SDDP is combined with a master problem to solve generation and transmission expansion planning are presented in Campodónico et al. [2003] and Thomé et al. [2019]. An extension of SDDP to multistage integer programming models (SDDiP proposed by Zou et al. [2019]) that addresses the long-term planning of electric power infrastructure under multiscale uncertainty is presented by Lara et al. [2019]. In that work it is assumed that the cost of fuels, the availability of renewable generation and the load demand are stage-wise independent in the scenario tree, and the decisions of investments in generation and storage units are integer variables. However, the time dependence governing the original stochastic process is not

considered within SDDiP.

In this paper, we present a multi-stage multi-sector strategic energy planning model in which both the end-use demand for energy services and resource costs are considered uncertain. These two sets of uncertain parameters were identified in [Guevara et al., 2020] as the ones with the most significant impact on the optimal investment strategy. The main contribution of our work is a novel methodology to model the time-dependence of these uncertain parameters. Our methodology takes, as its input, ranges defined in the literature for the uncertain parameters (as well as their nominal values), and constructs auto-regressive (AR) models with the property that the random variables corresponding to each period are—conditionally on the value of the previous observation— distributed *uniformly* within those pre-specified ranges. The latter property creates new challenges since the noise of the AR process becomes dependent on the variables of the previous stage; to circumvent this issue, we introduce a coefficient α in the AR construction for controlling the zigzag effects in the evolution of uncertain processes, and show that the resulting process can be represented in the SDDP model with stage-wise independent variables. As it turns out, the coefficient α can be used to control the likelihood of extreme scenarios. To preserve the convexity of the stochastic problem, we discretize the AR models associated with the cost parameters involved in the objective function using Markov chains. The resulting formulation is then solved with an advanced SDDP algorithm available in the literature that handles finite-state Markov chains. Finally, we perform numerical experiments on a Swiss case study to assess the benefit of considering time-dependence in stochastic processes.

The rest of the paper is organized as follows. In Section 2, we describe the deterministic formulation of the multi-stage strategic energy planning problem from the literature, which does not take uncertainties into account, and we introduce a stochastic multi-stage formulation. In Section 3 we present our methodology for modeling uncertainty based on auto-regressive (AR) models. Section 4 discusses the implementation of these AR models within an SDDP framework and how to approximate an AR process by finite state-space Markov chains (in discrete time). Section 5 is dedicated to a numerical application to the Swiss energy system, and in Section 6 we provide summary results and conclusions for future research.

2 Multi-stage strategic energy planning

2.1 A deterministic multi-stage model

In this section, we present the deterministic formulation of the multi-stage strategic energy planning model which is an extension of the one-period model, EnergyScope, first introduced by Moret et al. [2017] and applied to the Swiss energy system. EnergyScope is a multi-sector and multi-energy model that is driven by end-user demand in energy services (electricity, heating and transportation), by the efficiency and cost of technologies (generation and storage) and by the cost and availability of resources (imported and local). In this model, heating demand is split into centralized, decentralized or industrial. The first two heating demands include end-use demand for space heating and hot water and belongs to the category of low-temperature heat. The third is the end-use demand for industrial process heating and belong to the category of high-temperature heat. Transport end-use demand is also divided into passenger mobility (public and private) and freight mobility (rail and road). The model seeks an optimal policy that minimizes investment and operating costs over the entire planning horizon, i.e. 2020-2050. The operating decisions are taken on a monthly basis to represent long-term energy storage through hydroelectric dams, which is relevant for Switzerland. Investment decisions are related to the installation of new technologies, which are affected by strategic decisions. Strategic decisions concern to (i) the proportion of low-temperature heat demand to be dedicated to decentralized heat (district heating network (DHN)) and centralized heat demand, (ii) the proportion of passenger mobility to be dedicated to private and public sector demand, and (iii) the proportion of freight mobility to be dedicated to road and rail freight demand.

A simplified formulation of the model is presented below. A detailed formulation is presented in Appendix A.

$$\min_{\substack{\mathbf{C}_t^{INV},\mathbf{C}_t^{SV},\mathbf{C}_t^{O\$M},\mathbf{x}_{t,p},\\ \mathbf{y}_{t,p},\mathbf{z}_{t,r,m},\mathbf{Sto}^{+}_{t,u,l,m},\\ \mathbf{Sto}^{-}_{t,u,l,m},\mathbf{D}_{t,l,m},\mathbf{Loss}_{t,l,m},\\ \mathbf{GWP}^{OP}_{t,r}}} \quad \sum_{t\in\mathcal{T}} \tau(t)\left(\mathbf{C}_t^{Inv} - \mathbf{C}_t^{SV} + \mathbf{C}_t^{O\&M}\right) \tag{m.1}$$

s.t.

$$\mathbf{C}_t^{Inv} = \sum_{p\in\mathbf{P}} c_{t,p}^{Inv}\cdot \mathbf{y}_{t,p} \qquad \forall t\in\mathcal{T}, \tag{m.2}$$

$$\mathbf{C}_t^{O\&M} = \varsigma_t \cdot \left(\sum_{p\in\mathbf{P}} c_{t,p}^{Maint}\cdot \mathbf{x}_{t,p} + \sum_{r\in\mathbf{RES}}\sum_{m\in\mathbf{M}} c_{t,r}^{Op}\cdot \mathbf{z}_{t,r,m}\cdot h_m^{Op}\right) \qquad \forall t\in\mathcal{T}, \tag{m.3}$$

$$\mathbf{C}_t^{SV} = \sum_{p\in\mathbf{P}} sv_{t,p}\cdot c_{t,p}^{Inv}\cdot \mathbf{y}_{t,p}\cdot \mathbb{1}_{\{t+lt_p>T+1\}}(t) \qquad \forall t\in\mathcal{T}, \tag{m.4}$$

$$\mathbf{x}_{0,p} = res_{0,p}, \qquad \forall p\in\mathbf{P}, \tag{m.5}$$

$$\mathbf{x}_{t,p} = \mathbf{x}_{t-1,p} + \mathbf{y}_{t,p} - res_{t-1,p} + res_{t,p} - \mathbf{y}_{t-lt_p,p}\cdot \mathbb{1}_{\{t>lt_p\}}(t) \qquad \forall p\in\mathbf{P}, \forall t\in\mathcal{T}, \tag{m.6}$$

$$f_{t,p}^{min} \le \mathbf{y}_{t,p} \le f_{t,p}^{max} \qquad \forall p\in\mathbf{P}, \forall t\in\mathcal{T}, \tag{m.7}$$

$$\bar{f}_p^{min} \le \mathbf{x}_{t,p} \le \bar{f}_p^{max} \qquad \forall p\in\mathbf{P}, \forall t\in\mathcal{T}, \tag{m.8}$$

$$\mathbf{z}_{t,p,m} \le \mathbf{x}_{t,p}\cdot k_{p,m} \qquad \forall p\in\mathbf{P}, \forall m\in\mathbf{M}, \forall t\in\mathcal{T}, \tag{m.9}$$

$$\sum_{m\in\mathbf{M}} \mathbf{z}_{t,p,m}\cdot h_m \le \mathbf{x}_{t,p}\cdot \widehat{k}_p \sum_{m\in\mathbf{M}} h_m \qquad \forall p\in\mathbf{P}, \forall t\in\mathcal{T}, \tag{m.10}$$

$$\sum_{m\in\mathbf{M}} \mathbf{z}_{t,r,m}\cdot h_m \le avail_r \qquad \forall r\in\mathbf{RES}, \forall t\in\mathcal{T}, \tag{m.11}$$

$$\mathbf{z}_{t,u,m} = \mathbf{z}_{t,u,m-1} + h_m \sum_{\substack{l\in\mathbf{L}\\ \eta^{+}_{u,l}>0}} \mathbf{Sto}^{+}_{t,u,l,m}\eta^{+}_{u,l} - h_m \sum_{\substack{l\in\mathbf{L}\\ \eta^{-}_{u,l}>0}} \mathbf{Sto}^{-}_{t,u,l,m}/\eta^{-}_{u,l} \qquad \forall u\in\mathbf{STO}, \forall m\in\mathbf{M}, \forall t\in\mathcal{T}, \tag{m.12}$$

$$\sum_{i\in\mathbf{P}\cup\mathbf{RES}\setminus\mathbf{STO}} f_{i,l}\cdot \mathbf{z}_{t,i,m} = -\sum_{u\in\mathbf{STO}} (\mathbf{Sto}^{-}_{t,u,l,m} - \mathbf{Sto}^{+}_{t,u,l,m}) + \mathbf{D}_{t,l,m} + a_l\cdot \mathbf{Loss}_{t,l,m} \forall l\in\mathbf{L}, \forall m\in\mathbf{M}, \forall t\in\mathbf{T}, \tag{m.13}$$

$$\mathbf{Loss}_{t,l,m} = \sum_{i\in\mathbf{RES}\cup\mathbf{P}\setminus\mathbf{STO}, f_{i,l}>0} f_{i,l}\cdot \mathbf{z}_{t,i,m}\cdot \%loss_l \qquad \forall m\in\mathbf{M}, \forall l\in\mathbf{L}\setminus\mathbf{RES}, t\in\mathbf{T}, \tag{m.14}$$

$$\mathbf{GWP}_{t,r}^{Op} = \sum_{m\in\mathbf{M}} gwp_{r,m}^{Op}\cdot \mathbf{z}_{t,r,m}\cdot h_m \qquad \forall r\in\mathbf{RES}, \forall t\in\mathcal{T} \tag{m.15}$$

The objective function of the energy problem minimizes the total discounted system cost (i.e., investment and operation costs) minus the salvage value of the residual life of installed technologies at the end of the planning horizon (m.1), where $\tau(t)$ is the discount factor at stage t, with $t \in \mathcal{T} = \{1, ..., T\}$. The investment cost $\mathbf{C}_t^{Inv}$ in period $t \in \mathcal{T}$ is the sum of the product of the investment cost of technology $p \in \mathbf{P}$ with the new capability of its corresponding technology (m.2). The O&M cost $\mathbf{C}_t^{O\&M}$ in period t is equal to the amount of the system's maintenance and operating costs multiplied by the annualization factor ς_t, where $r \in \mathbf{RES}$ represents the energy resource and $m \in \mathbf{M}$ represents the months of the year (m.3). The discounted salvage $\mathbf{C}_t^{SV}$ in period t assigns a non-zero value to the technologies still available at the end of the planning horizon (m.4).

The set of restrictions (m.5)-(m.6) refers to the initial design of the capacity and its subsequent expansion, which considers the new capacity installed and the decommissioning plan. The available capacity $\mathbf{x}_{t,p}$, as shown in (m.6), is continuously updated by balancing the accumulated capacity $\mathbf{x}_{t-1,p}$ in period $t-1$, the new capacity installed $\mathbf{y}_{t,p}$, the elimination of unused capacity $res_{t,p} - res_{t-1,p}$ at the beginning of the period t and the retirement capacity of technologies that were installed for optimization but have reached their useful life. When $t = 0$, the initial capacity is the existing infrastructure at that moment as shown in (m.5). Once a technology is installed, it can be operated until its end-of-life. The indicator function $\mathbb{1}_{\{condition\}}(t)$ takes the value 1 when the *condition* is met and the value 0 when it is not met. Constraint (m.7) sets the upper and lower bounds to the new capacity installed for each period t and technology p. For the available capacity, we also use the upper and lower bounds over the entire planning horizon as shown (m.8). The decision variable $\mathbf{z}_{t,\cdot,m}$ determines how resources, the unit's storage and technologies are used in each period. Constraints (m.9)-(m.10) link the total capacity available of a technology to its actual use in each month and period via two capacity factors: a capacity factor for each month $k_{p,m}$ depending on resource availability (e.g. renewable) and a yearly capacity factor $\widehat{k}_p$ accounting for technology downtime and maintenance. The constraint (m.11), is used to limit the total use of resources by their availability during the period *avail*. Constraint (m.12) is the balancing equation for storage units $u \in \mathbf{STO}$. In this constraint when $m = 1$ then the initial condition $\mathbf{z}_{t,u,0}$ is equal to $\mathbf{z}_{t,u,12}$, $\forall t \in \mathcal{T}$ and $\forall u \in \mathbf{STO}$.

The variables $(\mathbf{Sto}^{+}_{t,u,l,m}, \mathbf{Sto}^{-}_{t,u,l,m})$ determine the inputs to or outputs from the storage units. Layers $\mathbf{L}$ are defined as the elements in the system that need to be balanced in each month, such as resources and end-use demand. For example, electricity imported or produced on the system (as well as end-use demand for electricity) can be stored in hydroelectric dams or used as inputs to other energy conversion technologies (such as heat pumps) to meet the end-use demand for electricity. Constraint (m.13) expresses the balance for each layer: all outputs from resources and technologies (including storage) are used to satisfy the end-uses-demand $\mathbf{D}_{t,l,m}$ or as inputs to other resources and technologies. The matrix $f \in \mathcal{M}_{|\mathbf{P} \cup \mathbf{RES} \setminus \mathbf{STO}| \times |L|}(\mathbb{R})$ defines for all technologies and resources outputs to (positive) and inputs from (negative) layers. Losses $\mathbf{Loss}_{t,l,m}$ are considered for the electricity grid and for the DHN (District heating network). Constraint (m.14) calculates the amount of electricity that is lost from both produced and imported electricity in the corresponding layers (a_l is equal to one for these corresponding layers). The total emissions of resources $\mathbf{GWP}^{Op}_{t,m}$ is calculated as the sum of the use $\mathbf{z}_{t,r,m}$ over different months multiplied by the month duration h_m and the emissions of the resource $gwp^{Op}_{r,m}$ (m.15). Note that in our model there are no emission targets, but this expression allows us to estimate the amount of emissions produced by the system in each period t.

2.2 A stochastic multi-stage model

The problems of real world energy systems are usually of a dynamic nature, where uncertain parameters are revealed sequentially and decisions must be adjusted to the recent realization. Multistage Stochastic Linear Programming captures well the nature of sequential decision-making under uncertainty, where the uncertainty is modeled by the overall stochastic process. Let T be a time horizon that spans several years, then consider the following T-stage stochastic linear programming model, written in a recursive

form:

$$\min_{(y_1(\xi_1),z_0(\xi_1))\in\mathcal{X}_1(y_0(\xi_0),\xi_1)} f_1(y_1(\xi_1),z_0(\xi_1),\xi_1)+\mathbb{E}_{\xi_2|\xi_1}\left[\min_{(y_2(\xi_{[2]}),z_1(\xi_{[2]}))\in\mathcal{X}_2(y_1(\xi_1),\xi_2)} f_2(y_2(\xi_{[2]}),z_1(\xi_{[2]}),\xi_2)+\right.$$
$$\left.\mathbb{E}_{\xi_3|\xi_{[2]}}\left[\ldots+\mathbb{E}_{\xi_T\xi_{[T-1]}}\left[\min_{(y_T(\xi_{[T]}),z_{T-1}(\xi_{[T]}))\in\mathcal{X}_T(y_{T-1}(\xi_{[T-1]}),\xi_T)} f_T(y_T(\xi_{[T]}),z_{T-1}(\xi_{[T]}),\xi_T)\right]\right]\right]. \quad (2)$$

In the above formulation, ξ_t represents a random vector containing some of (or all) the stochastic parameters at stage t, $\xi_{[t]} = (\xi_1, \xi_2, ..., \xi_t)$ denotes the history of the stochastic process up to time t and $\mathbb{E}_{\xi_t|\xi_{[t-1]}}$ denotes the conditional expectation operation in stage t with respect to ξ_t given the information $\xi_{[t-1]}$ of stage $t-1$. We distinguish two sets of decision variables in each stage, namely, the state variable, denoted by $y_t(\xi_{[t]}) = [\mathbf{y}_{t,p}, \mathbf{x}_{t,p}]$, which links successive stages (e.g. capacity installed, available capacity), and the local or stage variable, denoted by $z_t(\xi_{[t]}) = [\mathbf{C}_t^{INV}, \mathbf{C}_t^{O\&M}, \mathbf{C}_t^{SV}, \mathbf{z}_{t,\cdot,m}, \mathbf{Sto}^+_{t,u,l,m}, \mathbf{Sto}^-_{t,u,l,m}, \mathbf{D}_{t,l,m}, \mathbf{Loss}_{t,l,m}, \mathbf{GWP}^{Op}_{t,p}]$, which is only contained in the subproblem at stage t. To ease the notation, we will henceforth use y_t and z_t as a shorthand for $y_t(\xi_{[t]})$ and $z_t(\xi_{[t]})$ respectively. The function f_t and the set $\mathcal{X}_t$ denote the objective function and set of feasible decisions associated with stage t, respectively. The function f_1 and the set $\mathcal{X}_1$ are deterministic (i.e. ξ_1 is a known value). y_0 represent the initial state of the system. In addition, we assume that the objective function f_t is linear of the form $f_t(y_t, z_{t-1}, \xi_t) = c_t^\top(\xi_t)y_t + d_t^\top(\xi_t)z_{t-1}$ which represent the cost function associated with stage t, and the $\mathcal{X}_t$ set is of the form

$$\mathcal{X}_t(y_{t-1}, \xi_t) := \{(y_t, z_{t-1}) : B_t(\xi_t)y_{t-1} + A_t(\xi_t)y_t + C_t(\xi_t)z_{t-1} = b_t(\xi_t), y_t \geq 0, z_{t-1} \geq 0\}, \forall t = 2, ..., T \quad (3)$$

where $\mathcal{X}_t(\cdot,\cdot)$ describes the feasible set for decisions to be made at stage t. We introduce two different time scales, a long-term time scale for investment decisions and a short-term time scale for operations decisions. As an example, the dynamics between investment and operation decisions using these two time scales are illustrated in the Figure 1.

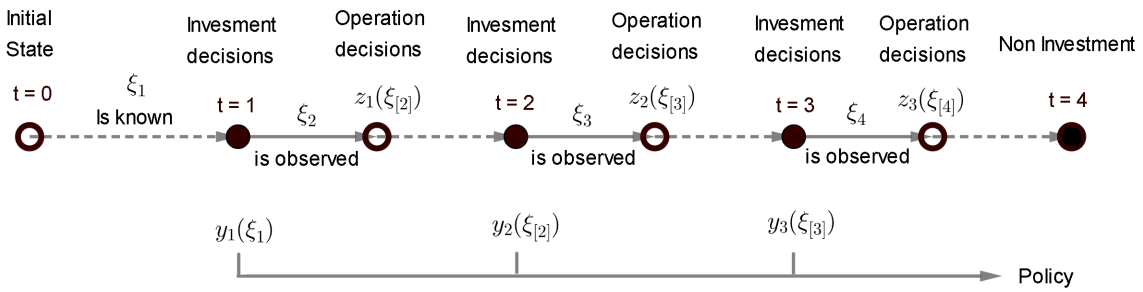


Figure 1: A small example of four stages with three periods of investment, followed by operating decisions conditioned on revealed uncertainty, which shows the dynamic process of the model.

As shown in the figure above, $t = 0$ represents the initial state of the system, that is, the total available capacity, which is considered as input data. Uncertainty is sequentially revealed over time. Investment decisions are made at the beginning of each stage before the uncertainties are revealed. On the other hand, operational decisions are made during the next stage after the uncertainties are realized. Here $\{y_t, z_t\}_{t=1}^T$ describe a policy, a solution to the MSLP. We assume relatively complete recourse for any policy y_t, z_t. $\forall t = 2, ..., T$. Such an assumption is reasonable in our context since it is equivalent to assuming that one can always import energy (albeit at a high cost) to satisfy demand.

3 Uncertainty modelling through auto-regressive processes

3.1 The goals of the model

Multi-stage models under uncertainty imply a sequence of decisions to be made at each period, considering the state of the system and some uncertain parameters that have a serial correlation. Since in our model (2), the investment and operational decisions are made sequentially based on the information available on the random variables up to stage t and the state variables, we need a model that describes the evolution of stochastic processes in each stage. [Moret et al., 2017] only defines the ranges of variation of the uncertain parameters, but does not describe how the process evolves over the planning horizon. Assuming that the uncertain parameters are independent across stages may be unrealistic. Therefore, to describe the evolution of the processes, we use the ranges of variation of the uncertain parameters reported by [Moret et al., 2017] to construct several auto-regressive processes. In this section, we assume that the end-use demand for energy services and the costs of all resources are time-dependent stochastic processes.

Suppose that a stochastic process ξ_t is driven by an auto-regressive process of order 1, denoted by AR(1), and has the following representation

$$\xi_{t+1} = a_t \cdot \xi_t + \epsilon_{t+1}, \quad \forall t = 0, \dots, T-1, \tag{4}$$

where the AR coefficient a_t and the standard deviation of the innovation (or error term) ϵ_{t+1} are both allowed to depend on time t; ϵ_t is independent across t. The initial realization ξ_0 may be deterministic. This AR model tries to explain the future value of ξ_{t+1} using the response variable of the previous stage ξ_t as a predictor. Forecasting models such as the one in (4) are often used to model inflows in multi-stage models for hydroelectric planning (see, e.g., Shapiro et al. [2013], Löhndorf and Shapiro [2019]).

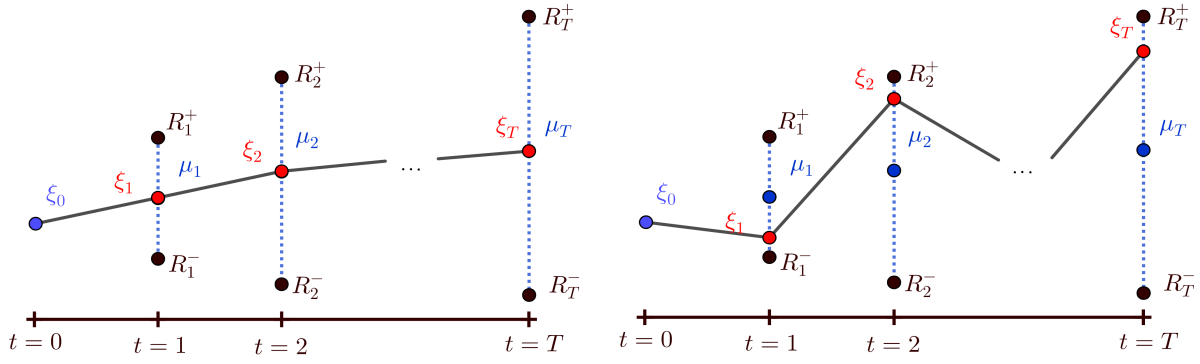
In this study, we want to use AR models to estimate the future energy demand and resource costs, but due to the lack of information from historical data, approaches such as maximum likelihood estimation or least-squares estimation cannot be applied directly to estimate each parameter a_t, and therefore we cannot know the distributions of the error term ϵ_t. To estimate the a_t parameters and the distribution of the errors ϵ_t, we use the ranges of variation and the nominal values of the uncertain parameters found in Moret et al. [2017]. A range of variation is defined by assigning a lower bound R_t^- and an upper bound R_t^+ to the process ξ_t at stage t, and the nominal value μ_t is the most probable realization corresponding to the range $[R_t^-, R_t^+]$. Nominal values, in addition to providing the most probable value of the range at stage t, also indicate that the data is distributed 50% above and below it. When a range is symmetric (mean equals nominal value), we construct the corresponding AR process so that it has uniform distribution within each range $[R_t^-, R_t^+]$ (conditionally on the value of the process in stage $t-1$), while for an uncertain parameter with an asymmetric range (e.g., cost of resources), we use a stochastic process $\xi_t^{'}$ defined as:

$$\xi_t^{'} = \begin{cases} \mu_t \cdot (1 + \xi_t) & \text{if } \xi_t \geq 0, \\ \\ \mu_t / (1 - \xi_t) & \text{otherwise,} \end{cases}$$

where $\{\xi_t\}$ is the process drawn from (4)..

In summary, in order to be able to build a continuous AR stochastic process of the form (4), we assume that the ranges of variation $[R_t^-, R_t^+]$ (or $[-R_t^{'}, R_t^{'}]$), nominal values μ_t of the uncertain parameters and the starting point of the processes at stage $t = 0$ are given. Figure 2a represents the input data necessary to estimate the parameters of the AR model. We want to construct a sequence of $\{\xi_t\}_{t=0}^T$ (or $\{\xi_t^{'}\}_{t=0}^T$) using the following strategy: when the stochastic process realization is near the upper bound of the range in step t, then in the next stage, the process realization cannot be near the lower bound of the range; similarly, when the stochastic process realization is near the lower bound in stage t, then in the next stage, the process realization cannot be near the upper bound. This proposal tries to avoid the zigzag shape that a process presents when it is considered stage-wise independent. Figure 2b shows the ranges

of variation and the nominal values of each process $\{\xi_t\}_{t=1}^T$ through black and blue dots, respectively. An arbitrary realization of the process $\{\xi_t\}_{t=0}^T$ is represented by red dots.



(a) Input data

(b) Evolution of the stochastic process $\{\xi_t\}_{t=0}^T$

Figure 2: In this figure $\mu_1, \mu_2, ..., \mu_T$ are the nominal values, R_t^- and R_t^+ for $t = 1, ..., T$ are the lower and upper bounds of the ξ_t parameter, respectively. ξ_0 is deterministic and represents the initial realization of the process. The figure on the left represents the evolution of the processes in the deterministic case, while the figure on the right, a possible realization of the stochastic process is shown.

3.2 The construction of the AR model

We start by defining a model that describes the evolution of the process without considering the error term ϵ_t. Then we introduce randomness in the initial model to finally derive an AR process with which the degree of dependency can be controlled, thereby avoiding the zigzag shape in the stochastic process realizations.

3.2.1 The deterministic process

Without loss of generality suppose that the range $[R_t^-, R_t^+]$ represents all symmetric ranges (including $[-R_t^{'}, R_t^{'}]$). Since ξ_0 is a given constant, we define ξ_1 as the midpoint of $[R_1^-, R_1^+]$. The idea of our construction is to transport each realization of ξ_t, $t = 1, \ldots, T-1$ to an equivalent place within the range of the next stage while keeping proportions. More specifically, let $t \geq 1$ be the current stage and let $Z(\xi_t)$ be the equivalent place of ξ_t within the range of the next stage. We impose that the ratio between $(R_t^+ - R_t^-)$ and $(R_{t+1}^+ - R_{t+1}^-)$ be proportional to the ratio between $(\xi_t - R_t^-)$ and $(Z(\xi_t) - R_{t+1}^-)$, i.e.,

$$\frac{R_t^+ - R_t^-}{R_{t+1}^+ - R_{t+1}^-} = \frac{\xi_t - R_t^-}{Z(\xi_t) - R_{t+1}^-}. \tag{5}$$

Thus, we obtain the model without the random factor:

$$\xi_{t+1} := Z(\xi_t) := R_{t+1}^- + \frac{R_{t+1}^+ - R_{t+1}^-}{R_t^+ - R_t^-}\left(\xi_t - R_t^-\right). \tag{6}$$

It is easy to see that in the case of deterministic processes each element ξ_t, $t = 1, \ldots, T$ is equal to the midpoint of the range corresponding to stage t, which is also the nominal value μ_t (see Figure 2a).

3.2.2 The stochastic process

Our goal now is to add a random factor to the equation (6). To do so, we need to construct a random variable ω_t that allows us to move $Z(\xi_t)$ without leaving the range corresponding to the stage t. For

that purpose, we introduce for $t \geq 1$ a *variability-controlling* parameter $\alpha \in [0,1]$ that limits the support of $Z(\xi_t) + \omega_t$ to a proportion α of the range $[R_{t+1}^-, R_{t+1}^+]$. More specifically, we define ξ_1 as a uniform random variable on $[R_1^-, R_1^+]$; then, for $t \geq 1$, given the value of $Z(\xi_t)$ computed in (6), we add a random perturbation ω_t to $Z(\xi_t)$ as follows:

$$\xi_{t+1} := Z(\xi_t) + \omega_t, \text{ with } \omega_t \sim \text{Uniform}(-\alpha_t \cdot \rho_t, \alpha_t \cdot (1-\rho_t)), \tag{7}$$

where $\alpha_t := \alpha(R_{t+1}^+ - R_{t+1}^-)$, and

$$\rho_t := \frac{Z(\xi_t) - R_{t+1}^-}{R_{t+1}^+ - R_{t+1}^-}.$$

It follows that

$$\xi_{t+1} \sim \text{Uniform}\left(\alpha R_{t+1}^- + (1-\alpha)Z(\xi_t),\ \alpha R_{t+1}^+ + (1-\alpha)Z(\xi_t)\right), \tag{8}$$

so we see that, for any given value of ξ_t, the support ξ_{t+1} is contained in $[R_{t+1}^-, R_{t+1}^+]$ and has length $\alpha(R_{t+1}^+ - R_{t+1}^-)$ as desired. Figure 3 illustrates the basic idea for the construction of $\{\xi_t\}$, with $\alpha = 0.75$ being used in this example.

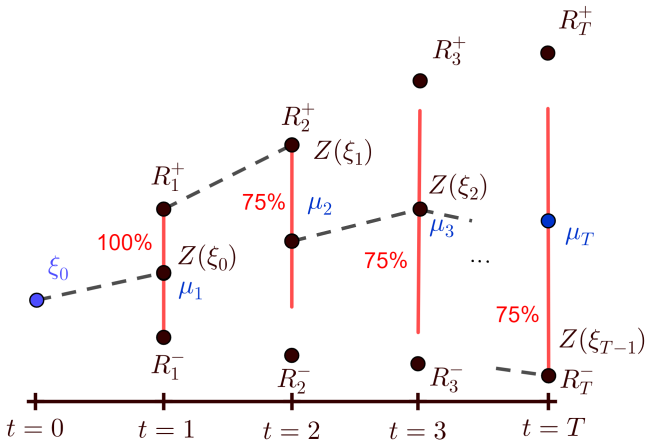


Figure 3: The figure represents the possible relations of $\{\xi_t\}$ in red lines. The first stage covers 100% of the range, while the following stages always cover 75% of the range. It also shows a realization of process.

Equations (6)-(8) make the role of α clear. It allows the decision maker to control the variability of the underlying stochastic processes based on his/her beliefs, but at the same time α has a number of properties, listed below:

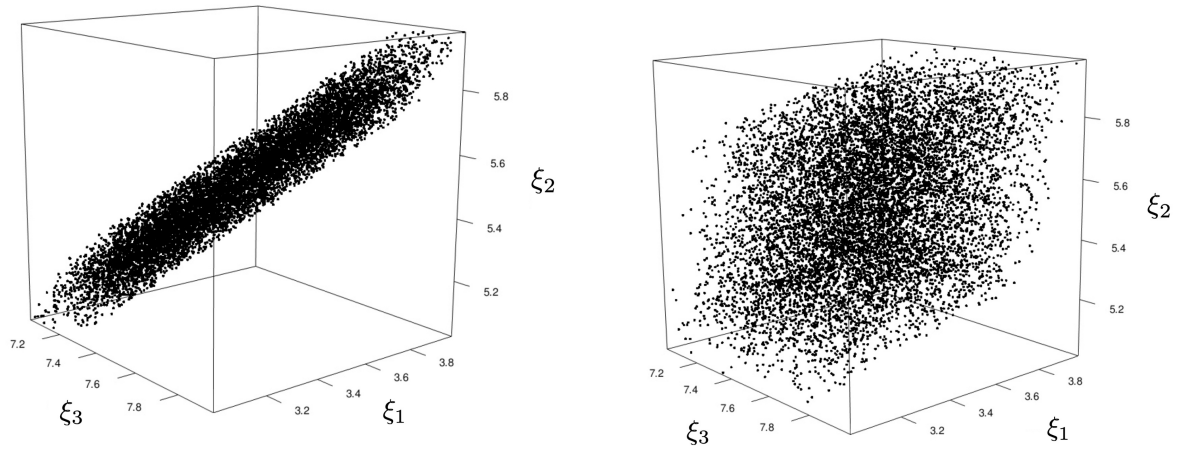
(i) If $\alpha = 0$, then for all $t \geq 2$ we have that ξ_t is a linear transformation of the uniform random variable ξ_1 with a positive multiplicative constant; thus, ξ_t is (unconditionally) uniformly distributed, has correlation one with ξ_1, and by applying an inductive argument in (6) we see that the support of ξ_t is the range $[R_t^-, R_t^+]$.

(ii) If $\alpha = 1$, the process $\{\xi_t\}$ is independent across periods and each ξ_t is uniformly distributed on $[R_t^-, R_t^+]$.

(iii) If $\alpha \in (0,1)$, then the unconditional distribution of ξ_t is symmetric but not uniformly distributed, as it is defined as the sum of uniformly distributed random variables.

(iv) As α decreases, the probability that either all ξ_t, $t = 1, \ldots, T$ are close to their maximum values or all ξ_t are close to their minimum values, increases.

Property (iv) above—which is a direct consequence of the fact that the correlation among the ξ_ts increase as α decreases—is particularly important in the context of strategic energy planning, since the scenarios with high investment and operations costs and high demand clearly are the scenarios with the highest system cost; conversely, those scenarios with low costs and low demand clearly are the scenarios with the lowest system cost. Thus, by choosing lower values of α, the decision-maker can make sure that scenarios that are close to the ones with extreme costs are generated more often in the course of the algorithm. To illustrate property (iv) above, Figure 4 depicts 10,000 realizations of a process $\{\xi_t\}_{t=1}^3$ defined with $[R_t^-, R_t^+] = [3, 4], [5, 6], [7, 8]$, $t = 1, 2, 3$ respectively, for $\alpha = 0.25$ and $\alpha = 0.75$. We see that in the case of lower α the correlation among the ξ_t is high and indeed the extreme scenarios are generated more often.



(a) $\alpha = 0.25$. (b) $\alpha = 0.75$.

Figure 4: 10,000 realizations of $\{\xi_t\}$, $t = 1, 2, 3$ for two different values of α.

Finally, we write the AR model in the form (4). For that, we need to rewrite (7), since in the latter formulation the random "noise" term ω_t depends on the previous values of the process. One way to avoid this is by using the inverse transform method according to which we can write

$$\text{Uniform}(a, b) \;=\; a + (b - a) \cdot U, \;\text{ with } U \sim \text{Uniform}(0, 1).$$

It follows that

$$\xi_{t+1} = Z(\xi_t) - \alpha_t \cdot \left(\frac{Z(\xi_t) - R_{t+1}^-}{R_{t+1}^+ - R_{t+1}^-} \right) + \alpha_t \cdot U, \;\text{ with } U \sim \text{Uniform}(0, 1). \tag{9}$$

Then, by substituting $Z(\xi_t)$ from (6) into (9) and doing some simplifications in the latter equation, we obtain the following expression

$$\begin{aligned} \xi_{t+1} &= \left(\frac{R_{t+1}^+ - R_{t+1}^- - \alpha_t}{R_t^+ - R_t^-} \right) \xi_t + \left(-\frac{R_{t+1}^+ - R_{t+1}^- - \alpha_t}{R_t^+ - R_t^-} \right) R_t^- + R_{t+1}^- + \alpha_t \cdot U \\ &= \underbrace{\left(\frac{(1-\alpha)\Delta_{t+1}}{\Delta_t} \right)}_{a_t} \xi_t + \underbrace{R_{t+1}^- - \frac{(1-\alpha)\Delta_{t+1}}{\Delta_t} R_t^- + \alpha \Delta_{t+1} \cdot U}_{\epsilon_{t+1}}, \end{aligned} \tag{10}$$

where $\Delta_t := R_t^+ - R_t^-$.

4 SDDP and Markov Chain

We discuss now solution methods for the multi-stage stochastic model described in Section 2, where the uncertainty components are modeled using the techniques presented in Section 3. As discussed in Section 1, we use the Stochastic Dual Dynamic Programming (SDDP) algorithm for that purpose. The SDDP algorithm performs a series of iterations, each consisting of one step forward and one step backward through the stages. It decomposes the problem into stages to construct an approximately optimal policy using Benders' cuts. In addition, it incorporates Monte Carlo sampling to approximate the original process using only a subset of the scenarios at each iteration, which makes it a method capable of solving large-scale problems. When the stochastic process is stage-wise independent the expected future cost function is independent of the realization of the process, and therefore the cuts that are generated for a subproblem are also valid cuts for other subproblems in the same stage. This property of sharing cuts is fundamental for an efficient implementation of the SDDP. Despite its advantages, an important limitation of the SDDP is related to the modeling of the dynamics of the state variables.

When the stage-wise independent condition is relaxed, then the expected future cost functions $\mathcal{Q}_t(y_t, \xi_{t+1})$ also depend on past realizations of stochastic process. The most common way to preserve dependence within the SDDP is to assume that the stochastic process is modeled as an AR process of order 1 or AR(1). An AR(1) is incorporated into the SDDP model by defining the stochastic process as a state variable, the noise or innovation ξ_t is considered stage-wise independent and the recursive equation (as given by equation (4)) is added to the model. Note however that, if the new state variable multiplies any other decision variable, bilinear terms destroy the convexity of the future cost function. Therefore, to maintain the mathematical properties that make the method efficient, only the processes that are on the right-hand side of the constraints can be expressed as an affine function of the errors. The price of such reformulation is an increase in the number of state variables.

Another way to preserve the dependence in the SDDP without destroying the convexity of the problem is to model the stochastic process as a finite-state Markov chain [Löhndorf and Shapiro, 2019, Philpott and de Matos, 2012]. A stochastic process $\{\xi_t\}$ is called Markovian if the conditional distribution of ξ_t given $\xi_{[t-1]}$ is the same as that of ξ_t given ξ_{t-1} for $t = 1, ...T$. When a data process is assumed to be Markovian, a future cost function must be enumerated for each value taken by the state variable, which increases the dimensionality of the problem. It is the main computational disadvantage of this approach. Also, if the process is multi-dimensional, the number of Markov states needed to obtain a good enough solution becomes prohibitive.

In the previous section, we constructed a general AR(1) process to model the energy demands and resource costs. Given the ways in which the uncertainty of stochastic processes can be included in the SDDP, we use the state-variable approach for the uncertainty in the demand, which appears only in the right-hand side of the constraints; for the uncertainty in costs, we apply the aforementioned strategy to approximate the AR(1) process of the costs by a Markov chain to maintain the convexity of the model. Therefore, in the next section, we explain how to construct Markov Chains with finite state space from AR(1) processes. As this process is multidimensional, we reduced the dimensionality of the process by choosing two uncertain parameters, specifically the cost of fossil fuel (all fossil fuels costs are assumed to be correlated) and electricity, which turn out to be the parameters with the most significant impact on investment decisions [Guevara et al., 2020]. Furthermore, we assume that the two stochastic processes are independent of each other, which allows us to perform the discretization of each one separately.

4.1 Finite state Markov-chain approximations

It is well known that Markov chains can be used to approximate AR processes of order 1. When an AR(1) process is stationary, this means that the transition probability distribution is constant for all time

instants. Different methods have been proposed to discretize a stationary process AR(1). Among them, we can mention Tauchen [1986], Tauchen and Hussey [1991], Kopecky and Suen [2010] y Rouwenhorst [1995]. An extension of the Tauchen method (Tauchen [1986]) to a non-stationary process AR(1) was proposed by Fella et al. [2019], for a problem class where the variance increases with time and the persistence component is non-stationary. The main difference between the extension and its stationary counterpart is that the range of the equidistant state space can vary over time and therefore so do the transition probabilities.

Fella et al. [2019] consider an AR(1) process of the following form

$$\eta_t = \rho_t \eta_{t-1} + \epsilon_t, \epsilon_t \sim \text{Normal}(0, \sigma_{\epsilon t}), \tag{11}$$

where the standard deviation of the innovation $\sigma_{\epsilon t}$ and the AR coefficient ρ_t are both allowed to depend on time t, and where the unconditional standard deviation of η_t is σ_t. In addition, it is not required that $|\rho_t| < 1$ (as in the stationary case). Then, they approximate the non-stationary AR(1) process by a Markov chain with N states in each stage t, but with a time-dependent state space $\Upsilon_t^N = \{\bar{\eta}_t^1, ..., \bar{\eta}_t^N\}$ (uniformly-spaced) and transition matrix Π_t^N. The time-varying grid points are defined as

$$\bar{\eta}_t^N = -\bar{\eta}_t^1 = \Omega \sigma_t, \tag{12}$$

where Ω is a positive constant. The step size $h_t = 2\Omega\sigma_t/(N-1)$ and the transition probabilities for any $i, j = 1, ..., N$ are

$$\pi_t^{ij} = \begin{cases} \Phi\left(\frac{\bar{\eta}_t^j - \rho\bar{\eta}_{t-1}^i + h_t/2}{\sigma_{\epsilon t}}\right) & \text{if } j = 1, \\ 1 - \Phi\left(\frac{\bar{\eta}_t^j - \rho\bar{\eta}_{t-1}^i - h_t/2}{\sigma_{\epsilon t}}\right) & \text{if } j = N, \\ \Phi\left(\frac{\bar{\eta}_t^j - \rho\bar{\eta}_{t-1}^i + h_t/2}{\sigma_{\epsilon t}}\right) - \Phi\left(\frac{\bar{\eta}_t^j - \rho\bar{\eta}_{t-1}^i - h_t/2}{\sigma_{\epsilon t}}\right) & \text{otherwise.} \end{cases}$$

where Φ denotes the cumulative distribution function for the standard normal distribution. Our goal is to use this method to discretize the AR(1) processes constructed in Section 3 to model the evolution of costs; however, since the AR(1) processes in (10) involve uniform distributions instead of normal ones, we must adapt this method to our setting, as we explain below.

Discrete approximations of AR(1) processes

In this section, we explain how to adapt the Fella et al. [2019] method to discretize a one-dimensional AR process, where innovations are independent and identically distributed non-Gaussian random variables.

Let d_t be the process that describes the costs of natural gas (or electricity). Before discretizing d_t, it is necessary to specify an N-state Markov chain in each stage t. Let $\widehat{d}_t$ be the discrete-valued process that approximates d_t in each stage t. Let $\Upsilon_t^N = \{\bar{d}_t^1, \bar{d}_t^2, ..., \bar{d}_t^N\}$ be the finite set of possible realizations of $\widehat{d}_t$ in time t, where $\bar{d}_t^1 < \bar{d}_t^2 < ... < \bar{d}_t^N$. To determine the time-varying points of the grid, we set $\bar{d}_t^N$ and $\bar{d}_t^1$ to be equal to the maximum and minimum values of the range $[-R_t^{'}, R_t^{'}]$ associated with the process ξ_t. The other values are calculated using the formula

$$\bar{d}_t^i = -R_t^{'} + (i-1) \cdot s_t \tag{13}$$

where $s_t = (2R_t^{'})/(N-1)$. It is the first modification we make to the method described in Fella et al. [2019] since we rely directly on the range of the parameter. Note that for any t, $\bar{d}_t^1 = -R_t^{'}$ and $\bar{d}_t^N = R_t^{'}$. Another important change in the method is the assumption of the uniform distribution instead of the

normal distribution to describe the error distribution ϵ_t. However, the change in the distribution causes the original step length h_t to lie outside the domain of the distribution at extreme points. To fix this problem, we choose a step length that constructs intervals of equal size but within the domain of the chosen distribution.

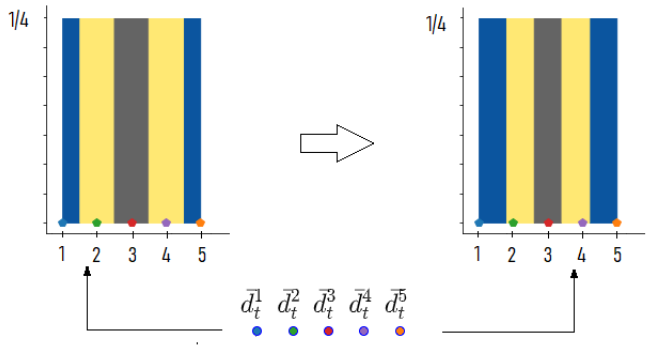


Figure 5: An example with a 5-state Markov chain located at the base of the uniform distribution. The uniform distribution is represented with different colors. The colors represent the intervals formed with the selected step length.

Figure 5 shows two ways of assigning probabilities to a 5-state Markov chain using different step lengths h_t. Both plots represent the density of the uniform distribution on the interval [1,5]. Each state variable d_t^t is mapped to the support of ϵ_t and is represented with the same initial color but with a different shape. Basically, the discretization of a process consists of assigning a probability to $\widehat{d}_t$ of falling in the interval $(\bar{d}_t^j - h_t, \bar{d}_t^j + h_t)$ conditionally on $\widehat{d}_{t-1} = \bar{d}_{t-1}^i$. In addition, the colored bars represent probabilities formed by the intervals. Note that if the step length $h_t = \frac{1}{2}|\bar{d}_t^N - \bar{d}_t^1|/(N-1)$ for $N \leq 5$, then the method tends to give more probability to the central values than to the extreme values. For example, in the figure on the left side, the probability that $\widehat{d}_t$ falls in the interval $(\bar{d}_t^4 - h_t, \bar{d}_t^4 + h_t)$ is much higher than the probability that $\widehat{d}_t$ falls in the interval $(\bar{d}_t^5 - h_t, \bar{d}_t^5 + h_t)$ since half of the interval is out of distribution support. But by redistributing intervals of equal size within the domain of the uniform distribution, now the probability of falling into the extremes will be slightly higher. The figure on the right also shows that the step length at the extreme states of the Markov chain is the same, whereas the step lengths around the other Markov states are different.

Given that the step length is not constant for all t and that the parameters of the uniform distribution depend on the time t, we denote $\mathcal{G}_t$ as the cumulative distribution function of ϵ_t in time t. So for any two states i and j, the associated transition probability is

$$\pi_t^{ij} = \begin{cases} \mathcal{G}_t\left(\bar{d}_t^j - a_t \bar{d}_{t-1}^i + h_t\right) & \text{if } j = 1, \\ 1 - \mathcal{G}_t\left(\bar{d}_t^j - a_t \bar{d}_{t-1}^i - h_t\right) & \text{if } j = N, \\ \mathcal{G}_t\left(\bar{d}_t^j - a_t \bar{d}_{t-1}^i + h_t^{j+}\right) - \mathcal{G}_t\left(\bar{d}_t^j - a_t \bar{d}_{t-1}^i - h_t^{j-}\right) & \text{otherwise,} \end{cases} \tag{14}$$

where $h_t = (d_t^N - d_t^1)/N$, $h_t^{j+} = |(j-1)\cdot(d_t^2 - d_t^1) - j \cdot h_t|$, and $h_t^{j-} = |(j-1)\cdot(d_t^2 - d_t^1) - (j-1)\cdot h_t|$.

We can simplify further the expressions in (14) by observing that, as seen in (10), we have

$$\epsilon_t \sim R_t^- - \frac{(1-\alpha)\Delta_t}{\Delta_{t-1}} R_{t-1}^- + \alpha\Delta_t \cdot \text{Uniform}(0,1).$$

By using the fact that (with $U \sim \text{Uniform}(0,1)$ and $b > 0$)

$$P(a + b \cdot U \leq z) \;=\; P\left(U \leq \frac{z-a}{b}\right) \;=\; \begin{cases} 0 & \text{if } z < a \\ \frac{z-a}{b} & \text{if } a \leq z < a+b \\ 1 & \text{if } z > a+b, \end{cases}$$

it follows that we can write

$$\mathcal{G}_t(z) = \begin{cases} 0 & \text{if } z < R_t^- - \left(\frac{(1-\alpha)\Delta_t}{\Delta_{t-1}}\right) R_{t-1}^- \\ \frac{z - R_t^- + \left(\frac{(1-\alpha)\Delta_t}{\Delta_{t-1}}\right) R_{t-1}^-}{\alpha \Delta_t} & \text{if } R_t^- - \left(\frac{(1-\alpha)\Delta_t}{\Delta_{t-1}}\right) R_{t-1}^- \leq z < R_t^- - \left(\frac{(1-\alpha)\Delta_t}{\Delta_{t-1}}\right) R_{t-1}^- + \alpha\Delta_t \\ 1 & \text{if } z > R_t^- - \left(\frac{(1-\alpha)\Delta_t}{\Delta_{t-1}}\right) R_{t-1}^- + \alpha\Delta_t. \end{cases} \tag{15}$$

Now, by using the expression for $\mathcal{G}_t(z)$ derived above in (14), it is straightforward to calculate the transition probabilities π_t^{ij} for each t.

We have now a complete description of our model and the method used to solve it. We call our approach "SDDP-α", to emphasize the importance of the parameter α in the uncertainty modeling, as discussed in Section 3. In the next section, we present an application of our approach.

5 Application to the Swiss energy System

In this section, we apply the proposed time-dependent stochastic SDDP-α approach to the Swiss energy system in which resource costs and demands are uncertain. The objective is twofold. First, we assess the benefit of considering time dependence in the stochastic processes by comparing the optimal solutions and the out-of-sample performances of the SDDP-α model with those obtained with the classical stage-wise independent SDDP approach, called SDDP-IND. In the latter, one optimizes the model assuming that all processes are stage-independent and uncorrelated. Secondly, we analyze in more detail the sensitivity of the α setting on the SDDP-α solutions. Note that the coefficient $\alpha \in (0,1)$ is used to control the amplitude of the zigzag effect in the AR(1) processes. When α is small, only modest variations in time of the stochastic processes are allowed, thus leading to limited zigzag effects. An α close to 1 will on the contrary generate strong variations over time of the values of the uncertain parameters. Note that the particular case of $\alpha = 1$ corresponds to the SDDP-IND model. As discussed in Section 3, small values of α make the generation of extreme scenarios more likely. In our experiments, we consider three intermediate values for α, i.e., 0.25, 0.5 and 0.75.

The uncertainty on the demands (electricity, lighting, and heating), and on the cost of fossil fuels and electricity is modeled through AR(1) processes using the α coefficient. We assume a correlation one among the different fossil fuel costs (gas, coal and diesel). Then the two AR(1) models for fossil fuels and electricity costs are discretized each one into a three-state Markov Chain at each stage, yielding a total of nine states per stage. Assumptions on uncertain parameters (i.e., nominal, min and max values) and the resulting Markov chains are given in Appendices B and D, respectively. Appendix C illustrates potential trajectories predicted by the AR processes as defined in Section 3 for the realization of the stochastic processes.

Existing capacities for generation and end-use technologies have been calculated based on information from [Moret et al., 2016]. The time horizon of the study [2010,2050] is decomposed into five 10-year periods with four investment periods from 2020. For interested reader, all data can be found in Moret et al. [2017][1] and on the website `https://www.energyscope.ch/`.

[1]The model is described in detail in chapter 1 of that thesis, while the data are documented in detail in Appendix A of the same thesis. The code and data are publicly available at `https://github.com/energyscope/EnergyScope/tree/v1.0`

5.1 Implementation details

All stochastic models are written in Julia 1.4.0 and solved with the open-source library SDDP.jl [Dowson and Kapelevich, 2020]. This library implements a Stochastic Dual Dynamic Programming algorithm with possible extensions to AR models and Markov Chain models. We use the single-cut version of the algorithm and only one scenario in each forward pass. The sampling method applied to generate the sample space Ω_t for each stage was the Latin Hypercube Sampling (LHS) [McKay et al., 1979], since it has proven to yield more accurate results in the SDDP context [Homem-de Mello et al., 2011] as it covers the uncertain space more uniformly than Monte Carlo sampling. To obtain an upper bound estimator (UB) on the models arising from SDDP-IND and SDDP-α models, we used the methodology presented by de Matos et al. [2017]. The formula of the optimality gap for the SDDP algorithm is the one proposed by Ding et al. [2019]. The first policy evaluation is done at iteration 500 and then every 50 iterations until the desired optimality gap is reached.

We solve each instance using a single core of Intel Core i7-8750H CPU 2.20 GHz $\times$ 12 with 32 GB RAM and CPLEX(V12.9) to solve linear programming subproblems. For a fair comparison between SDDP-IND and SDDP-α results, we use the same number of realizations and random vector samples (or scenarios) in all methods.

Our goal is to obtain optimal policies, where a policy is a rule that describes what decisions must be made at each stage for each realization of the stochastic process at that point in time. More specifically, a policy is a sequence of functions $\{\Pi_t\}_{t=1}^{T}$ yielding, for any state y_t and observation of the random variable ξ_t, a feasible control z_t. Because the implemented methods are an approximation of the real problem, it is necessary to adopt an evaluation procedure to assess the quality of the policies. Therefore, we generated from the original AR(1) processes (resource demands and costs) 30,000 out-of-sample(α) scenarios with $\alpha \in \{0.75, 0.50, 0.25\}$. To evaluate SDDP-$\alpha$ policies in all out-of-sample(α) scenarios, we search for a Markov Chain scenario in some sense closer to a considered scenario of the continuous process, to substitute Markov Chain states for the considered sample. For all simulations, we employ the same scenario sequence for the forward pass.

5.2 Numerical results

All stochastic models have been solved with a 2% optimality gap objective. Table 1 reports the model characteristics, the final objective function values, the CPU times for the training and the upper bound evaluations and the number of upper bound evaluations performed during the optimization process (first one at iteration 500 and then every 50 iterations).

	SDDP-IND	SDDP-0.75	SDDP-0.50	SDDP-0.25
# of subproblems by stage	40	40	40	40
Objective value (MCHF)	175,478	172,995	171,482	168,753
Upper bound (MCHF)	177,166	174,917	172,952	168,977
Standard deviation (MCHF)	28,926	33,995	41,503	56,561
Optimality gap (%)	1.55	1.81	1.72	1.33
Training-CPU time (sec)	18,721	25,932	59,651	14,503
# of iterations	550	650	1,850	3,250
CPU time for UB evaluations (sec)	1,334	3,019	16,294	11,526
# of UB estimations	2	4	28	56

Table 1: Characteristics and detailed numerical results for the SDDP-IND and SDDP-α models.

As more zigzag effect is permitted, i.e., for SDDP-IND and higher values of α, we observe that the total cost of the solution (i.e. the objective value) increases. We also notice that the number of iterations

increases when α decreases while the total CPU time does not change monotonically with α. CPU times for training vary from 4 hours for SDDP-0.25 to more than 16 hours for SDDP-0.50.

5.2.1 Out-of-sample performances

We now compare the performances of the different stochastic solutions using samples of 30,000 scenarios generated from the original AR processes with $\alpha \in \{0.75, 0.50, 0.25\}$. Figure 6 summarizes the cost performances decomposed among its different components, i.e., Investment, Maintenance, Operation and Total costs.

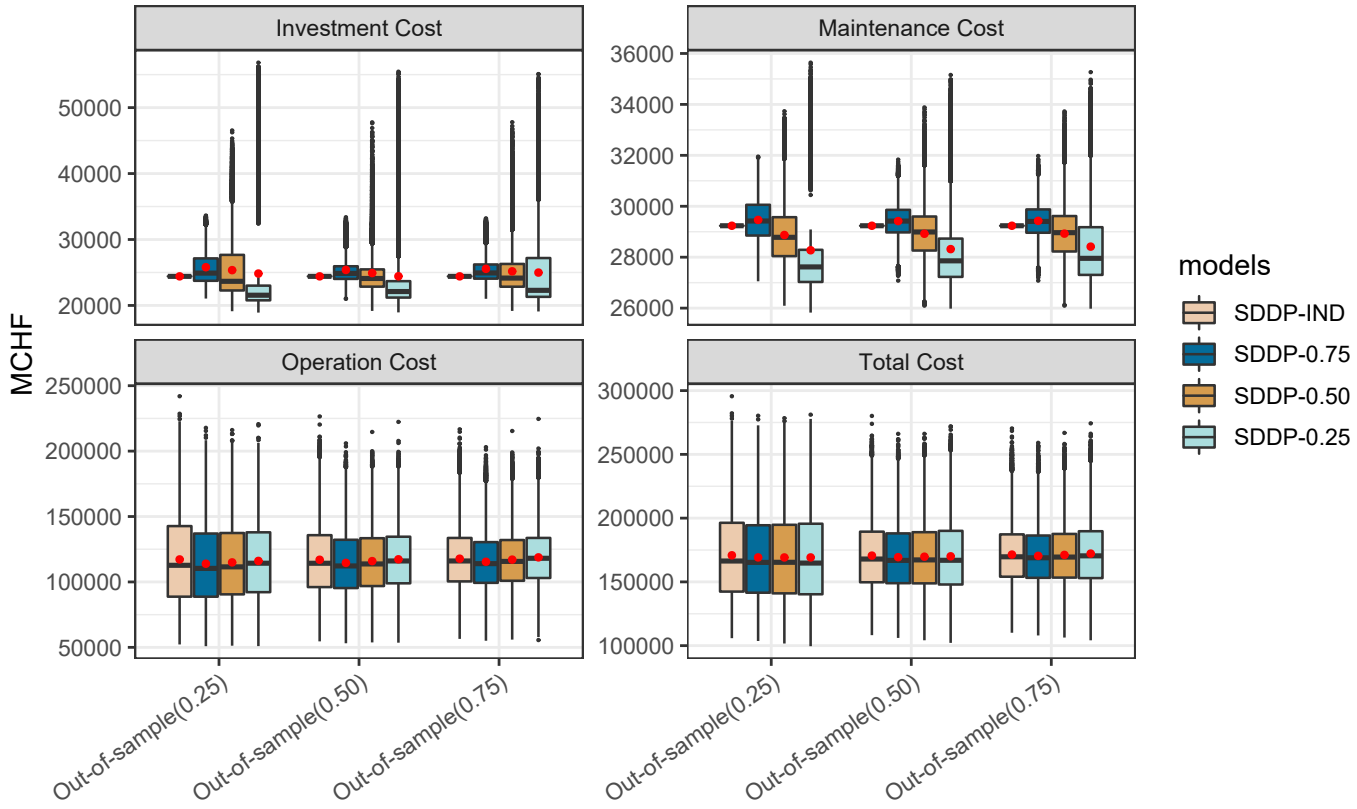


Figure 6: Boxplot of total investment, maintenance, operation and total costs. The red dot represents the sample mean. For each out-of-sample simulation, 30000 scenarios were generated from the original continuous AR process with $\alpha \in \{0.75, 0.50, 0.25\}$.

At first sight, the four stochastic solutions lead to similar performance in terms of total costs although the SDDP-IND policy costs seem to have a higher variability with higher extreme total costs. The decomposition of the total cost into its different components is however completely different. It turns out that the SDDP-IND solution does not adapt investment decisions over time to the realizations of the uncertain parameters. Since the model structure assumes temporal independence, the SDDP-IND model results in sequence a two-stage investment/operations problems in each period that are independent. This leads, as it can be seen on Figure 1 and then for investment decisions on Figures 7 and 9, to a unique investment and maintenance cost and a larger variation in operations costs.

SDDP-α solutions produce much more realistic investment strategies that adapt to the realizations

of uncertainties. We notice that, when α decreases, the investment and maintenance costs have a greater dispersion but a much lower average value. We will see later that the adaptability of the SDDP-α policy allows for investing in efficient technologies according to the degree of correlation with which they have been trained. The smaller the α, the greater the investment diversification, a very desirable result to reduce the risks associated with uncertainty.

5.2.2 Scenario analysis

We now investigate in more detail the role of the α coefficient on its degree of protection against extreme scenarios of parameter realizations. Note that in the sets of 30,000 scenarios generated above with the continuous AR processes, extreme scenarios have a very low probability to appear, i.e., a maximum (or minimum) scenario implies that the sequence of the AR processes takes the maximum (or minimum) value at each stage, which is unlikely. We therefore study specifically the SDDP-α and SDDP-IND policies on the two extreme scenarios (MIN and MAX) together with the nominal scenario and we discuss the resulting costs (Table 2), installed capacities (Figures 7 and 9) and system operations (Figure 8). The MIN (MAX) scenario corresponds to the scenario with the lowest (highest) possible realizations of costs and demands in all stages.

Scenarios	Policies	Inv.	Maint.	Op.	Total
MAX	SDDP-IND	24,412	29,235	341,875	395,521
MAX	SDDP-0.75	32,643	32,059	300,867	365,569
MAX	SDDP-0.50	47,803	34,074	273,114	354,991
MAX	SDDP-0.25	58,124	35,589	259,174	352,887
MIN	SDDP-IND	24,412	29,235	34,563	88,210
MIN	SDDP-0.75	22,534	28,625	34,153	85,312
MIN	SDDP-0.50	22,090	28,093	34,947	85,129
MIN	SDDP-0.25	20,402	26,933	35,810	83,145
Nominal	SDDP-IND	24,412	29,235	112,333	165,979
Nominal	SDDP-0.75	23,210	28,575	113,517	165,302
Nominal	SDDP-0.50	21,023	27,218	116,964	165,204
Nominal	SDDP-0.25	20,079	26,485	118,650	165,214

Table 2: Costs associated to the SDDP-IND and SDDP-α policies for the maximum (MAX), minimum (MIN), nominal (Nominal) scenarios.

First, we remark in Table 2 that the extreme scenarios lead to more extreme total costs than those observed in the previous simulations. This confirms our suspicion that they are not present in the samples. We also observe that the SDDP-α policies perform better than the SDDP-IND one, since the total costs in all scenarios are lower than those of the SDDP-IND, especially on the extreme ones. Indeed, considering the time dependencies of the stochastic process yields a 3%-8% cost reduction on the extreme scenarios. On the one hand, the non-adaptability of the SDDP-IND policy causes the power system to have excess capacity when minimum scenarios are forecast. On the other hand, the power system is not sufficiently protected when maximum scenarios are forecast; indeed, the power system with the SDDP-IND policy has the worst operating cost for the MAX scenario. The operating and investment costs of the SDDP-α policies behave appropriately, i.e., when there is an increase in capacity, the operating costs become lower, but when there is a low capacity, the operating costs become higher. We can notice that the SDDP-0.25 policy performs better than all other policies. It confirms that with a small α the SDDP-α model is trained on a scenario sample that is more likely to include extreme scenarios as it avoids the zigzag effect. Finally, in the nominal scenario, all policies follow a similar behavior as mentioned above and do not show much difference in the actual total cost.

Figure 7 plots installed capacities resulting from the SDDP-IND and SDDP-α policies for the

maximum (MAX), minimum (MIN), nominal (Nominal) scenarios. For the first 20 years, policies invest little in new capacity. Then, when the nuclear plants reach the end of their useful life in 2030, all policies begin to determine which technologies and resources will meet future energy needs.

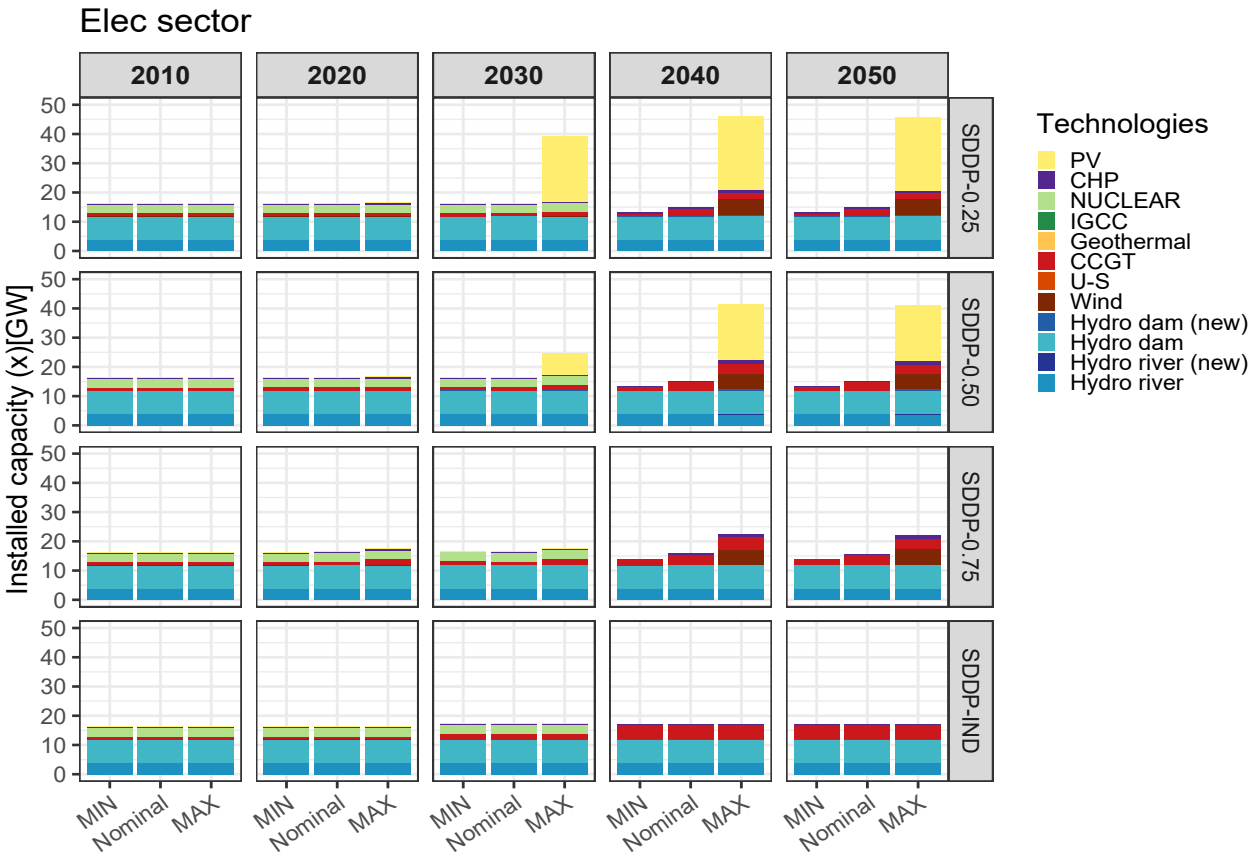


Figure 7: Installed capacities for the SDDP-IND and SDDP-α policies for the maximum (MAX), minimum (MIN), nominal (Nominal) scenarios.

When the demands and resource costs are considered stage-wise independent, the optimal policy solution of SDDP-IND prefers conventional technologies (e.g., based on natural gas) to renewables (e.g., photovoltaic, wind) or cogeneration (CHP). In contrast, the SDDP-α policies will have a more diversified set of technologies by 2050, including PV, Cogeneration, Wind and CCGT at the MAX scenario. In the MIN scenario, the solution will only invest in CCGT and Cogeneration.

Figure 8 shows the power generation by technology in the electricity sector. As observed previously, the SDDP-IND policy depends heavily on natural gas. In the MAX scenario, the optimal SDDP-IND policy uses all available capacity from the most economical technologies and then supplements generation by importing electricity or using existing conventional technologies, regardless of resource costs. On the other hand, the electricity generation resulting from the SDDP-α policies shows a lower dependence on imported electricity. Imports even completely disappear with $\alpha = 0.5$ but then reappear with $\alpha = 0.25$ to satisfy a more important electrification of the system due certainly to a consideration of more extreme scenarios in the optimization.

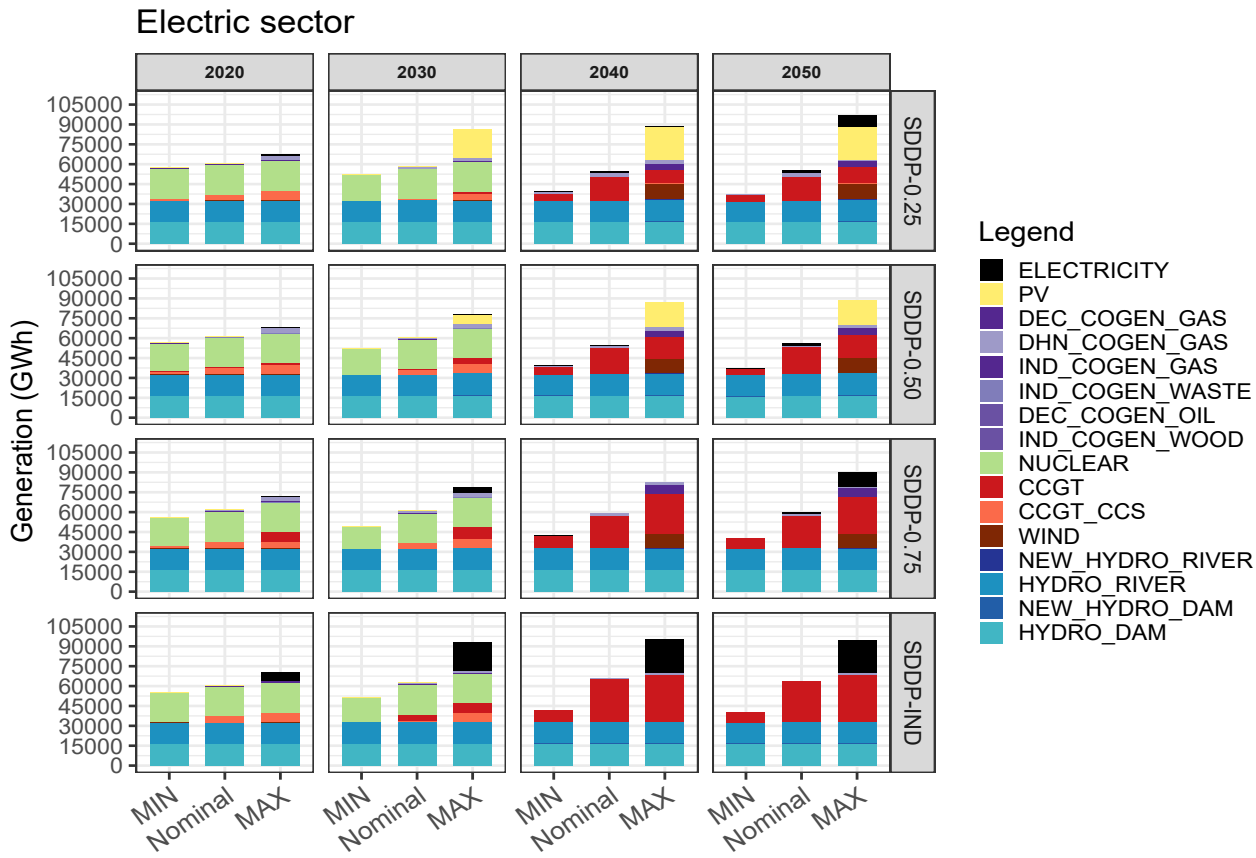


Figure 8: Power generation for the SDDP-IND and SDDP-α policies for the maximum (MAX), minimum (MIN), nominal (Nominal) scenarios.

Finally, Figure 9 shows the evolution of the heating sector. The electrification of the heating sector is confirmed in the MAX scenario as the α decreases. We thus observe a higher penetration of the Electrical Heat Pumps, Direct Elec. and Solar Thermal technologies.

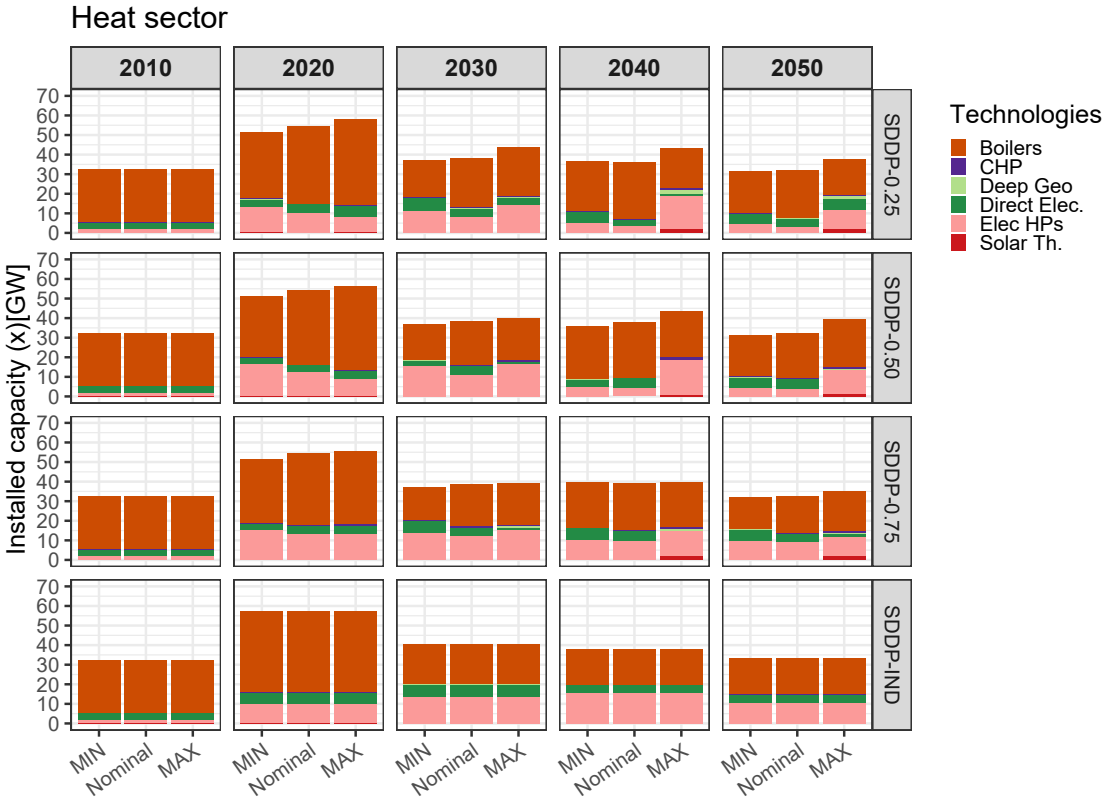


Figure 9: Installed capacity for the heating sector for the SDDP-IND and SDDP-α policies for the maximum (MAX), minimum (MIN), nominal (Nominal) scenarios.

6 Conclusions and Future Work

Strategic energy planning models are undoubtedly important to indicate the investments that must be made in technologies in the current context where there is considerable uncertainty about the evolution of costs and demand for energy. Multi-stage stochastic programming provides a very useful framework in that regard, as it allows for investment decisions to be revised periodically as a function of the uncertainty observed up to that point. The use of multi-stage models in this setting, however, faces two challenges: first, defining the input stochastic processes for the model requires modeling the evolution of costs and demand over time; second, the total number of scenarios grows quickly since strategic energy planning models are characterized by the presence of many sources of uncertainty— both in the objective function and in the constraints—which requires the use of specialized algorithms to solve the problem.

In this paper, we have presented an approach that tackles the two challenges presented above. The key to our methodology is to start from ranges of values for the uncertain components—information that typically can be obtained from external studies in the literature—and then to develop auto-regressive processes that are restricted to those ranges. Such development requires special care, as it differs from standard auto-regressive models with normally distributed noise; indeed, processes with normal noise have unbounded support and hence scenarios with unrealistic values might be generated. In contrast, our approach not only stays within the pre-specified ranges in each period, but also introduces a modeling parameter α that allows the user to control the variability of the process from one period to the next.

We also show that the resulting model can be solved with the well-known Stochastic Dual Dynamic Programming (SDDP) algorithm by using a combination of extra state-variables and Markov chains to represent the AR processes.

We have applied our models to the Swiss energy system. The results show the benefits of using a multi-stage model as opposed to two-stage or robust optimization approaches proposed for this problem in the literature, as the latter models require making all the investment decisions at the beginning of the horizon, whereas our multi-stage model allows for making the decisions over time, thereby adapting them to the observed uncertainty. The results also demonstrate the importance of modeling the dependence of the input processes across stages; in fact, the case where the uncertainty is assumed to be independent across stages does not allow for adaptability—although the investment decisions can change from period to period, they do not adapt to the particular scenarios observed in each period. Finally, the results show the crucial role played by the variability-controlling parameter α. As discussed in the paper, lower values of α allow for protection against extreme scenarios.

The limitations that exist when using time dependence within SDDP are that processes that are only on the right-hand side of the constraints can be represented as affine functions of the errors, and with the Markov chain (which is not limited to considering only the right-hand side as uncertain), it is not possible to consider a large number of states Markov chain. According to the papers presented in the literature review, most of them are formulated as a mixed-integer programming model, since the conversion units (technologies) are integer and not continuous. Therefore for future work, it is to be able to consider the integrality of the technologies and compare how these solutions impact this kind of relaxation. Also, as the Swiss government proposes as a policy to establish zero emissions by 2050, it is necessary to include restrictions in the model to reduce emissions during the planning horizon.

Acknowledgemnts

First and third authors gratefully acknowledge the support provided by FONDECYT 1171145, Chile. The second author gratefully acknowledges partial support from Qatar National Research Fund under Grant Agreement no NPRP10-0212–170447 and from FONDECYT 1190325, Chile. Finally, all three authors acknowledge the support provided by ANILLO ACT192094, Chile.

References

Juan C. Rojas-Zerpa and Jose M. Yusta. Methodologies, technologies and applications for electric supply planning in rural remote areas. *Energy for Sustainable Development*, 20:66 – 76, 2014. ISSN 0973-0826. doi: https://doi.org/10.1016/j.esd.2014.03.003. URL http://www.sciencedirect.com/science/article/pii/S0973082614000271.

Simon Hilpert, Cord Kaldemeyer, Uwe Krien, Stefan Günther, Clemens Wingenbach, and Guido Pleßmann. The open energy modelling framework (oemof) - A new approach to facilitate open science in energy system modelling. *CoRR*, abs/1808.08070, 2018. URL https://github.com/oemof/oemof.

Stefan Pfenninger and Bryn Pickering. Calliope: a multi-scale energy systems modelling framework. *Journal of Open Source Software*, 3:825, 2018. doi: 10.21105/joss.00825. URL https://github.com/calliope-project/calliope.

Gauthier Limpens, Stefano Moret, Hervé Jeanmart, and Francois Maréchal. Energyscope td: A novel open-source model for regional energy systems. *Applied Energy*, 255:113729, 2019. ISSN 0306-2619. doi: https://doi.org/10.1016/j.apenergy.2019.113729. URL http://www.sciencedirect.com/science/article/pii/S0306261919314163.

Joanna Krzemień. Application of markal model generator in optimizing energy systems. *Journal of Sustainable Mining*, 12(2):35–39, 2013. doi: https://doi.org/10.7424/jsm130205.

Mark Howells, Holger Rogner, Neil Strachan, Charles Heaps, Hillard Huntington, Socrates Kypreos, Alison Hughes, Semida Silveira, Joe DeCarolis, Morgan Bazillian, and Alexander Roehrl. Osemosys: The open source energy modeling system: An introduction to its ethos, structure and development. *Energy Policy*, 39(10):5850–5870, 2011. doi: https://doi.org/10.1016/j.enpol.2011.06.033.

Frédéric Babonneau, Michael Caramanis, and Alain Haurie. ETEM-SG: Optimizing regional smart energy system with power distribution constraints and options. *Environmental Modeling & Assessment*, 22 (5):411–430, 2017. doi: 10.1007/s10666-016-9544-0.

Patrick Sullivan, Volker Krey, and Keywan Riahi. Impacts of considering electric sector variability and reliability in the message model. *Energy Strategy Reviews*, 1(3):157–163, 2013. doi: https://doi.org/10.1016/j.esr.2013.01.001.

Warren B. Powell, Abraham George, Hugo Simão, Warren Scott, Alan Lamont, and Jeffrey Stewart. Smart: A stochastic multiscale model for the analysis of energy resources, technology, and policy. *INFORMS Journal on Computing*, 24(4):665–682, 2012. doi: 10.1287/ijoc.1110.0470.

Stefano Moret, Frédéric Babonneau, Michel Bierlaire, and François Maréchal. Decision support for strategic energy planning: A robust optimization framework. *European Journal of Operational Research*, 280(2):539 – 554, 2020. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2019.06.015. URL http://www.sciencedirect.com/science/article/pii/S0377221719304904.

Esnil Guevara, Fréderic Babonneau, Tito Homem-de-Mello, and Stefano Moret. A machine learning and distributionally robust optimization framework for strategic energy planning under uncertainty. *Applied Energy*, 271:115005, 2020. ISSN 0306-2619. doi: https://doi.org/10.1016/j.apenergy.2020.115005. URL http://www.sciencedirect.com/science/article/pii/S0306261920305171.

Stefano Moret, Michel Bierlaire, and François Maréchal. Strategic energy planning under uncertainty: a mixed-integer linear programming modeling framework for large-scale energy systems. In Zdravko Kravanja and Miloš Bogataj, editors, *26th European Symposium on Computer Aided Process Engineering*, volume 38 of *Computer Aided Chemical Engineering*, pages 1899–1904. Elsevier, 2016. doi: https://doi.org/10.1016/B978-0-444-63428-3.50321-0.

Stefano Moret, Víctor Codina Gironès, Michel Bierlaire, and François Maréchal. Characterization of input uncertainties in strategic energy planning models. *Applied Energy*, 202:597 – 617, 2017. ISSN 0306-2619. doi: https://doi.org/10.1016/j.apenergy.2017.05.106. URL http://www.sciencedirect.com/science/article/pii/S0306261917306116.

Wolfram Wiesemann, Daniel Kuhn, and Melvyn Sim. Distributionally robust convex optimization. *Operations Research*, 62(6):1358–1376, 2014. doi: 10.1287/opre.2014.1314.

John R. Birge and François Louveaux. *Multistage Stochastic Programs*, pages 265–287. Springer New York, New York, NY, 2011. ISBN 978-1-4614-0237-4. doi: 10.1007/978-1-4614-0237-4_6. URL https://doi.org/10.1007/978-1-4614-0237-4_6.

G.E.P. Box, G.M. Jenkins, G.C. Reinsel, and G.M. Ljung. *Time Series Analysis: Forecasting and Control*. Wiley Series in Probability and Statistics. Wiley, 2015. ISBN 9781118674925. URL https://books.google.cl/books?id=rNt5CgAAQBAJ.

Shan Jin, Sarah Ryan, Jean-Paul Watson, and David Woodruff. Modeling and solving a large-scale generation expansion planning problem under uncertainty. *Energy Systems*, 2:209–242, 11 2011. doi: 10.1007/s12667-011-0042-9.

Rahul Marathe and Sarah Ryan. On the validity of the geometric brownian motion assumption. *The Engineering Economist*, 50, 04 2005. doi: 10.1080/00137910590949904.

Holger Heitsch and Werner Roemisch. Scenario tree modeling for multistage stochastic programs. *Math. Program.*, 118:371–406, 05 2009. doi: 10.1007/s10107-007-0197-2.

George Tauchen. Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters*, 20(2):177 – 181, 1986. ISSN 0165-1765. doi: https://doi.org/10.1016/0165-1765(86)90168-0. URL http://www.sciencedirect.com/science/article/pii/0165176586901680.

J. Singh, Kavinesh, B. Philpott, Andy, and Kevin Wood. Dantzig-wolfe decomposition for solving multistage stochastic capacity-planning problems, 2009. URL https://calhoun.nps.edu/handle/10945/38417.

Heejung Park and Ross Baldick. Multi-year stochastic generation capacity expansion planning under environmental energy policy. *Applied Energy*, 183:737 – 745, 2016. ISSN 0306-2619. doi: https://doi.org/10.1016/j.apenergy.2016.08.164. URL http://www.sciencedirect.com/science/article/pii/S0306261916312739.

Anastasia Ioannou, Gulistiani Fuzuli, Feargal Brennan, Satya Widya Yudha, and Andrew Angus. Multi-stage stochastic optimization framework for power generation system planning integrating hybrid uncertainty modelling. *Energy Economics*, 80:760 – 776, 2019. ISSN 0140-9883. doi: https://doi.org/10.1016/j.eneco.2019.02.013. URL http://www.sciencedirect.com/science/article/pii/S0140988319300702.

Y. Liu, R. Sioshansi, and A. J. Conejo. Multistage stochastic investment planning with multiscale representation of uncertainties and decisions. *IEEE Transactions on Power Systems*, 33(1):781–791, 2018.

Alexander Shapiro and Arkadi Nemirovski. *On Complexity of Stochastic Programming Problems*, pages 111–146. Springer US, Boston, MA, 2005. ISBN 978-0-387-26771-5. doi: 10.1007/0-387-26771-9_4. URL https://doi.org/10.1007/0-387-26771-9_4.

M. V. Pereira and L. M. Pinto. Multi-stage stochastic optimization applied to energy planning. *Math. Program.*, 52(1–3):359–375, May 1991. ISSN 0025-5610.

Cristiana L. Lara, Dharik S. Mallapragada, Dimitri J. Papageorgiou, Aranya Venkatesh, and Ignacio E. Grossmann. Deterministic electric power infrastructure planning: Mixed-integer programming model and nested decomposition algorithm. *European Journal of Operational Research*, 271(3):1037 – 1054, 2018. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2018.05.039. URL http://www.sciencedirect.com/science/article/pii/S0377221718304466.

Cristiana L. Lara, John D. Siirola, and I. Grossmann. Electric power infrastructure planning under uncertainty: stochastic dual dynamic integer programming (sddip) and parallelization scheme. *Optimization and Engineering*, pages 1–39, 2019.

Steffen Rebennack. Generation expansion planning under uncertainty with emissions quotas. *Electric Power Systems Research*, 114:78–85, 2014.

Jikai Zou, Shabbir Ahmed, and Xu Andy Sun. Stochastic dual dynamic integer programming. *Math. Program.*, 175(1–2):461–502, May 2019. ISSN 0025-5610. doi: 10.1007/s10107-018-1249-5. URL https://doi.org/10.1007/s10107-018-1249-5.

Fernanda Thomé, Ricardo C. Perez, Lucas Okamura, Alessandro Soares, and Silvio Binato. Stochastic multistage co-optimization of generation and transmission expansion planning, 2019.

J. R. Birge. Decomposition and partitioning methods for multistage stochastic linear programs. *Oper Res*, 33(5):989–1007, 1985.

N. Campodónico, Silvio Binato, Rafael Kelman, Mario Pereira, Martin Tinoco, F. Montoya, M. Zhang, and F. Mayaki. Expansion planning of generation and interconnections under uncertainty. *Proc. Ann. 3rd Balkans Power Conf., 2003.*, 2003.

Alexander Shapiro, Wajdi Tekaya, Joari Paulo da Costa, and Murilo Pereira Soares. Risk neutral and risk averse stochastic dual dynamic programming method. *European Journal of Operational Research*, 224(2):375 – 391, 2013. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2012.08.022. URL http://www.sciencedirect.com/science/article/pii/S0377221712006455.

Nils Löhndorf and Alexander Shapiro. Modeling time-dependent randomness in stochastic dual dynamic programming. *European Journal of Operational Research*, 273(2):650 – 661, 2019. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2018.08.001. URL http://www.sciencedirect.com/science/article/pii/S0377221718306787.

A.B. Philpott and V.L. de Matos. Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion. *European Journal of Operational Research*, 218(2):470 – 483, 2012. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2011.10.056. URL http://www.sciencedirect.com/science/article/pii/S0377221711010332.

George Tauchen and Robert Hussey. Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica*, 59(2):371–396, 1991. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/2938261.

Karen A. Kopecky and Richard M.H. Suen. Finite state markov-chain approximations to highly persistent processes. *Review of Economic Dynamics*, 13(3):701 – 714, 2010. ISSN 1094-2025. doi: https://doi.org/10.1016/j.red.2010.02.002. URL http://www.sciencedirect.com/science/article/pii/S1094202510000128.

K. Rouwenhorst. Asset pricing implications of equilibrium business cycle models. In Thomas F. Cooley, editor, *In Frontiers of Business Cycle Research*, page 294–330, 1995. Princeton University Press.

Giulio Fella, Giovanni Gallipoli, and Jutong Pan. Markov-chain approximations for life-cycle models. *Review of Economic Dynamics*, 34:183 – 201, 2019. ISSN 1094-2025. doi: https://doi.org/10.1016/j.red.2019.03.013. URL http://www.sciencedirect.com/science/article/pii/S1094202519301565.

O. Dowson and L. Kapelevich. SDDP.jl: a Julia package for stochastic dual dynamic programming. *INFORMS Journal on Computing*, 2020. doi: https://doi.org/10.1287/ijoc.2020.0987. Articles in Advance.

M. D. McKay, R. J. Beckman, and W. J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979. ISSN 00401706. URL http://www.jstor.org/stable/1268522.

Tito Homem-de Mello, Vitor de Matos, and Erlon Finardi. Sampling strategies and stopping criteria for stochastic dual dynamic programming: A case study in long-term hydrothermal scheduling. *Energy Systems*, 2:1–31, 03 2011. doi: 10.1007/s12667-011-0024-y.

Vitor L. de Matos, David Morton, and Erlon C. Finardi. Assessing policy quality in a multistage stochastic program for long-term hydrothermal scheduling. *Annals of Operations Research*, 253(2):713–731, June 2017. ISSN 0254-5330. doi: 10.1007/s10479-016-2107-6.

L. Ding, S. Ahmed, and A. Shapiro. A python package for multi-stage stochastic programming. In *Optimization online*, pages 1–41, 2019.

A Appendix

A.1 Model formulation

In this subsection, we define the sets, the parameters and the variables that are used in the multi-stage strategic energy planning model. Figure 10 summarizes the sets with their relative indices used throughout the paper.

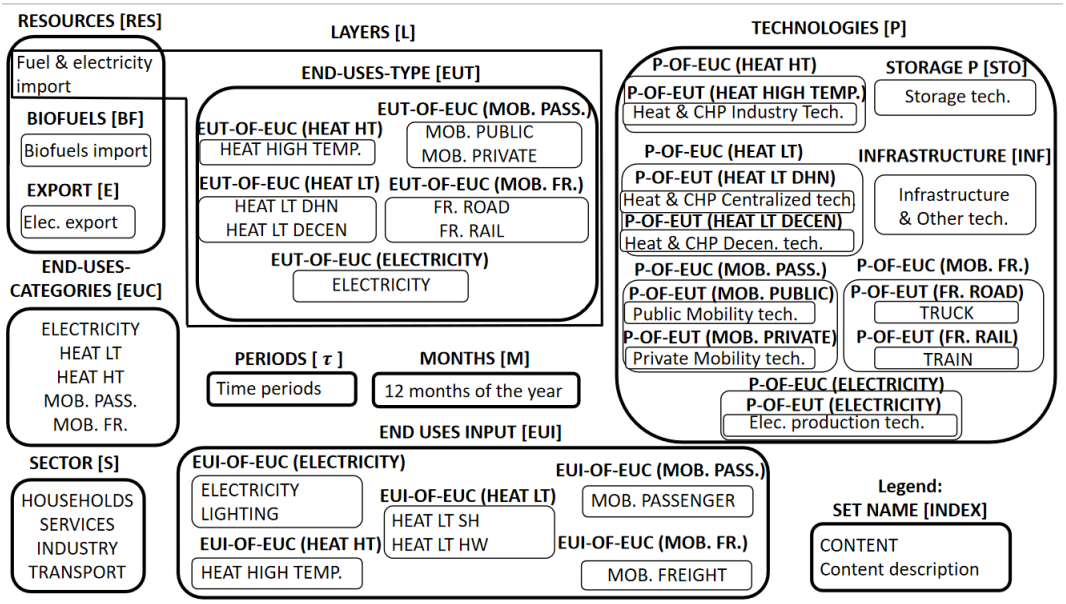


Figure 10: Diagram of sets and indices used in the MSLP model (source: Moret et al. [2017]). Abbreviations: space heating (SH), hot water (HW), temperature (Temp.), mobility (MOB), Low Temperature (LT), High Temperature (HT), Freight (FR.), Passenger (PASS.).

In Figure 10, we represent the elements defining the set by a square with rounded edges. For example, the set **EUC** is one of the main sets of the model and the elements of this set will be used as indexes for new subsets. The set **EUT** is the union of the subsets **EUT-OF-EUC** indexed by elements of the set **EUC**. In addition, the end-use demands are calculated based on these elements. The set **EUI** is the union of the subsets **EUI-OF-EUC** indexed by the elements of the set **EUC**. The set **P** (or processes) also contains subsets indexed by elements of the **EUC** set, but in turn, the subsets formed by the elements HEAT LT, MOB. PASS. and MOB. FR. contain two subsets indexed by their corresponding $eut \in \mathbf{EUT}$ element.

The set **L** is a special set, since on these elements the model will have to balance what is generated, produced, consumed or stored in each month. Also each $l \in \mathbf{L}$ element defines a type of end-use demand that must be satisfied. As an example, in the case of decentralized heat production with a NG boiler, the amount of NG consumed by the boiler is an input layer, and the amount of heat produced by the boiler is an output layer. In the case of electricity, the model can import or produce electricity within the system to meet the electricity end use demand; in addition, electricity can be stored in hydroelectric dams or used as an input to other energy conversion technologies (such as electric vehicles) to meet other types of end-use demand.

(I) Sets (and Indices).

$\mathcal{T}$	time periods, $t \in \mathcal{T} = \{1, ..., T\}$
M	twelve-month period, $m \in \mathbf{M} = \{1, ..., 12\}$
RES	imported or local resources, $r \in \mathbf{RES} = \{coal, gas\ natural\ (NG), electricity, wood, ...\}$
INF	infrastructure: DHN, grid, and intermediate energy conversion technologies (i.e. not directly supplying end-use demand), $inf \in \mathbf{INF} = \{electrolysis, pyrolysis, ...\}$
E	exported electricity ($\subset$ **RES**), $ex \in \mathbf{E} = \{electricity\ export\}$
S	sectors of the energy system, $s \in \mathbf{S} = \{households, services, industry, transport\}$
STO	storage units, $u \in \mathbf{STO} = \{Power\text{-}to\text{-}Gas, Pumped\ Hydro\}$
EUC	categories of demand (end-uses), $euc \in \mathbf{EUC} = \{electricity, heat\ low\ temperature, ...\}$
EUI	input to the model, $eui \in \mathbf{EUI} = \{electricity, lighting, heat\ lt\ sh, heat\ lt\ hw, ...\}$ $\mathbf{EUI} = \bigcup_{euc \in \mathbf{EUC}} \boldsymbol{EUI\text{-}OF\text{-}EUC}_{euc}$
EUT	types of demand (end-uses), $eut \in \mathbf{EUT} = \{electricity, heat\ high\ temperature, ...\}$ $\mathbf{EUT} = \bigcup_{euc \in \mathbf{EUC}} \boldsymbol{EUT\text{-}OF\text{-}EUC}_{euc}$
L	layers are used to balance resources/processes in the system, $l \in \mathbf{L} = \mathbf{RES} \setminus \{\mathbf{BF} \cup \mathbf{E}\} \cup \mathbf{EUT}$
$\mathbf{P\text{-}OF\text{-}EUT}_{ELECTRICITY}$	energy conversion technologies for electricity generation with $p \in \{nuclear, ccgt, pv, wind, hydro\ dam, hydro\ river, coal, ...\}$
$\mathbf{P\text{-}OF\text{-}EUT}_{\text{HEAT HIGH TEMP.}}$	energy conversion technologies for heat production in industrial processes at high temperatures with $p \in \{boiler, CHP, elec.\ direct\ heating\}$
$\mathbf{P\text{-}OF\text{-}EUT}_{\text{HEAT LT DHN}}$	energy conversion technologies for centralized heat production (DHN) with $p \in \{boiler, CHP, heat\ pump, deep\ geothermal\}$
$\mathbf{P\text{-}OF\text{-}EUT}_{\text{HEAT LT DECEN}}$	energy conversion technologies for decentralized heat production with $p \in \{boiler, CHP, fuel\ cell, solar\ thermal, heat\ pump, elec.\ direct\ heating\}$
$\mathbf{P\text{-}OF\text{-}EUT}_{\text{MOB. PUBLIC}}$	energy conversion technologies for passenger mobility in the public sector with $p \in \{gasolina\ car, diesel\ car, NG\ car, electric\ car, ...\}$
$\mathbf{P\text{-}OF\text{-}EUT}_{\text{MOB. PRIVATE}}$	energy conversion technologies for passenger mobility in the private sector with $p \in \{diesel\ bus, hybrid\ bus, trolley\ bus, train/metro, ...\}$
$\mathbf{P\text{-}OF\text{-}EUT}_{\text{FR. ROAD}}$	energy conversion technologies for road freight transport with $p \in \{truck\}$
$\mathbf{P\text{-}OF\text{-}EUT}_{\text{FR. RAIL}}$	energy conversion technologies for rail freight transport with $p \in \{trainfreight\}$
$\mathbf{P\text{-}OF\text{-}EUC}_{ELECTRICITY}$	energy conversion technologies in the electricity category with $p \in \{k : k \in \mathbf{P\text{-}OF\text{-}EUT}_{ELECTRICITY}\}$
$\mathbf{P\text{-}OF\text{-}EUC}_{\text{HEAT HT}}$	energy conversion technologies in the heat high temp. category with $p \in \{k : k \in \mathbf{P\text{-}OF\text{-}EUT}_{\text{HEAT HIGH TEMP.}}\}$
$\mathbf{P\text{-}OF\text{-}EUC}_{\text{HEAT LT}}$	energy conversion technologies in the heat low temp. category with $p \in \{k : k \in \mathbf{P\text{-}OF\text{-}EUT}_{\text{HEAT LT DHN}} \cup \mathbf{P\text{-}OF\text{-}EUT}_{\text{HEAT LT DECEN}}\}$
$\mathbf{P\text{-}OF\text{-}EUC}_{\text{MOB. PASS.}}$	energy conversion technologies in the mobility passenger category with $p \in \{k : k \in \mathbf{P\text{-}OF\text{-}EUT}_{\text{MOB. PUBLIC}} \cup \mathbf{P\text{-}OF\text{-}EUT}_{\text{MOB. PRIVATE}}\}$
$\mathbf{P\text{-}OF\text{-}EUC}_{\text{MOB. FR.}}$	energy conversion technologies in the mobility freight category with $p \in \{k : k \in \mathbf{P\text{-}OF\text{-}EUT}_{\text{MOB. ROAD}} \cup \mathbf{P\text{-}OF\text{-}EUT}_{\text{MOB. RAIL}}\}$
P	processes (or technologies) with $p \in \left\{k : k \in \left(\bigcup_{euc \in EUC} \mathbf{P\text{-}OF\text{-}EUC}_{euc}\right) \cup \mathbf{STO} \cup \mathbf{INF}\right\}$
MOB	energy conversion technologies for passenger and freight transportation ($\subset$ **P**), $p \in \mathbf{MOB} = \mathbf{P\text{-}OF\text{-}EUC}_{MOB.PASS.} \cup \mathbf{P\text{-}OF\text{-}EUC}_{MOB.FR.}$

(II) Parameters

[2] [Mpkm] (millions of passenger-km) for passenger, [Mtkm] (millions of ton-km) for freight mobility end-uses

[3] [Mpkm/h] for passenger, [Mtkm/h] for freight mobility end-uses

Name	Description	Units
	Parameters used in investments	
T	number of planning periods	
a_l	$a_l = 1$ if $l =$ **HEAT LT DHN**, $a_l = 0$ otherwise. $l \in \mathbf{L}$	
Δ_T	number of years in each period	
i_{rate}	discount rate	
$\overline{g}_{\%_{public}}$, $\underline{g}_{\%_{public}}$	Upper and lower limit to $\%_{public}$ (mobility public)	
$\overline{g}_{\%_{rail}}$, $\underline{g}_{\%_{rail}}$	Upper and lower limit to $\%_{rail}$ (mobility rail)	
$\overline{g}_{\%_{dhn}}$, $\underline{g}_{\%_{dhn}}$	Upper and lower limit to $\%_{dhn}$ (heat low temp. DHN)	
$c^{Inv}_{t,p}$	investment cost of technology $p \in \mathbf{P}$ in the time period $t \in \mathcal{T}$	[MCHF/GW][2,3]
lt_p	lifetime of technology $p \in \mathbf{P}$	
ς_t	annualization factor in time period $t \in \mathcal{T}$	
$res_{t,p}$	residual capacity of technology $p \in \mathbf{P}$ in time period $t \in \mathcal{T}$	[GW]
$\bar{f}^{min}_p$, $\bar{f}^{max}_p$	minimum/maximum available capacity of the technology installed $p \in \mathbf{P}$	[GW]
	Parameters used for system operations	
$f^{min}_{t,p}$, $f^{max}_{t,p}$	minimum/maximum capacity of installed technology $p \in \mathbf{P}$ in time period $t \in \mathcal{T}$	[GW][2,3]
$f^{min,\%}_p$, $f^{max,\%}_p$	it expresses the minimum and maximum yearly output shares of each technology $p \in P$ for each type of end uses demand	
$avail_r$	available amount of the resource $r \in$ **RES**	[GWh]
$k_{p,m}$	capacity factor of technology $p \in \mathbf{P}$ in month $m \in \mathbf{M}$	
$\widehat{k}_p$	yearly capacity factor of technology $p \in \mathbf{P}$	
$c^{Op}_{t,r,m}$	cost of the resource $r \in$ **RES** in month $m \in \mathbf{M}$ and time period $t \in \mathcal{T}$	[MCHF/GWh]
gwp^{Op}_r	emissions associated to fuels (from cradle to combustion) and imports of electricity. $r \in$ **RES**	$[ktCO_2\text{-}eq./GWh]$
$\eta^+_{u,l}$, $\eta^-_{u,l}$	efficiency [0;1] of storage $u \in$ **STO** input from/output to layer $l \in \mathbf{L}$	
$\%loss_{eut}$	electrical and thermal energy loss factor [0,1] (GRID and DHN). $eut \in \{\text{ELECTRICITY, HEAT LT DHN}\}$	
$\%Peak_{DHN}$	factor used to avoid underestimating the cost of centralized heat production	
$eUYear_{t,eui,s}$	annual end-uses in energy services $eui \in$ **EUI** per sector $s \in \mathbf{S}$ in the time period $t \in \mathcal{T}$	[GWh][1]
$eUI_{t,eui}$	total annual end-uses in energy services $eui \in$ **EUI** in the time period $t \in \mathcal{T}$ $eUI_{t,eui} = \sum_{s \in S} eUYear_{t,eui,s} \quad \forall eui \in \mathbf{EUI}, t \in \mathcal{T}$	[GWh][2]
h_m	duration of the month $m \in \mathbf{M}$ in hours	[h]
$\%lighting_m$	monthly distribution factor for lighting end use input (adding up to 1, $\sum_{m \in M} \%lighting_m = 1$)	
$\%sh_m$	monthly distribution factor for space heating end use input (adding up to 1, $\sum_{m \in M} \%sh_m = 1$)	
$f_{i,l}$	input/output resources/technologies to layers. $i \in \mathbf{RES} \cup \mathbf{P} \setminus \mathbf{STO}, l \in \mathbf{L}$	[GW][3]
$c^{Maint}_{t,p}$	technology maintenance cost $p \in \mathbf{P}$ in time period $t \in \mathcal{T}$	[MCHF/GW][2,3]

(III) Variables

Name	Description	Units
	variables used for investments	
$\mathbf{x}_{t,p}$	total available capacity of technology $p \in \mathbf{P}$ in time period $t \in \mathcal{T}$	[GW][4]
$\mathbf{y}_{t,p}$	new installed capacity of technology $p \in \mathbf{P}$ in time period $t \in \mathcal{T}$	[GW][3]
$\mathbf{G}_{t,\%_{public}}$	ratio [0; 1] public mobility over total passenger mobility in time period $t \in \mathcal{T}$	
$\mathbf{G}_{t,\%_{rail}}$	ratio [0; 1] rail transport over total freight transport in time period $t \in \mathcal{T}$	
$\mathbf{G}_{t,\%_{dhn}}$	ratio [0; 1] centralized over total low-temperature heat in time period $t \in \mathcal{T}$	
$\mathbf{C}^{Inv}_t$	discounted total investment cost in time period $t \in \mathcal{T}$	[MCHF]
$\mathbf{C}^{SV}_t$	discounted salvage value in time period $t \in \mathcal{T}$	[MCHF]
	Variables used for system operations	
$\mathbf{z}_{t,p,m}$	the operation of resources and technologies in each month $m \in \mathbf{M}$ and time period $t \in \mathcal{T}$	[GW][2,3]
$\mathbf{Sto}^+_{t,u,l,m}$	input to storage units $u \in$ **STO**, $l \in \mathbf{L}$, $m \in \mathbf{M}$, $t \in \mathcal{T}$	[GW]
$\mathbf{Sto}^-_{t,u,l,m}$	output from storage units $u \in$ **STO**, $l \in \mathbf{L}$, $m \in \mathbf{M}$, $t \in \mathcal{T}$	[GW]
$\mathbf{D}_{t,l,m}$	end-use demand $l \in \mathbf{L}, m \in \mathbf{M}, t \in \mathcal{T}$. Set to 0 if $l \notin$ **EUT**	[GW]
$\mathbf{GWP}^{op}_{t,r}$	GHG emissions of resources $r \in$ **RES** in time period $t \in \mathcal{T}$	$[ktCO_2\text{-}eq.]$
$\mathbf{Loss}_{t,eut,m}$	losses in the networks (grid and DHN), $t \in \mathcal{T}, eut \in$ **EUT**, $m \in \mathbf{M}$	[GW]
$\mathbf{C}^{O\&M}_t$	discounted $O\&M$ cost in time period $t \in \mathcal{T}$	[MCHF]

[4] [GWh] if $p \in$ **STO**

A.2 Objective function

The objective function of the energy problem minimizes the total discounted system cost (i.e., investment and operation costs) minus the salvage value of the residual life of installed technologies at the end of the planning horizon.

$$\min \sum_{t=1}^{T} \frac{\mathbf{C}_t^{Inv} - \mathbf{C}_t^{SV} + \mathbf{C}_t^{O\&M}}{(1+i_{rate})^{(t-1)\cdot\Delta_T}} \tag{m.1}$$

The investment cost $\mathbf{C}_t^{Inv}$ on period t is given by

$$\mathbf{C}_t^{Inv} = \sum_{p\in\mathbf{P}} c_{t,p}^{Inv} \cdot \mathbf{y}_{t,p}. \tag{m.2}$$

The O&M cost $\mathbf{C}_t^{O\&M}$ on period t is defined as follows

$$\mathbf{C}_t^{O\&M} = \varsigma_t \cdot \sum_{p\in\mathbf{P}} c_{t,p}^{Maint} \cdot \mathbf{x}_{t,p} + \varsigma_t \cdot \sum_{r\in\mathbf{RES}} \sum_{m\in\mathbf{M}} c_{t,r,m}^{Op} \cdot \mathbf{z}_{t,r,m} \cdot h_m^{Op}, \tag{m.3}$$

where parameter ς_t is an annualization factor. Maintenance and operation costs are linked to the available capacities and resource activities, respectively. Finally the salvage value $\mathbf{C}_t^{SV}$ is defined as

$$\mathbf{C}_t^{SV} = \sum_{p\in\mathbf{P}} sv_{t,p} \cdot c_{t,p}^{Inv} \cdot \mathbf{y}_{t,p} \cdot \mathbb{1}_{\{t+lt_p>T+1\}}(t), \tag{m.4}$$

with

$$sv_{t,p} = \frac{1-(1+i_{rate})^{-\Delta_T\cdot(t+lt_p-T-1)}}{\frac{(1+i_{rate})^{\Delta_T\cdot(T+1-t)}}{\left(1-(1+i_{rate})^{-\Delta_T\cdot lt_p}\right)}}.$$

A.3 Constraints

The first set of restrictions refers to the initial design of the capacity and its subsequent expansion, which considers the new installed capacity and the decommissioning plan. The available capacity $\mathbf{x}_{t,p}$ of technology p in time period t, as shown in (m.6), is continuously updated by balancing the accumulated capacity $\mathbf{x}_{t-1,p}$ in time period $t-1$, the new installed capacity $\mathbf{y}_{t,p}$, the elimination of unused capacity $res_{t,p} - res_{t-1,p}$ at the beginning of the time period t and the retirement capacity of technologies that were installed for optimization but have reached their useful life.

$$\mathbf{x}_{0,p} = res_{0,p}, \quad \forall p \in \mathbf{P} \tag{m.5}$$

$$\mathbf{x}_{t,p} = \begin{cases} \mathbf{x}_{t-1,p} + \mathbf{y}_{t,p} - res_{t-1,p} + res_{t,p} - \mathbf{y}_{t-lt_p,p} & if \quad t > lt_p \quad \forall p \in \mathbf{P}, t \in \mathcal{T} \\ \mathbf{x}_{t-1,p} + \mathbf{y}_{t,p} - res_{t-1,p} + res_{t,p} & if \quad 0 < t \leq lt_p \quad \forall p \in \mathbf{P}, t \in \mathcal{T} \end{cases} \tag{m.6}$$

At $t = 0$, the initial capacity is the existing infrastructure at that moment as show in (m.5). Once a technology is installed, it can be operated until its end-of-life. For the available capacity we use the upper and lower limits for the entire planning horizon (m.7), as follows

$$\bar{f}_p^{min} \leq \mathbf{x}_{t,p} \leq \bar{f}_p^{max} \quad \forall p \in \mathbf{P}, t \in \mathcal{T} \tag{m.7}$$

Long-term investment equations

Constraint (m.8) set the upper and lower bounds to the new installed capacity for each time period $t \in \mathcal{T}$ and technology $p \in \mathbf{P}$.

$$f^{min}_{t,p} \leq \mathbf{y}_{t,p} \leq f^{max}_{t,p} \quad \forall p \in \mathbf{P}, t \in \mathcal{T} \tag{m.8}$$

Constraint (m.9) forces the new installed capacity at each time period t to equal total annual end-use for passenger and freight mobility. The addition of this constraint is motivated by the fact that the investment cost of passenger and freight transport technologies is not accounted for in the model. This helps to avoid the installation of more capacity than necessary in the transportation sector.

$$\sum_{mob \in \mathbf{P\text{-}OF\text{-}EUC}\{euc\}} \mathbf{y}_{t,mob} = \frac{eUI_{t,euc}}{\sum_{m \in \mathbf{M}} h_m} \quad \forall euc \in \{\mathbf{MOB.\ PASS.}, \mathbf{MOB.\ FR.}\}, t \in \mathcal{T} \tag{m.9}$$

Constraints (m.10)-(m.12) set the upper and lower limits to the share public vs private mobility, train vs truck in freight transportation and DHN vs decentralized for low-temperature heating demand. Note that, although these variables are not investment decisions, they are related to them, since these variables divide mobility (passenger and freight) and heat low demands, which can affect investment decisions.

$$\underline{g}_{\%_{public}} \leq \mathbf{G}_{t,\%_{public}} \leq \overline{g}_{\%_{public}} \quad \forall t \in \mathcal{T} \tag{m.10}$$

$$\underline{g}_{\%_{rail}} \leq \mathbf{G}_{t,\%_{rail}} \leq \overline{g}_{\%_{rail}} \quad \forall t \in \mathcal{T} \tag{m.11}$$

$$\underline{g}_{\%_{dhn}} \leq \mathbf{G}_{t,\%_{dhn}} \leq \overline{g}_{\%_{dhn}} \quad \forall t \in \mathcal{T} \tag{m.12}$$

There are a number of other constraints related to investment in new capacity, but we present them later because they are constraints specific to the Swiss energy problem.

Short-term operating equations

The operation of resources and technologies are determined by the operating variable $\mathbf{z}$. Constraints (m.13)-(m.14) link the total capacity available of a technology to its actual use in each month and time period via two capacity factors: a capacity factor for each month ($k_{p,m}$) depending on resource availability (e.g. renewable) and a yearly capacity factor ($\widehat{k}_p$) accounting for technology downtime and maintenance.

$$\mathbf{z}_{t,p,m} \leq \mathbf{x}_{t,p} \cdot k_{p,m} \quad \forall p \in \mathbf{P}, \forall m \in \mathbf{M}, t \in \mathcal{T} \tag{m.13}$$

$$\sum_{m \in \mathbf{M}} \mathbf{z}_{t,p,m} \cdot h_m \leq \mathbf{x}_{t,p} \cdot \widehat{k}_p \sum_{m \in \mathbf{M}} h_m \quad \forall p \in \mathbf{P}, t \in \mathcal{T} \tag{m.14}$$

In the constraint (m.15), the total use of resources is limited by the yearly availability ($avail$).

$$\sum_{m \in \mathbf{M}} \mathbf{z}_{t,r,m} \cdot h_m \leq avail_r \quad \forall r \in \mathbf{RES}, t \in \mathcal{T} \tag{m.15}$$

Constraint (m.16) is the balancing equation for storage units. In this constraint when $m = 1$ then $\mathbf{z}_{t,u,0} = \mathbf{z}_{t,u,12}$.

$$\mathbf{z}_{t,u,m} = \mathbf{z}_{t,u,m-1} + h_m \sum_{\substack{l \in \mathbf{L} \\ \eta^+_{u,l} > 0}} \mathbf{Sto}^+_{t,u,l,m} \eta^+_{u,l} - h_m \sum_{\substack{l \in \mathbf{L} \\ \eta^-_{u,l} > 0}} \mathbf{Sto}^-_{t,u,l,m} / \eta^-_{u,l} \quad \forall u \in \mathbf{STO}, \forall m \in \mathbf{M}, t \in \mathcal{T} \tag{m.16}$$

Constraint (m.17)-(m.25) shows the calculation of the end-uses demand $\mathbf{D}$ in each month m and layer l, as a function of end-use input (eUI).

$$\mathbf{D}_{t,ELEC.,m} = \frac{eUI_{t,ELECTRICITY}}{\sum_{m' \in \mathbf{M}} h_{m'}} + eUI_{t,lighting} \frac{\%lighting_m}{h_m} + \mathbf{Loss}_{t,ELEC.,m} \quad \forall m \in \mathbf{M}, t \in \mathcal{T} \tag{m.17}$$

$$\mathbf{D}_{t,\text{HEAT LT DHN},m} = \left(\frac{eUI_{t,\text{HEAT LT HW}}}{\sum_{m' \in \mathbf{M}} h_{m'}} + eUI_{t,\text{HEAT LT SH}} \frac{\%sh_m}{h_m} \right) \cdot \mathbf{G}_{t,\%_{dhn}} \quad \forall m \in \mathbf{M}, t \in \mathcal{T} \tag{m.18}$$

$$\mathbf{D}_{t,\text{HEAT LT DEC},m} = \left(\frac{eUI_{t,\text{HEAT LT HW}}}{\sum_{m' \in \mathbf{M}} h_{m'}} + eUI_{t,\text{HEAT LT SH}} \frac{\%sh_m}{h_m} \right) \cdot \left(1 - \mathbf{G}_{t,\%_{dhn}}\right) \quad \forall m \in \mathbf{M}, t \in \mathcal{T} \tag{m.19}$$

$$\mathbf{D}_{t,\text{MOB. PUBLIC},m} = \frac{eUI_{t,\text{MOB. PASSENGER}}}{\sum_{m' \in \mathbf{M}} h_{m'}} \cdot \mathbf{G}_{t,\%_{public}} \quad \forall m \in \mathbf{M}, t \in \mathcal{T} \tag{m.20}$$

$$\mathbf{D}_{t,\text{MOB. PRIVATE},m} = \frac{eUI_{t,\text{MOB. PASSENGER}}}{\sum_{m' \in \mathbf{M}} h_{m'}} \cdot \left(1 - \mathbf{G}_{t,\%_{public}}\right) \quad \forall m \in \mathbf{M}, t \in \mathcal{T} \tag{m.21}$$

$$\mathbf{D}_{t,\text{FR. RAIL},m} = \frac{eUI_{t,\text{MOB. FREIGHT}}}{\sum_{m' \in \mathbf{M}} h_{m'}} \cdot \mathbf{G}_{t,\%_{rail}} \quad \forall m \in \mathbf{M}, t \in \mathcal{T} \tag{m.22}$$

$$\mathbf{D}_{t,\text{FR. ROAD},m} = \frac{eUI_{t,\text{MOB. FREIGHT}}}{\sum_{m' \in \mathbf{M}} h_{m'}} \cdot \left(1 - \mathbf{G}_{t,\%_{rail}}\right) \quad \forall m \in \mathbf{M}, t \in \mathcal{T} \tag{m.23}$$

$$\mathbf{D}_{t,\text{HEAT HIGH TEMP.},m} = \frac{eUI_{t,\text{HEAT HIGH TEMP.}}}{\sum_{m' \in \mathbf{M}} h_{m'}} \quad \forall m \in \mathbf{M}, t \in \mathcal{T} \tag{m.24}$$

$$\mathbf{D}_{t,r,m} = 0 \quad \forall m \in \mathbf{M}, \forall r \in \mathbf{RES} \setminus \{\mathbf{BF} \cup \mathbf{E}\}, t \in \mathcal{T} \tag{m.25}$$

As we can see, the variable $\mathbf{D}_{t,ELEC.,m}$ calculates the final electricity use which results from the sum of the electricity demand (constant in all months), the lighting demand, distributed over the months according to the % of lighting and the electricity loss. The low temperature heat demand is the sum of the annual demand for hot water (HW), evenly distributed over the year, and space heating (SH), distributed over the months according to the %sh. The percentage that multiplies the sum of the annual demand for hot water (HW) and space heating (SH) divides low temperature heat demand into centralized (district heating network (DHN)) and decentralized and is defined by the variable $\mathbf{G}_{t,\%_{dhn}}$. The high temperature heat and transport demand is evenly distributed among the months. Passenger mobility and freight demand are expressed in passenger-kilometers (pkms) and ton-kilometers (tkms), respectively. The variables $\mathbf{G}_{t,\%_{public}}$ defines the penetration of public transport in passenger mobility and $\mathbf{G}_{t,\%_{rail}}$ defines the penetration of rail in freight.

Layers are defined as all the elements in the system that need to be balanced in each month, such as resources and end-uses demand. For example, the electricity imported or produced in the system are layers that can be stored in hydroelectric dams or used as inputs to other energy conversion technologies (such as heat pumps) to meet the end-use demand for electricity which also represents a layer. Constraint (m.26) expresses the balance for each layer: all outputs from resources and technologies (including storage) are used to satisfy the end-uses-demand or as inputs to other resources and technologies. The matrix f defines for all technologies and resources outputs to (positive) and inputs from (negative) layers.

$$\sum_{i \in \mathbf{P} \cup \mathbf{RES} \setminus \mathbf{STO}} f_{i,l} \cdot \mathbf{z}_{t,i,m} + \sum_{u \in \mathbf{STO}} (\mathbf{Sto}^{-}_{t,u,l,m} - \mathbf{Sto}^{+}_{t,u,l,m}) - \mathbf{D}_{t,l,m} - a_l \cdot \mathbf{Loss}_{t,l,m} = 0 \quad \forall l \in \mathbf{L}, \forall m \in \mathbf{M}, t \in \mathcal{T} \tag{m.26}$$

Losses (**Loss**) are considered for the electricity grid and for the DHN, constraint (m.27) calculates the amount of electricity that is lost from both produced and imported electricity in the corresponding

layers.

$$\mathbf{Loss}_{t,eut,m} = \sum_{i \in \mathbf{RES} \cup \mathbf{P} \setminus \mathbf{STO}, f_{i,eut} > 0} f_{i,eut} \cdot \mathbf{z}_{t,i,m} \cdot \%loss_{eut} \quad \forall eut \in \mathbf{EUT}, \forall m \in \mathbf{M}, t \in \mathcal{T} \tag{m.27}$$

The constraint (m.28) shows the calculation of GHG emissions. It is calculated as the sum of the products of the amount of resource used (local or imported), the duration of each month (hour) and the emission related to the resource (gwp^{Op}).

$$\mathbf{GWP}^{Op}_{t,r} = \sum_{m \in \mathbf{M}} gwp^{Op}_{r,m} \cdot \mathbf{z}_{t,r,m} \cdot h_m \quad \forall r \in \mathbf{RES}, t \in \mathcal{T} \tag{m.28}$$

In our model there is no emissions target, so only the equation allows us to estimate the amount of emissions produced by the system in each time period.

A.4 Additional constraints

Equations (m.1)-(m.28) define the main constraints of the energy model, which can be adapted or forced by adding constraints that limit the degrees of freedom of the model. Additional restrictions necessary for the specific case of the Swiss energy system are introduced below.

Long-term investment equations

Constraint (m.29) links linearly the storage capacity to the new installed power. Constraint (m.30) associates the cost of investment the *PowerToGas* unit to the maximum size of two conversion units. This constraint is displayed in a compact nonlinear formulation.

$$\mathbf{y}_{t,StoHydro} \leq f^{max}_{StoHydro} \frac{\mathbf{y}_{t,NewHydroDam} - f^{min}_{NewHydroDam}}{f^{max}_{NewHydroDam} - f^{min}_{NewHydroDam}}, \quad \forall t \in \mathcal{T} \tag{m.29}$$

$$\mathbf{y}_{t,PowerToGas} = \max\left\{\mathbf{y}_{t,PowerToGas}; \mathbf{y}_{t,GasToPower}\right\}, \quad \forall t \in \mathcal{T} \tag{m.30}$$

Constraint (m.31) is used to calculate the energy efficiency as a fixed cost. Constraint (m.32) represent an additional investment cost of 9.4 billion CHF_{2015} is linked proportionally to the deployment of stochastic renewables. Constraint (m.33) links the DHN size to the total size of the installed centralized energy conversion technologies.

$$\mathbf{y}_{t,EFFICIENCY} = \frac{1}{1 + i_{rate}}, \quad \forall t \in \mathcal{T} \tag{m.31}$$

$$\mathbf{y}_{t,Grid} \geq 1 + \frac{9400}{c^{Inv}_{t,Grid}} \frac{\mathbf{y}_{t,Wind} + \mathbf{y}_{t,PV}}{f^{max}_{Wind} + f^{max}_{PV}}, \quad \forall t \in \mathcal{T} \tag{m.32}$$

$$\mathbf{y}_{t,DHN} \geq \sum_{p \in \mathbf{P\text{-}OF\text{-}EUT}_{\text{HEAT LT DHN}}} \mathbf{y}_{t,p}, \quad \forall t \in \mathcal{T} \tag{m.33}$$

Constraint (m.34) complies with the decision of the Swiss government to phase out the nuclear power plants at the end of their useful life. Therefore there is no investment in nuclear plants over the entire planning horizon

$$\mathbf{y}_{t,NUCLEAR} = 0 \quad \forall t \in \mathcal{T} \tag{m.34}$$

Short-term operating equations

Constraint (m.35) is used to avoid underestimating the cost of centralized heat production, a multiplication factor is introduced to account for peak demand, defined as a $\%Peak_{DHN}$ times the maximum monthly average heat demand (also expressed in a compact non-linear formulation[5]).

$$\sum_{i \in \textbf{P-OF-EUT}_{\text{HEAT LT DHN}}} \mathbf{x}_{t,i} \geq \%peak_{DHN} \max_{m \in \mathbf{M}} \left\{ \mathbf{D}_{t,\text{HEAT LT DHN},m} + \mathbf{Loss}_{t,\text{HEAT LT DHN},m} \right\}, \quad t \in \mathcal{T} \tag{m.35}$$

Constraints (m.36)-(m.37) are complementary to constraint (m.8), as it expresses the minimum ($f^{min,\%}$) and maximum ($f^{max,\%}$) yearly output shares of each technology for each type of end-use demand.

$$\sum_{m \in \mathbf{M}} \mathbf{z}_{t,p,m} h_m \geq f_p^{min,\%} \sum_{p' \in \textbf{P-OF-EUT}_{eut}} \sum_{m \in \mathbf{M}} \mathbf{z}_{t,p',m} h_m \quad \forall eut \in \mathbf{EUT}, \forall p \in \textbf{P-OF-EUT}_{eut}, t \in \mathcal{T} \tag{m.36}$$

$$\sum_{m \in \mathbf{M}} \mathbf{z}_{t,p,m} h_m \leq f_p^{max,\%} \sum_{p' \in \textbf{P-OF-EUT}_{eut}} \sum_{m \in \mathbf{M}} \mathbf{z}_{t,p',m} h_m \quad \forall eut \in \mathbf{EUT}, \forall p \in \textbf{P-OF-EUT}_{eut}, t \in \mathcal{P} \tag{m.37}$$

Constraint (m.38) imposes that the share of the different technologies for mobility be the same in each month.

$$\mathbf{z}_{t,mob,m} \sum_{m' \in \mathbf{M}} h_{m'} \geq \sum_{m' \in \mathbf{M}} \mathbf{z}_{t,mob,m'} h_{m'} \quad \forall m \in \mathbf{M}, \forall mob \in \mathbf{MOB}, t \in \mathcal{T} \tag{m.38}$$

Constraint (m.39) is a complement of constraint (m.29) and it is used for ensures that the shifted production in a given month does not exceed the electricity production by the dams in that month.

$$\mathbf{Sto}^{+}_{t,StoHydro,Elec,m} \leq \mathbf{z}_{t,HydroDam,m} + \mathbf{z}_{t,NewHydroDam,m} \quad \forall m \in \mathbf{M}, t \in \mathcal{T} \tag{m.39}$$

[5] All equations expressed in a compact non-linear form in this section (Eqs. m.37 y m.42) can be easily linearized

B Assumptions for uncertain parameters

Table 3: Nominal, min and max values for end-uses demands in heating, electricity and mobility sectors. Abbreviations: Temperature (T), Hot Water (HW), Space Heating (SH)

		$eUI_{year,EUIdemands}$ $[GWh]$		
EUI demands	Year	Nominal	Min	Max
HEAT_HIGH_T	2020	21,535	13,283	29,786
	2030	19,144	11,219	27,070
	2040	17,413	8,475	26,350
	2050	15,925	7,174	24,676
	2060	13,864	4,770	22,958
HEAT_LOW_T_SH	2020	61,432	42,379	80,486
	2030	52,416	31,806	73,026
	2040	44,451	20,280	68,622
	2050	37,013	14,235	59,791
	2060	33,352	9,646	57,059
HEAT_LOW_T_HW	2020	11,116	8,860	13,372
	2030	11,575	8,797	14,354
	2040	11,891	7,848	15,935
	2050	12,179	7,386	16,973
	2060	12,567	6,879	18,255
ELECTRICITY	2020	34,829	26,441	43,217
	2030	34,754	24,561	44,948
	2040	35,469	21,207	49,731
	2050	37,001	19,929	54,074
	2060	35,322	17,312	53,332
LIGHTING	2020	6,149	4,635	7,663
	2030	5,495	4,037	6,953
	2040	5,202	3,264	7,140
	2050	5,000	2,977	7,023
	2060	4,961	2,910	7,013
MOBILITY_PASSENGER	2020	131,144	115,239	147,049
	2030	141,147	118,635	163,659
	2040	150,250	119,851	180,649
	2050	151,351	111,305	191,396
	2060	160,904	113,611	208,197
MOBILITY_FREIGHT	2020	33,917	29,227	38,607
	2030	38,776	33,001	44,551
	2040	41,255	32,516	49,995
	2050	41,950	29,500	54,399
	2060	45,619	31,145	60,093

Table 4: Nominal, min and max values for resource costs. Abbreviations: Natural Gas (NG), Synthetic Natural Gas (SNG), Carbon Capture and Storage (CCS)

		$C^{Op}_{year,resources,\cdot}$ $[MCHF/GWh]$		
Resources	Year	Nominal	Min	Max
ELECTRICITY	2010	0.0662	0.0564	0.1058
	2020	0.0757	0.0595	0.1287
	2030	0.0833	0.0585	0.1541
	2040	0.0909	0.0578	0.1736
	2050	0.0927	0.0554	0.1780
GASOLINE/BIOETHANOL	2010	0.0709	0.0604	0.1134
	2020	0.0775	0.0662	0.1243
	2030	0.0849	0.0618	0.1527
	2040	0.0920	0.0604	0.1757
	2050	0.1002	0.0565	0.1964
DIESEL/BIODIESEL	2010	0.0686	0.0399	0.1029
	2020	0.0752	0.0591	0.1279
	2030	0.0822	0.0577	0.1520
	2040	0.0892	0.0567	0.1703
	2050	0.0972	0.0548	0.1904
LFO	2010	0.0488	0.0455	0.0781
	2020	0.0535	0.0420	0.0909
	2030	0.0584	0.0411	0.1081
	2040	0.0634	0.0403	0.1210
	2050	0.0690	0.0389	0.1332
NG/NG_CCS/SNG	2010	0.0248	0.0212	0.0422
	2020	0.0288	0.0202	0.0518
	2030	0.0329	0.0196	0.0608
	2040	0.0370	0.0187	0.0706
	2050	0.0414	0.0182	0.0811
WOOD	2010	0.0475	0.0421	0.0537
	2020	0.0658	0.0460	0.0710
	2030	00797	0.0436	0.0900
	2040	0.0937	0.0420	0.1246
	2050	0.0949	0.0407	0.1451
COAL/COAL_CCS	2010	0.0206	0.0201	0.0371
	2020	0.0244	0.0192	0.0452
	2030	0.0274	0.0189	0.0515
	2040	0.0303	0.0187	0.0573
	2050	0.0306	0.0183	0.0579
URANIUM	2010	0.0038	0.0032	0.0046
	2020	0.0039	0.0031	0.0051
	2030	0.0040	0.0028	0.0064
	2040	0.0041	0.0026	0.0081
	2050	0.0041	0.0024	0.0081

C Modeling the Price of Resources and Demands

Figure 11 below displays the distribution of 30'000 scenarios of demands of electricity, heating and lighting for different values of α using the AR processes. The figure also highlights 7 random trajectories (in blue color). Figure 12 is similar but for energy prices.

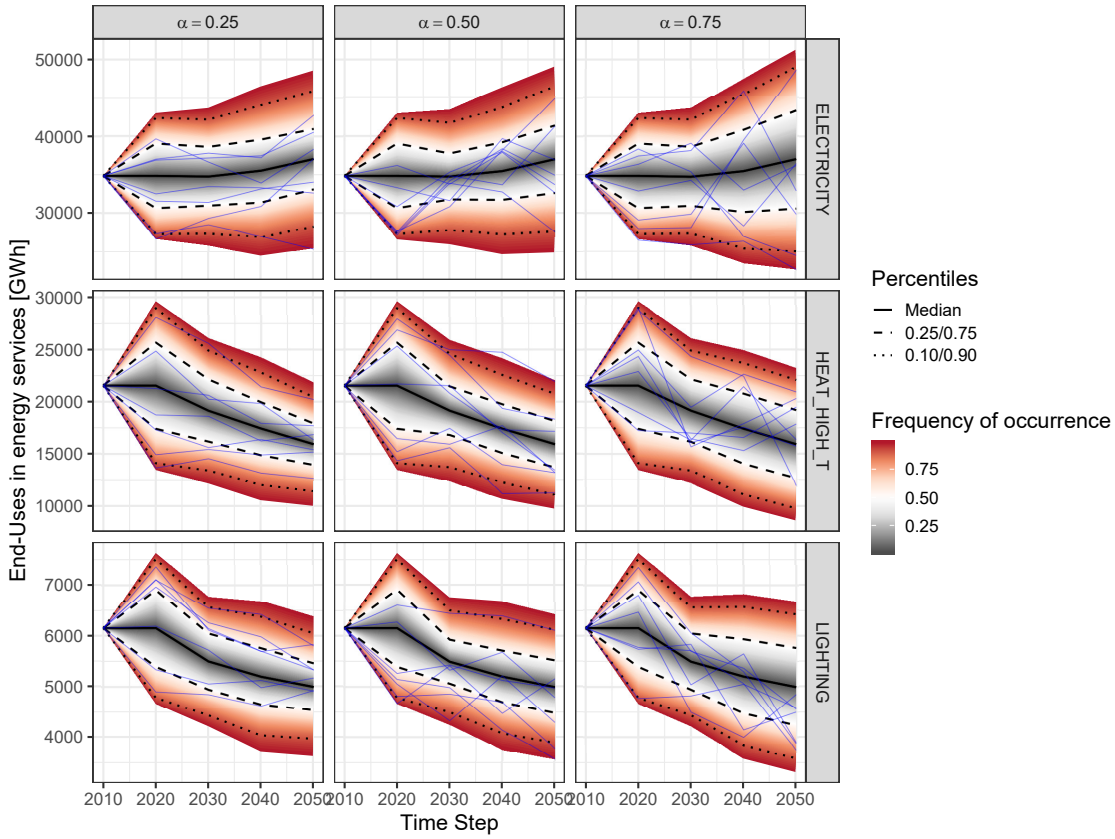


Figure 11: 30,000 scenarios generated from the original AR processes with $\alpha \in 0.75, 0.50, 0.25$ for demand of ELECTRICITY, HEAT_HIGH_T and LIGHTING. .

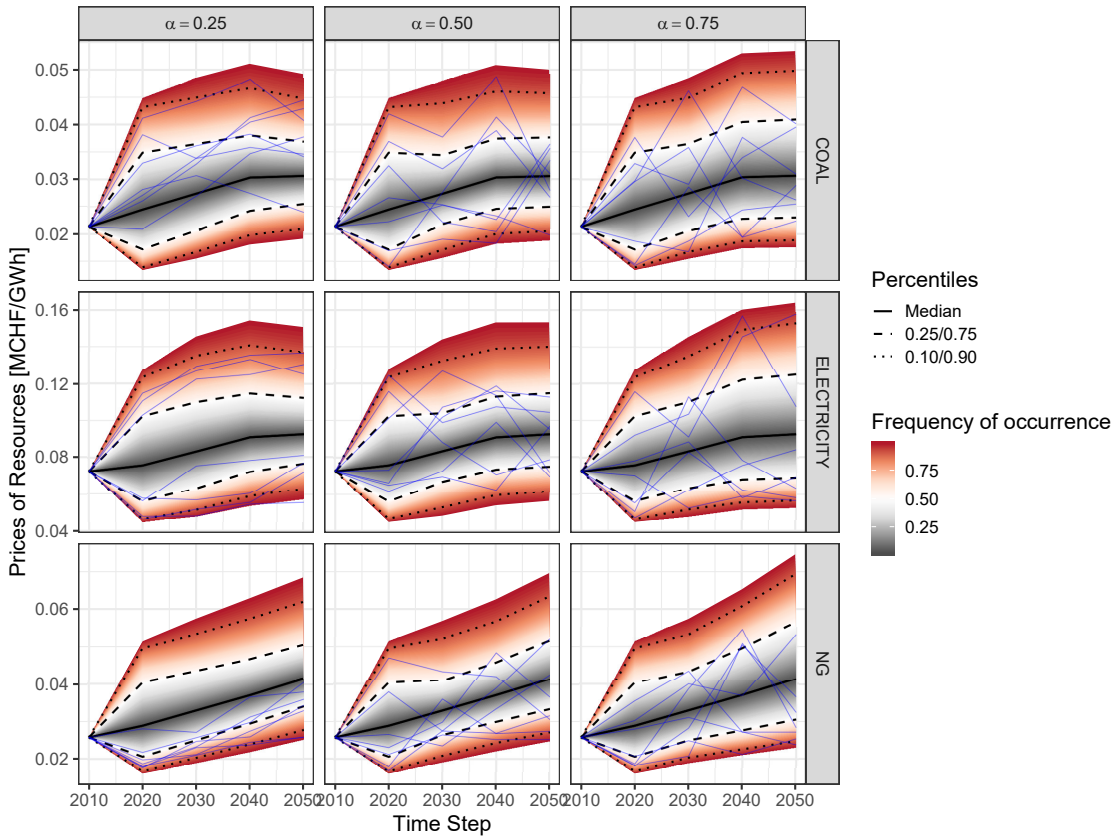


Figure 12: 30,000 scenarios generated from the original AR processes with $\alpha \in 0.75, 0.50, 0.25$ for price of electricity, coal and natural gas.

D Markov chains for electricity and natural gas prices

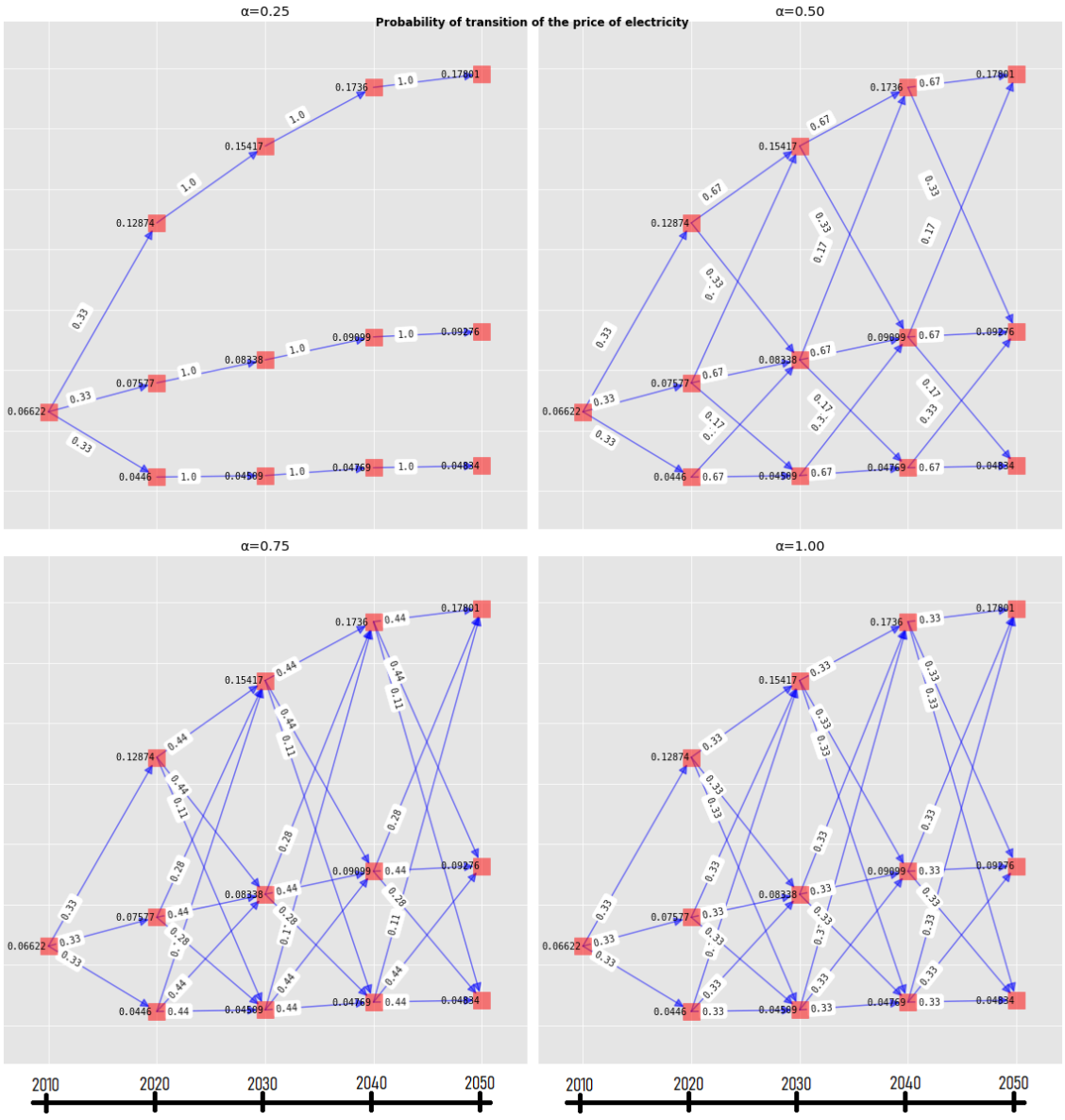


Figure 13: Markov Chains for electricity prices for different values of α.

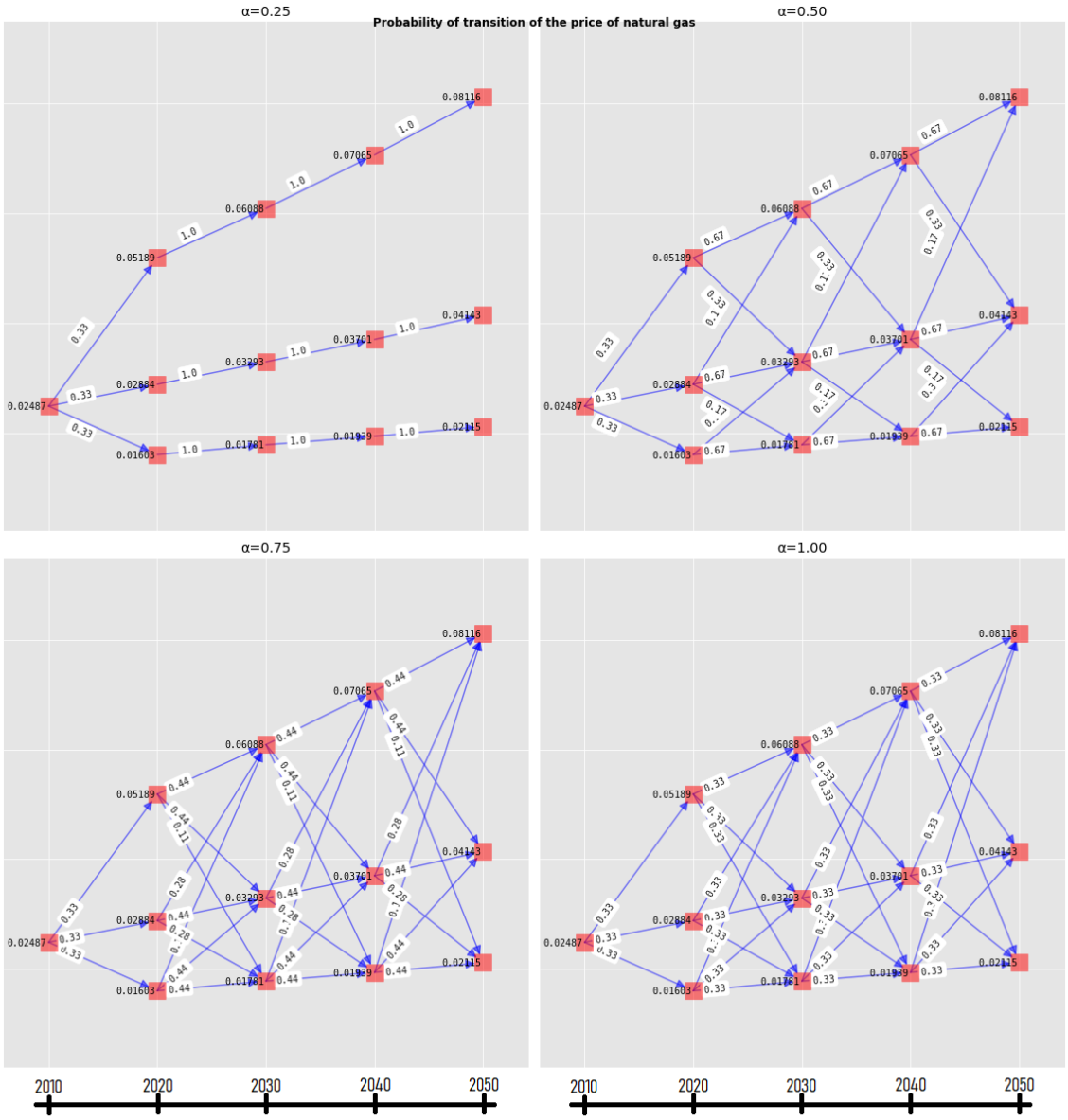


Figure 14: Markov Chains for Natural Gas prices for different values of α.