A tailored decomposition approach for optimization under uncertainty of carbon removal technologies in the EU power system

Valentina Negri[1], Daniel Vázquez[1], Ignacio E. Grossmann[2], Gonzalo Guillén-Gosálbez[1*]

[1]Institute for Chemical and Bioengineering, Department of Chemistry and Applied Biosciences, ETH Zurich, Vladimir-Prelog-Weg 1, 8093 Zurich, Switzerland
[2]Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA
*Corresponding author: Gonzalo Guillén Gosálbez: gonzalo.guillen.gosalbez@chem.ethz.ch

Abstract

The broad portfolio of technologies delivering negative emissions calls for integrated analyses to explore the synergies between them and the power sector, with which they display strong links. These analyses, encompassing carbon removal and power generation, should be conducted at a regional level, considering system uncertainties, to assess local benefits and the impact on carbon removal potential. This study investigates how electricity demand uncertainty affects the optimal design of integrated carbon removal and power generation systems using multi-stage stochastic programming. Given the model complexity, we propose a tailored decomposition approach that reduces the computational time by 90% enabling insights into various European scenarios. We find that a combination of 12 technologies and biomass could achieve up to 9 Gt of CO_2 of net removal from the atmosphere. Omitting uncertainties leads to an underestimation of the total cost and the selection of different technologies that might lead to suboptimal performance.

Keywords

EU energy system – carbon removal – exogenous uncertainty – multistage stochastic optimization - decomposition method

Nomenclature

Abbreviations	
BECCS	bioenergy with carbon capture and storage
CCS	carbon capture and storage
CO_2	carbon dioxide
CDR	carbon dioxide removal
DACCS	direct air capture with carbon capture and storage
EU	European Union
EV	expected value problem
GHG	greenhouse gas
IAMs	Integrated assessment models
LP	linear programming model
MILP	mixed-integer linear programming model
MSS	multistage stochastic programming model

NETPs	negative emissions technologies and practices
NACs	non-anticipativity constraints
RAPIDU	RemovAl oPtImization model under Uncertainty
SAA	sample average approximation
RP	solution to the fully stochastic problem
2SS	two-stage stochastic programming model
VSS	value of the stochastic solution

Sets
I $\{i \mid i$ is a technology$\}$
J $\{j \mid j$ is a country$\}$
$\mathcal{S}$ $\{s \mid s$ is a scenario$\}$
$\mathcal{SP}_F$ $\{s \mid s$ is a scenario included in the first period NACs$\}$
$\mathcal{S}P_X$ $\{s \mid s$ is a scenario included in the exogenous period NACs$\}$
S_{red} $\{s \mid s$ is a scenario included in the solution of step 1$\}$
$\mathcal{T}$ {t | t is a time period}

Variables	
y_{ijt}^s	first-stage investment decisions for technology i, country j and time period t in scenario s
b_{ijt}^s	first-stage binary decisions for technology i, country j and time period t in scenario s
x_{ijt}^s	second-stage operation decisions for technology i, country j and time period t in scenario s

Parameters	
N	number of exogenous scenarios s
${}^{y}A_{ijt}^s$	parameters matrix, dependent on the variable and scenario s, for technology i, country j and time period t
p^s	probability of occurrence of scenario s

1. Introduction

The urgent need to mitigate the adverse effects of climate change has led to a growing emphasis on reducing greenhouse gas (GHG) emissions, whose rising concentration is primarily attributed to anthropogenic activities (Daggash and Mac Dowell, 2019). Among all the economic industries, the energy sector, which includes heat and electricity, contributed to 57% of global emissions in 2016 (Ritchie et al., 2020), making it a key player in meeting the climate targets.

The European Union (EU), responsible for roughly 10% of global emissions and one major leader in climate policies (Meckling et al., 2017), published in 2019 the European Green Deal to push a political shift toward a carbon-neutral society in 2050 (European Commission, 2019; Schenuit et al., 2021). This includes efforts to decarbonize the energy sector through renewable energy deployment (amounting to 32% of the gross final energy consumption) and energy efficiency measures to cut GHG emissions by 55% (Victoria et al., 2020).

While these actions, already included in Integrated assessment models (IAMs), are underway, there is also the need to explore comprehensive strategies that not only reduce emissions but also actively reduce the concentration of GHG in the atmosphere (IPCC, 2022). This could be accomplished via carbon dioxide (CO_2) removal (CDR) strategies or negative emissions technologies and practices (NETPs). Some of these, namely bioenergy (Solano Rodriguez et al., 2017) and direct air capture with carbon capture and storage (CCS) (BECCS and DACCS, respectively), are already included in the IAMs (Realmonte et al., 2019). Yet, IAMs do not incorporate social or political dimensions (Doukas et al., 2018) and do not have the technological or spatial resolution for detailed energy systems planning. Moreover, when assessing multiple key performance indicators of these technologies, trade-offs appear, suggesting that a regionalized portfolio of NETPs integrated into the energy systems should be evaluated to exploit their complementary strengths (Cobo et al., 2023). However, only a few studies model this coupling explicitly (Bistline and Blanford, 2021; Creutzig et al., 2019). A few regional studies were carried out by Daggash et al. (2019) in the context of the United Kingdom and by Sagues et al. (2019) in the United States. Recently, a model that integrates the EU power system together with BECCS and DACCS was developed to shed light on the consequences of delaying the deployment of CDR options (Galán-Martín et al., 2021). Despite this work providing valuable insights into the optimal deployment pathways, the conclusions were drawn based on a deterministic model, which should be reevaluated at each time period due to the change in data, such as energy demand and technology learning rate, as explored by Rathi and Zhang (2022), and lacks the possibility of incorporating uncertain parameters.

Uncertainty is an inherent characteristic of energy systems, as they are influenced by a complex interplay of factors such as technological advancements, policy frameworks, economic fluctuations, and societal behaviors. Incorporating uncertainty analysis in energy system models is crucial for robust decision-making, particularly in the context of evaluating carbon removal options. Specifically for BECCS and DACCS, since they are often evaluated within long-term energy plans, the results are affected by a considerable degree of uncertainty (Fajardy et al., 2019), which is usually neglected, while in practice their large-scale deployment in the EU energy system is subject to numerous challenges. These challenges include technological readiness, high costs, land availability for biomass cultivation, scalability of the modularity of DACCS technologies, and social acceptance. Moreover, the variability and uncertainty in electricity demand pose additional complexities for the integration of these technologies into the energy system. Indeed, these CDR options are interconnected with the system because they either provide or require electricity. Thus, it is important to address the uncertain nature of the EU energy system, highlighted especially in recent years, due to the COVID-19 pandemic and the Ukraine invasion.

Traditional deterministic energy system models often overlook the uncertain nature of electricity demand, potentially leading to suboptimal or even infeasible decisions and unrealistic outcomes (Li and Grossmann, 2021). Hence, here we argue that optimization under uncertainty and, particularly, multistage stochastic programming, is essential to comprehensively assess the potential role of BECCS and DACCS in the EU energy system under uncertainty.

We refer to a recent review of stochastic programming published by Li and Grossmann (2021) for a list of selected articles on stochastic problems, where it became evident that the Process Systems Engineering literature on stochastic models is quite rich. Yet, the literature on sustainability problems under uncertainty is scarce, despite CCS being recognized in the field as pivotal to the future transition (Hasan et al., 2022). In particular, to our knowledge, energy system models that include CDR options and exogenous uncertainty are yet to be explored. Along these lines, Grant and co-authors (2021) carried out a pioneering work evaluating the policy implications of uncertain removal potential at a global scale using the TIAM-Grantham model in stochastic scenarios.
Analyzing the CDR-power nexus including uncertainty leads to much more complex problems than their deterministic counterparts. Despite recent developments in hardware and software, stochastic problems might still be intractable and require decomposition approaches to speed up the solution time to get insights from different case studies with reduced computational time. Traditional decomposition methods to solve large stochastic models include Benders and Lagrangean decomposition, sequential scenario decomposition (Apap and Grossmann, 2017), and shrinking and rolling horizon (Balasubramanian and Grossmann, 2004). Tailored approaches have also been developed, as a combination of these methods (Fusco et al., 2023; Lee et al., 2023; Meersman et al., 2023).

Here, we evaluate the impact of electricity demand uncertainty on the deployment of BECCS and DACCS in the EU energy system via multistage stochastic programming. A tailored algorithm is introduced to decompose the problem, reducing the computational time significantly. This allows us to generate insights into the optimal integration of these technologies within the EU energy system by investigating different scenarios of carbon removal by BECCS and DACCS.

The rest of the article is organized as follows. The problem statement is given next, followed by the optimization methods, including the multistage stochastic model and the decomposition algorithm. The results section includes a comparison of the computational performance of the stochastic model without and with the decomposition algorithm for different case studies. Lastly, the conclusions and outlook for future work are presented.

2. Problem statement

Given are 15 state-of-the-art power technologies and the most prominent CDR-engineered options. These technologies are divided into dispatchable, which can produce heat, electricity, or both, at demand, and non-dispatchable, which require a backup capacity of dispatchable technologies to account for the time they cannot be operated. Among these, we also consider conventional fossil technologies, i.e., coal and natural gas, and their retrofit with carbon capture and storage. Among the CDR options, we include biomass-based energy, which also acts as a dispatchable energy source, and direct air capture. We consider learning cost curves and realistic diffusion rates limiting all technologies' deployment. The region of interest of this study is the EU with the United Kingdom, where its member states are considered potential locations for the installation of power technologies. These 28 countries are modeled as load nodes with specific energy demands and resource availability.

The goal is to determine the optimal planning of the EU energy system under electricity demand uncertainty that satisfies the energy demand in each country, minimizing the system's total cost subject to a given cumulative CDR target at the end of the time horizon of 30 years. The decisions to be made include the capacity and location of the power technologies and the electricity, heat required by DACCS, and CO_2 flows between the EU states. The problem statement is presented graphically in Figure 1.

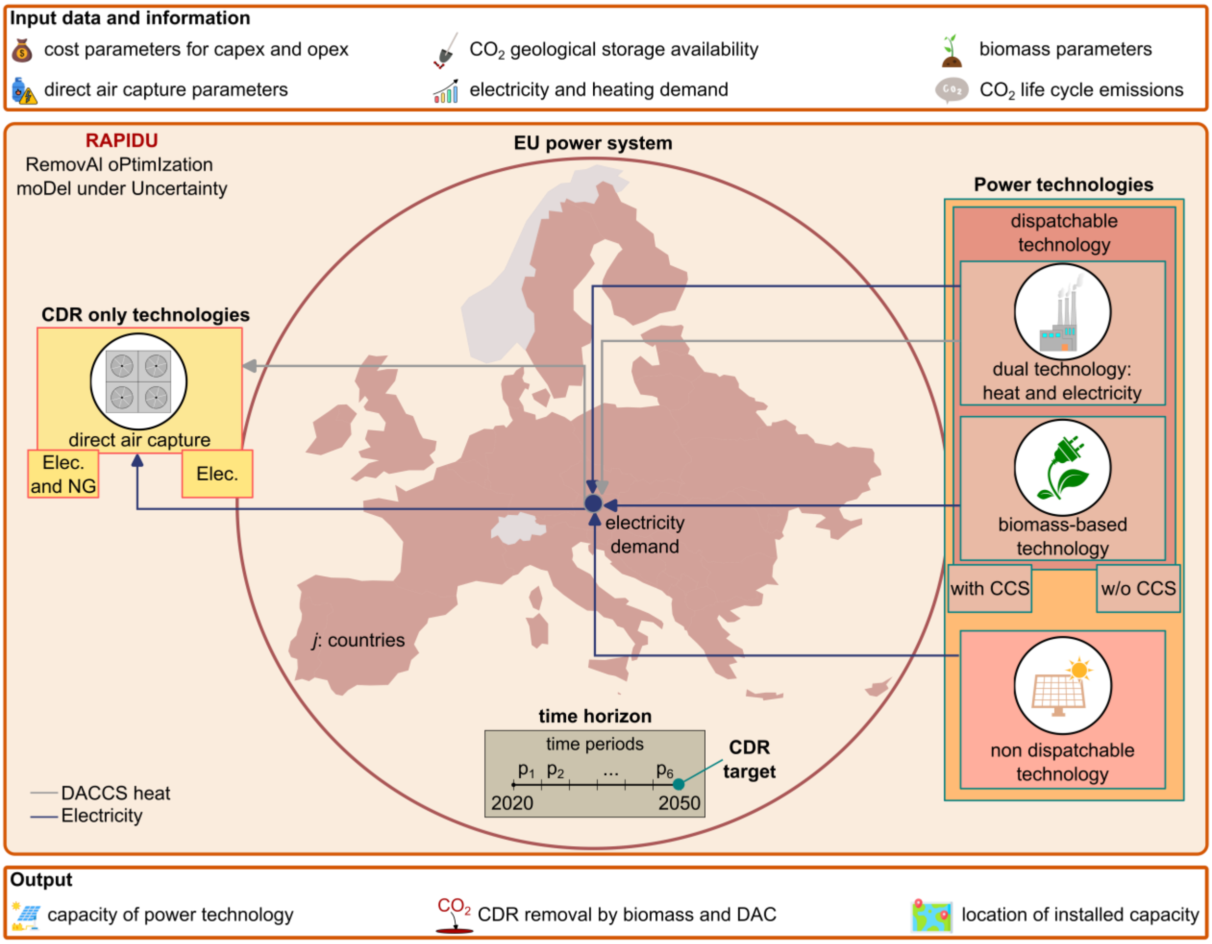

Figure 1. RAPIDU problem statement. We are given a set of power technologies and two carbon dioxide removal technologies already commercially available, namely BECCS and DACCS. We model the 27 EU countries and the UK as load nodes with specific energy demands and resource availability. The goal is to determine the optimal deployment pathway under electricity demand uncertainty (exogenous uncertainty) that meets a given CDR target at the end of the time horizon of 30 years, minimizing the system's total cost.

3. Methods

3.1 Deterministic model

We build on the RAPID model, previously proposed to understand the implications of delaying the deployment of CDR options at a regional level, focusing on BECCS and DACCS within the EU power system (Galán-Martín et al., 2021). RAPID was developed to investigate two scenarios: maximization of carbon removal and minimization of total cost subject to a carbon removal target. Perfect foresight for the input parameters was assumed in the time horizon 2020 – 2100, modeled as a set of discrete time periods. The carbon removal is computed as the difference between the positive CO_2 emissions and the CO_2 removed from the atmosphere by BECCS and DACCS. In essence, given a set of power technologies and BECCS and DACCS that can be deployed in each country of the EU, the goal is to find the optimal capacity installed and the removal by the two NETPs, so as to meet the CDR target at the end of the time horizon and the energy demand in each country.

In this work, we include binary variables (b_t^s) that account for the installation of the technologies at the beginning of each time period following (Guillén-Gosálbez et al., 2010; Sahinidis et al., 1989). We shorten the time horizon to 30 years, from 2020 to 2050, for a more realistic evaluation of the uncertainty within a net zero emissions target of the EU. Therefore, we impose a cumulative CDR target equal to zero in the year 2050, which is achieved jointly by all the countries in a cooperative strategy, and we solve for the minimum total cost of the system.

3.2 Uncertainty definition

Next, we consider exogenous uncertainty in the electricity demand by formulating a multistage stochastic mixed-integer linear programming (MSS - MILP) model (Apap and Grossmann, 2017). We consider that the electricity demand varies ± 20% from the nominal deterministic value. The choice of the demand as an uncertain parameter is justified by the fact that this parameter is dependent on multiple external factors, and hence can be highly uncertain. Some of the factors that affect demand are seasonality, economic growth and contraction, population growth and urbanization, technological innovation, e.g., increased efficiency, natural disasters, and even political and socio-cultural landscape changes.

Depending on how the uncertainty reveals, it can be characterized as exogenous or endogenous (Gupta and Grossmann, 2011). The former is decision independent and it is revealed automatically at each time period, e.g., market prices. In contrast, the latter is decision dependent; therefore, it is not associated with a particular time period, e.g., the size of an oil field, which is revealed only when the drilling starts (Grossmann et al., 2016). Multiple approaches to deal with uncertainty in optimization problems have been developed, including stochastic programming (Birge and Louveaux, 2011), chance-constrained (Li et al., 2008), and robust optimization (Lappas and Gounaris, 2016), which differ in the way uncertainty is characterized and the degree of risk aversion. Stochastic programming is a risk-neutral approach, where the uncertainty is characterized by a given probability distribution and in which the expected value of the objective function is optimized. In contrast, in chance-constrained there is the possibility to deal with reliability and risk management. In essence, it requires solving a stochastic programming problem with some probabilistic constraints. Robust optimization is also a risk-averse approach, which tries to find an optimal solution to the "worst-case scenario" that satisfies given constraints over a defined uncertainty set. The method of choice depends, among others, on the information available on the uncertain data, whether the emphasis is on feasibility or optimality, and if corrective actions can be taken. For more details on approaches for decision-making under uncertainty, we refer the reader to Apap and Grossmann (2017) and Li and Grossmann (2021).

3.3 Multistage stochastic model

In two-stage stochastic programming with exogenous uncertain parameters, we distinguish between first-stage and second-stage decision variables. The former are also called "here and now" variables, and are fixed at the beginning of the time period before knowing how the uncertainty will unfold. The latter, also known as recourse actions ("wait-and-see"), perform a corrective action after the uncertainty is revealed. In a multistage stochastic problem decisions, realizations, and recourse actions occur sequentially as represented in Figure 2, and multiple recourse actions can be taken as uncertainties are gradually revealed.

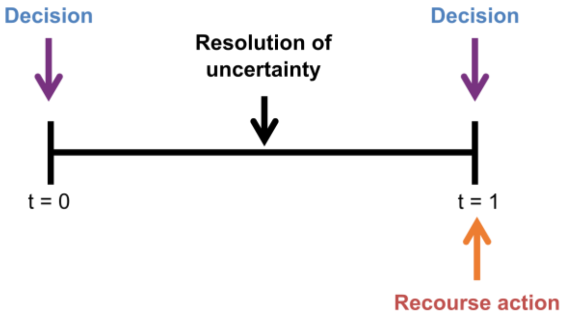

Figure 2. Sequence of events in stochastic programming with one endogenous uncertain parameter (Apap and Grossmann, 2017).

In what follows, we first describe the general modeling approach, and then provide the stochastic formulation taken as a basis. In stochastic programming, uncertainty is assumed to be described via scenario trees obtained from discretized probability distribution functions. Let us consider three time periods and two realizations of one uncertain parameter, high and low. For these assumptions, the standard scenario tree is represented in Figure 3a. Hereafter the following notation is used. A node is a possible state in a time period t. An arc is a possible transition from a state in t to a new state in t+1. A scenario is the complete path from the root node to a leaf node. We refer to the scenario probability as the probability of reaching a leaf node from the root node.

The number of exogenous scenarios (N) is equal to the product of the number of realizations for each exogenous parameter. If we consider two realizations of the uncertain parameter at each time period, the cardinality of the scenarios set is computed as 2^t and the number of nodes is $2^{t+1} - 1$.

An alternative form of the scenario tree in Figure 3b proposed by Ruszczyński (1997) gives each scenario a unique set of nodes. When moving from the standard to the alternative representation we create several copies of the same states that contain the same information at that point in time. Scenarios with the same information at time t are said to be indistinguishable at that time. Therefore, in indistinguishable scenarios, we must make the same decisions. This behavior is enforced through non-anticipativity constraints (NACs) represented by the red horizontal lines in Figure 3b.

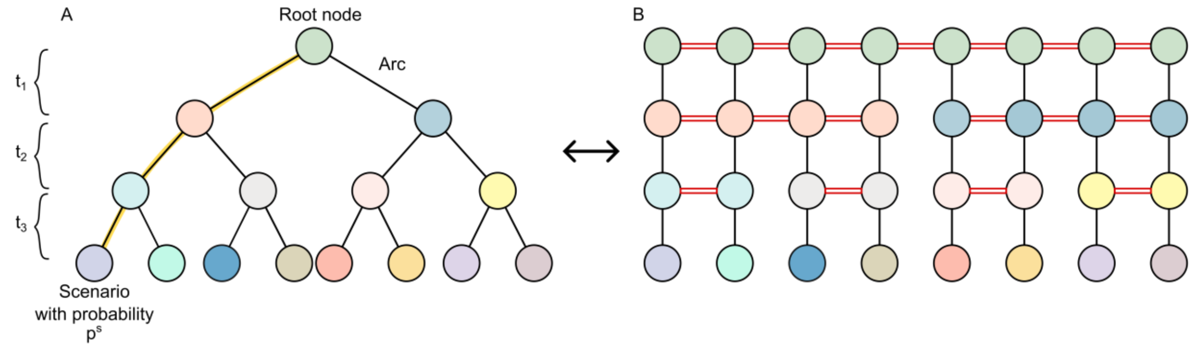

Figure 3. Exogenous uncertainty representation: standard (A) and alternative scenario tree (B). t represent the time periods, within t_{start} and t_{end}, each dot is a node and the black lines connecting the dots are called arc, which represent a possible transition between two states from time t to time t+1. The node at the top is the root node and the ones at the bottom are the leaf nodes. A complete path from the root node to a leaf node corresponds to a scenario, which occurs with a given probability.

We report in Eqs. ((*1*) – ((5) the general mathematical formulation of RAPIDU (RemovAl oPtImization model under Uncertainty), hereafter referred to as MSS1. The stochastic model is represented in the deterministic-equivalent form using non-anticipativity constraints ((3a)-(4c)).

$$\min_{b,y,x} \phi = \sum_{s\in\mathcal{S}} p^s \sum_{t\in\mathcal{T}} \left({}^{y}c_t^s y_t^s + {}^{x}c_t^s x_t^s + {}^{b}c_t^s b_t^s\right) \tag{1}$$

$$\text{s.t.}\ {}^{y}A_t^s y_t^s + {}^{x}A_t^s x_t^s + {}^{b}A_t^s b_t^s \leq a_t^s \quad \forall\, t\in\mathcal{T},\ \ s\in\mathcal{S} \tag{2}$$

$$b_1^s = b_1^{s'}\ \forall\,(s,s')\in\mathcal{SP}_F \tag{3a}$$

$$y_1^s = y_1^{s'}\ \forall\,(s,s')\in\mathcal{SP}_F \tag{3b}$$

$$x_1^s = x_1^{s'}\ \forall\,(t,s,s')\in\mathcal{S}P_X \tag{4a}$$

$$b_{t+1}^s = b_{t+1}^{s'}\ \forall\,(t,s,s')\in\mathcal{S}P_X \tag{4b}$$

$$y_{t+1}^s = y_{t+1}^{s'}\ \forall\,(t,s,s')\in\mathcal{S}P_X \tag{4c}$$

$$b_t^s \in \{0,1\},\ \ y_t^s \in \mathcal{Y}_t^s,\ \ x_t^s \in \mathcal{X}_t^s \quad t\in\mathcal{T},\ \ s\in\mathcal{S} \tag{5}$$

The objective function in Eq. (*1*) is the total expected cost, computed as the weighted sum of the costs in each scenario multiplied by the probability in each scenario, p^s. The general techno-economic constraints correspond to Eq. (2). Eqs. (3a) represent the first-period NACs, and Eqs. (4a) the exogenous NACs. These constraints enforce that the same decisions are taken in all the nodes that are indistinguishable. Lastly, in Eq. (5) the variables' bounds and integrality restrictions are specified.

Here, we adopt the same nomenclature used in Apap and Grossmann (2017) for the mathematical formulation. The multistage stochastic optimization model includes y_t^s first-stage investment decisions at the beginning of each time period, e.g., power technology capacity installed; x_t^s second-stage operation decisions that follow the investment decisions, e.g., the electricity generated from the installed capacity. These decisions are optimized over every country j ∈ J considering a set of technologies i ∈ I whose sets are omitted in the mathematical formulation for clarity (i.e., y_t^s corresponds to y_{ijt}^s). Each variable is scenario-dependent, identified by the set $\mathcal{S}$, indicated by the superscript on the right of each variable. MSS1 is presented in its most general form, where the left superscript of the A matrix indicates to which variable the parameters refer and even the cost parameters are indexed for s to allow for different values according to the scenario. E.g., ${}^{y}A_t^s$ means that the parameters included in A refer to the y_t^s decision variable at time period t in scenario s. We refer to Apap and Grossmann (2017) for the mathematical definition of the sets $\mathcal{SP}_F$ and $\mathcal{S}P_X$.

The number of scenarios, corresponding to the cardinality of set $\mathcal{S}$, is 64 (2^t) given two realizations of the uncertainty, high and low, at each time period. The realization of the uncertainty is the same for each country, i.e., in all EU countries and the United Kingdom the electricity demand either increases or decreases, which is a sensible assumption given the strong political ties in the region and assuming that the extent to which the economy will be electrified, particularly concerning transportation but also in hard-to-abate sectors like petrochemicals production, will be similar across countries. Therefore, here we do not consider the combination of all the scenarios, which would increase the complexity of the problem substantially.

The time horizon 2020 - 2050 is divided into six time periods $t \in \mathcal{T}$. We first consider all the time periods to be of equal length, and then we differentiate the lengths depending on the time period, with shorter time periods closer to the beginning of the time horizon and a rougher discretization towards the end. We adjust the input data accordingly to account for CAPEX and OPEX to match the new definition of the time periods.

A detailed description of the general techno-economic model constraints that are included in Eq. (2) can be found in the supplementary material of Galán-Martín et al. (2021). In essence, load-meeting and operation constraints are defined in Eqs. 1 – 31, and account for the distinction between dispatchable and non-dispatchable technologies, and make sure that the electricity demand is fulfilled. We note that in this work the electricity demand is met as an equality constraint. The modeling of the supply chain of biomass is also included, regarding land use, availability of residues, etc. CO_2 emission constraints include Eqs. 32 – 42, and model the carbon balance regarding positive emissions and negative emissions, as well as the trade, capture, and storage of CO_2 and its sequestration underground. Finally, cost equations include Eqs. 43 – 48, and model how the CAPEX and OPEX of the different power energy sources are obtained, as well as the cost of negative technologies.

3.4 Decomposition algorithm

The full space multistage stochastic programming in Eqs. (*1*) - (5) leads to large-scale problems that quickly become intractable. For example, for six periods and 28 countries, the MILP has more than 8.2 million variables and 4.2 million constraints. Hence, we propose the decomposition algorithm sketched in Figure 4 for its efficient solution.

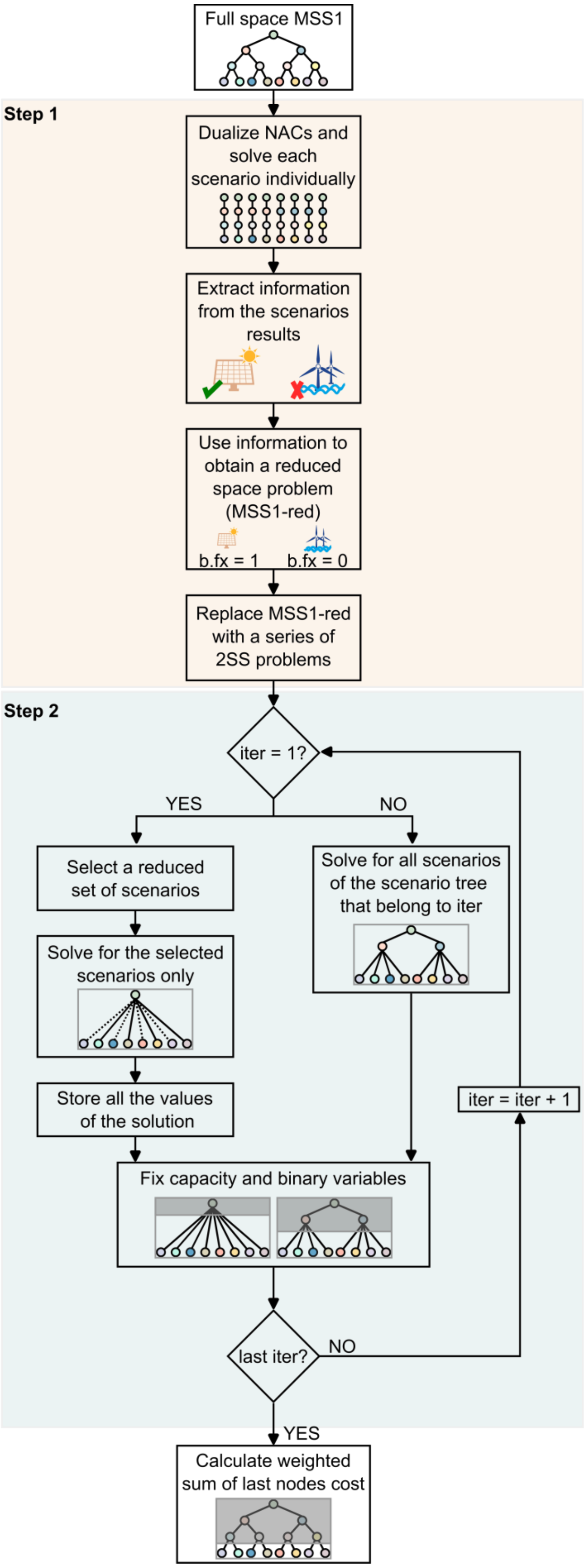

Figure 4. Decomposition algorithm steps. In the first step, we decompose the original problem into scenarios and we solve them individually. Afterward, we extract relevant information from the single scenarios. We use this information to obtain a reduced form of the original problem that we can solve faster by solving two-stage stochastic (2SS) problems in a rolling horizon approach.

The decomposition algorithm comprises two steps. First, we decompose the problem into single scenarios dualizing the first-period and exogenous NACs. This step is equivalent to the first iteration of the Lagrangean decomposition where the NACs constraints are transferred to the objective function using penalty terms multiplied by Lagrange multipliers (Apap and Grossmann, 2017). We note that for the problem presented in Section 2, the conventional Lagrangean decomposition algorithm proved to be not efficient, leading to high computational times to close the gap between the lower and upper bound as described in Apap and Grossmann (2017). This can be caused by the relatively low number of scenarios, as already observed in Apap and Grossmann (2017), which led to the same oscillatory behavior reported by Uribe-Rodríguez et al. (2023) for the convergence.

For simplicity, let us consider only the decision variables y_t^s. Then, after dualizing the NACs, Eq.(*1*) becomes

$$\begin{aligned} \min_{y} \phi_{LR}(\lambda) = \sum_{s\in\mathcal{S}} pr^s \sum_{t\in\mathcal{T}} {}^{y}c_t^s y_t^s + \sum_{(s,s')\in\mathcal{SP}_F} {}^{F}\lambda_1^{s,s'}(y_1^s - y_1^{s'}) \\ + \sum_{(t,s,s')\in\mathcal{SP}_X} {}^{X}\lambda_t^{s,s'}(y_{t+1}^s - y_{t+1}^{s'}) \end{aligned} \tag{6}$$

We rearrange the equations of MSS1 for all the variables as in Eq.(*6*) to obtain a number of problems equal to the cardinality of $\mathcal{S}$.

At this point, we can explore the solutions of the scenarios to collect information about the technology deployment. In particular, we want to know which technologies are not selected in any scenario or the technologies selected in every scenario in all countries in a given time period for capacity expansion. We note that if the binary variable that indicates the capacity expansion of a given technology is zero, it does not imply that that technology is not deployed at all. We then use the information gathered in step 1 to reduce the size of the original problem by eliminating elements from the set of technologies, i.e., setting the binary variables to zero or fixing investment decisions. However, this problem, which we call MSS1-red, might still be very large. Therefore, we decompose it further using heuristics.

In the second step of our approach, we approximate the multistage stochastic problem with a series of two-stage stochastic problems (2SS) that we solve iteratively in a shrinking horizon approach, as described in Balasubramanian and Grossmann (2004). Moreover, in the first iteration, which corresponds to the first node over the entire time horizon, we solve for a subset of scenarios. The number of scenarios and the method for their selection can be chosen depending on the application. In our case studies, we use a heuristic approach dependent on the shape of the decision tree. Out of $|S|$, we select ten representative scenarios of the whole set.

An alternative approach, although computationally more expensive, is to use sample average approximation (SAA). In SAA, a sample of scenarios $\mathrm{N} < |S|$ that best represents the initial problem is selected to form a reduced set S_{red}. The probability of each scenario is adjusted to sum up to one by dividing the probability of each scenario in S_{red} by the sum of the probabilities of all scenarios in S_{red} (Ehrenstein et al., 2019). Since the probability distribution of the uncertainty is known a priori, here we choose N scenarios such that they are representative of the full set S. This set is used to solve an approximation of MSS1.

We then store the decision variables from each scenario (y_t^s) and fix the ones at the first time period (y_1^s) according to the values of the solution obtained in the first iteration. From the second iteration onward, we solve one problem that considers all the scenarios of the corresponding 2SS problems. At each iteration, we fix the decision variables at the beginning of the time period until no more time periods are to be fixed. Notably, all the integer variables are fixed when solving the leaf nodes. Therefore, the last 64 iterations are linear programming (LP) problems, which can be solved efficiently. Lastly, we compute the total expected cost as a probability-weighted sum of the costs at the final nodes.
At each iteration, a 5% optimality gap is enforced as termination criteria and convergence is checked before solving the next node.

To speed up the algorithm further, priorities on the discrete variables are also used. We expect that BECCS is deployed before DACCS, because of its lower cost, until there is not enough capacity to meet the CDR target. This translates into heuristic-based constraints involving binary variables that reduce the combinatorial complexity of the problem.

RAPIDU is solved using GAMS 41.5.0 on an Intel(R) Core(TM) i7-10700 CPU @ 2.90GHz with 32.0 GB RAM using 16 parallel threads with the solver CPLEX 22.1.

3.5 Value of the stochastic solution for multistage stochastic models

To quantitatively assess the additional value of adding uncertainty to the optimization problem, we compute the value of the stochastic solution (VSS) (Birge and Louveaux, 2011). Indeed, a decision-maker might argue that including recourse decisions in the model might not be worth the additional computational effort, to which the VSS provides insights into the potential value left on the table when not considering uncertainty in the decision-making process.

In the case of two-stage stochastic problems, the procedure to calculate VSS is straightforward and widely used. Firstly, the solution to the model with a mean value of the uncertain parameter is computed. This is called the expected value problem or mean value problem (EV). Once the solution to this problem is known, it is used to solve the stochastic model by fixing the first stage “here and now” variables. This is known as the expected result of using the EV solution.

For a minimization problem, the VSS is computed as the difference between EEV and RP (Eq. (7)), where RP is the solution to the fully stochastic problem. A small VSS denotes that the deterministic model is a good approximation of the stochastic one.

$$\mathrm{VSS} = \mathrm{EEV} - \mathrm{RP} \tag{7}$$

In the case of multistage problems, however, computing the VSS is not as simple because each time period has “here and now” decision variables and it is not clear which variables should be fixed. Therefore, obtaining EEV would require solving a sequence of models. The issues in computing the values above and an approach to calculate the VSS for multistage stochastic problems are discussed in Escudero et al. (2007).

In this work, we use an approximation of the procedure to calculate the VSS, also reported in Escudero et al. (2007), where we fix only the first stage decisions in all time periods to obtain EEV. Then, the VSS is calculated as the difference between EEV and RP as in Eq. (*(7)*).

4. Computational results and discussion

4.1 Homogeneous discretization of the time horizon

First, we analyze the case for time periods of equal lengths. The results are discussed in this section comparing the full space model to the decomposed version.

Hereafter, we refer to the full space model, i.e., the problem which includes all equations and variables and is solved at once, as MSS1, and the version solved by applying a decomposition algorithm as MSS1-D.

4.1.1 Net zero target

We impose a net-zero cumulative CDR target, i.e., the positive emissions are fully balanced by the CO_2 removed by BECCS and DACCS, at the end of the time horizon.

The number of variables, equations, and solution time of MSS1 are summarized in Table 1. Notably, the computational time to solve MSS1 in its full form to a 5% optimality gap is considerable (around 13 hours). The total cost of the system is 10.7 trillion Euros calculated as the weighted average of each scenario objective function value, and computed according to a net present value calculation, accounting for fixed and variable costs (see (Galán-Martín et al., 2021) for more information on the cost calculation). Compared to the deterministic solution (9.9 trillion Euros) it represents an increase of 7.2%, which is not negligible. In the context of the EU, this economic burden is borne by all the Members according to some fairness principle.

Table 1. Model statistic of the multistage stochastic problem MSS1 for the minimization of the expected cost.

	MIP - MSS 64 scenarios
Number variables [millions]	7.9
Number binary variables [thousands]	300.7
Number equations [millions]	4.2
Resource usage (solution + generation time) [hr]	~13
Solver	CPLEX 22.1
Termination criteria: opt. gap [%]	5
Optimal objective function value [trillion Eur]	10.7

Given the large computational time, we apply our decomposition approach and solve the problem again as MSS1-D.

From step 1, which is based on the first iteration of the Lagrangean decomposition, we find that many technologies are not selected for capacity expansion in any country at any time period. We group all these technologies in a set and we force their decision binary variables to be zero when solving step 2. In order to further reduce the size of the model, we identify technologies that are selected for capacity expansion in the first time period in specific countries in all time scenarios. Again, we define a set for these technologies and we impose that the decision variable in the first period is equal to one. Both sets of technologies are reported in Table 2.

Notably, many technologies are not selected because the time horizon is relatively short and the target is not very ambitious compared to what it is possible to achieve (Galán-Martín et al., 2021).

Table 2. Information on technology expansion from the Lagrangean decomposition (step 1) for a net-zero CDR target in 2050.

Technologies not selected in any scenario for capacity expansion at any time period t	**Technologies selected (country) in all the scenarios for capacity expansion at t=t_1**
•Coal •Coal with CCS •Hydropower •Hydropower reservoir •Nuclear •Concentrated solar power •Solar PV roof •Switchgrass •Wind offshore	•Geothermal (Germany) •Natural gas (Denmark) •Forest residue with CCS (Poland) •Woody residues with CCS (Greece) •Solar PV open (Luxembourg and Malta)

We make use of the information obtained in step 1 as described above and proceed with step 2. We solve step 2 iterating over the number of nodes (127). At the first iteration, we use heuristics to determine the subset of scenarios used to solve MSS1-red. Indeed, MSS1-red is still very large.

We obtain a 90% reduction in the computational time while achieving the same objective function value (10.7 trillion Euros). The latter is calculated as a probability-weighted sum of the leaf nodes, as explained in Section 3.3. We compare in Table 3 the solutions between the full space model and the decomposed one in terms of computational time and optimal objective function value. The time reported in the table includes step 1 and 2 where the Lagrangean decomposition step is approximately 14 minutes. The precise resource usage of generation and solution time in seconds is reported in Table A5.

Table 3. Comparison of full space and decomposed multistage stochastic models for the minimization of the expected cost. A reduction in computational time of 83% is achieved by implementing the decomposition algorithm. * probability-weighted sum at final nodes

	MIP - MSS 64 scenarios	Decomposed MIP - MSS
Resource usage (generation + solution time) [hr]	∼13	∼1.2
Optimal objective function value [trillion Eur]	10.7	10.7*

Despite reaching the same expected total cost while meeting the electricity demand, there are differences in the decisions taken during the time horizon. For example, in MSS1 forest residues and natural gas with CCS capacities are also chosen and expanded at the first time period (Table A1). Additionally, comparing the information in Table 2 and Table A1, we notice that the algorithm only selects a few countries j where a technology i is selected in all the scenarios. In other words, in the decomposed problem, it is preferred to increase the capacity installed in selected countries *j* by 100% or more, while in MSS-1 the capacity of more technologies is increased less but across different countries.

Lastly, we compare the stochastic results with the deterministic ones.
First, we fix the binary decision variables from the solution of the deterministic problem in all time periods of the stochastic problem to obtain EEV (Table 4). We find that the cost of EEV is 10.98 trillion Eur. Compared to the value of RP in Table 3, the VSS is 285 billion Eur, which is roughly 3% of the RP, on the same order of magnitude as the examples presented in Birge and Louveaux (2011) and Li and Grossmann (2021).

Table 4. Objective function value for different case studies. Given the solution of the deterministic problem, we find the EEV value to later compute the VSS.

Deterministic variables to be fixed	Time period	Value of objective function [B Eur]
Binary decisions	Up to t_5	10791.2
Binary decisions	Up to t_6	10982.6
Binary and capacity decisions	Up to t_5	Infeasible
Binary and capacity decisions	Up to t_6	Infeasible

Then, we provide a graphical summary of the two sets of solutions. Let us assume that the deterministic model is run only one time at the beginning of the time horizon, although in practice it would be run in a rolling horizon fashion. Then, the solution that we obtain can be outside of the range of all the possible solutions obtained considering uncertainty.

In Figure 5 we show the electricity generated per country (subplot A) and per technology type (subplot B). We notice that the majority of the total electricity generated comes from a reduced set of countries (subplot A), i.e., France, Germany, Italy, Spain, and the United Kingdom, and also affected by the greatest variability among all the countries. Additionally, the deterministic solution lies outside of the range observed for the given scenarios in some countries such as the Netherlands, where the point lies outside of the boxplot of stochastic solutions.
Subplot B represents the same information, this time aggregated by technology. Here it is even more evident that the deterministic solution is not included within the stochastic values, in particular, in the case of wind onshore and solar PV open. Wind onshore and nuclear follow natural gas in production volume. Among all the technologies, natural gas is the main source of electricity and has also the highest variability while the capacity of wind offshore does not change across the scenarios.

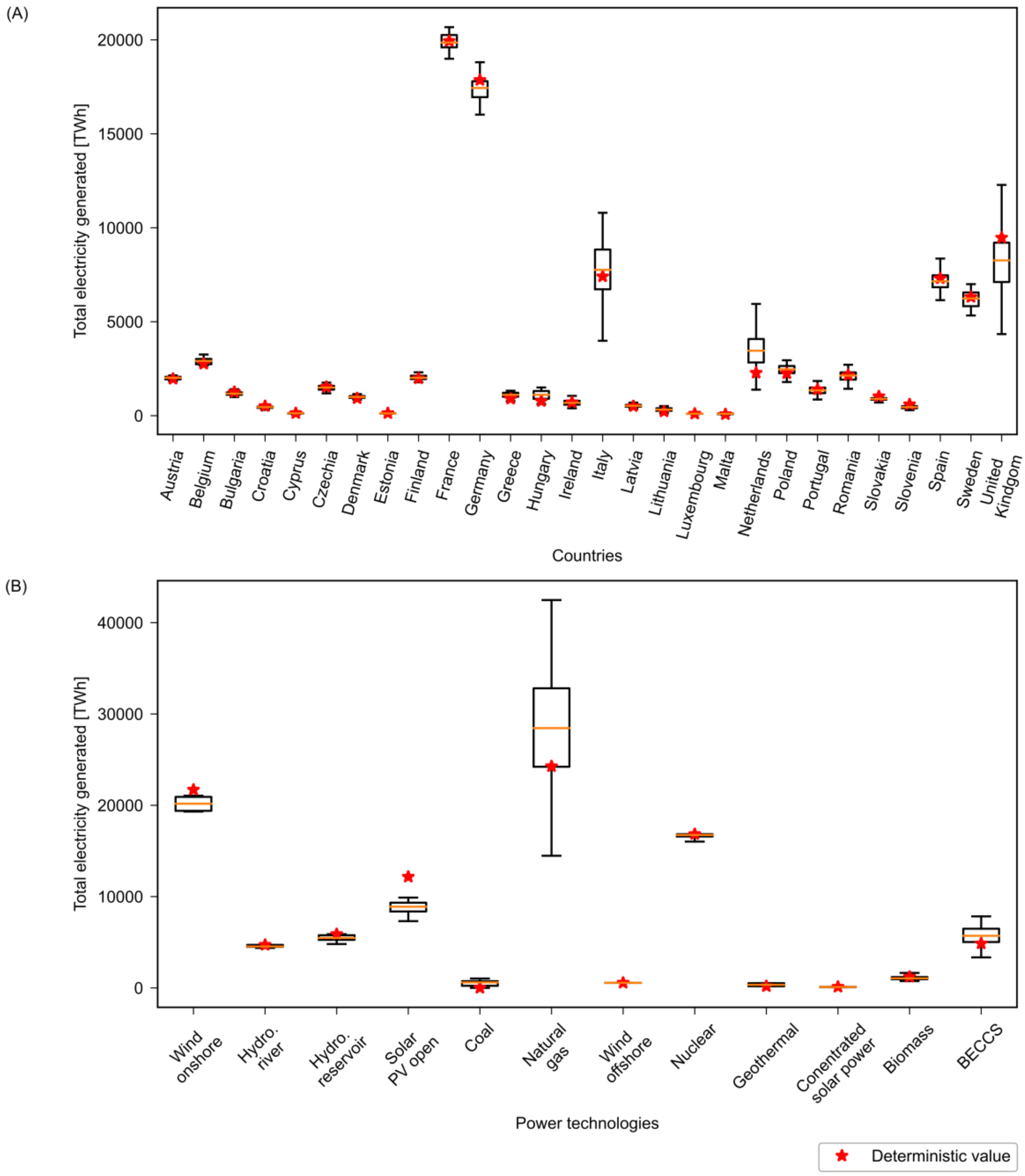

Figure 5. Total electricity generated aggregated by country (subplot A) and technology (subplot B). Stochastic range vs deterministic value. The box and whiskers plots are generated using the solution of all 64 scenarios of the stochastic model and they show ±25 the median value.

4.1.2 CDR scenarios

The decomposition algorithm presented in Section 3.3 allows us to more easily explore a range of case studies. We can do so by imposing a different CDR target at each iteration. In particular, our interest lies in assessing the feasibility of a carbon balance beyond the net zero target and proving that it is possible to attain such levels by mid-century. This would imply that the electricity sector would be responsible for offsetting emissions from hard-to-abate industries, such as cement and steel. Considering that the electricity sector accounted for approximately 30% of the total CO_2 emissions in the EU in 2021 (Statista, 2021), it seems reasonable for it to contribute significantly to the overall reduction in CO_2 emissions.

Our findings demonstrate that the deployment of BECCS and DACCS can potentially enable the removal of up to 9 Gt of CO_2 by 2050 by deploying BECCS and DACCS within the constraints of our model. In other words, the global carbon balance of the integrated sector, i.e., including in the system boundaries the power technologies in addition to NETPs, would result in the net removal of 9 Gt. Figure 6 illustrates the expected total cost of implementing solutions for the various CDR scenarios explored. A negative sign of emissions indicates that more CO_2 is removed via BECCS and DACCS than the amount emitted from all the technologies, i.e., natural gas, solar, and wind, resulting in negative emissions. We observe that the steepest increase in total cost occurs when we push the system to achieve more ambitious targets, ranging from 0 to –9 Gt of CO_2, while the cost does not decrease significantly for different positive targets, where the deployment of NETPs is minimal. Additionally, we find that DACCS, powered by electricity and heating, is only selected and deployed starting at the –9 Gt CO_2 target.

It is important to note that even in scenarios where the emissions balance is positive, there is still a deployment of bioenergy and BECCS (Table A2). This is because the advantage of bioenergy is twofold: it removes CO_2 from the atmosphere while simultaneously generating electricity. Thus, it serves as a valuable technology for achieving emissions reductions. This matches previous observations made by the authors (Negri et al., 2021) where even considering the impacts of a highly detailed BECCS supply chain, negative emissions were achieved in the EU.

Our analysis reveals that achieving a target beyond 10 Gt CO_2 removal is not feasible (considering the assumptions and technologies in our analysis) within the computational time limit allowed (12 hours). This limitation is primarily driven by the rate of technology deployment rather than the availability of CO_2 storage. The high electricity demand in the most uncertain scenarios poses a significant challenge in achieving higher levels of CO_2 removal within the given timeframe.

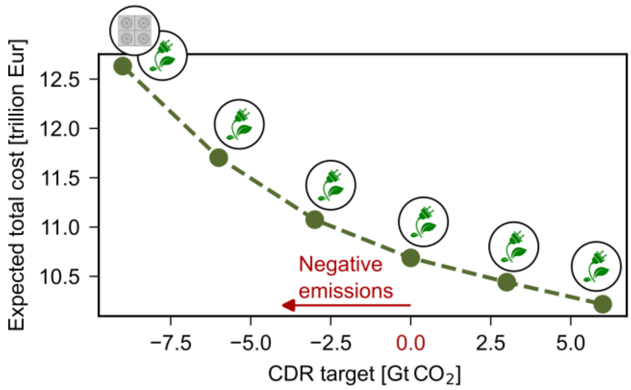

Figure 6. Expected total cost and technology deployed for different CDR targets. The icons of BECCS (green plant) and DACCS (gray box) indicate which CDR option is proving the required removal. Only in the case of 9 $GtCO_2$ removal, both technologies are deployed.

We report in Appendix A.1 the information relative to the technology selection in the decomposition algorithm step 1 and the computational time. We see that as the target becomes less ambitious, and eventually the system is even allowed to reach a positive emissions balance, fewer technologies are installed.

4.2 Inhomogeneous discretization of the time horizon with net-zero target

As a last step, we are interested in the implications of discretizing the time horizon in time periods of different lengths (see Figure 7 A and B). This modeling decision is driven by the fact that the consequences of the years closer to the start of the horizon have greater economic and social impacts over the ones of the years further in time.

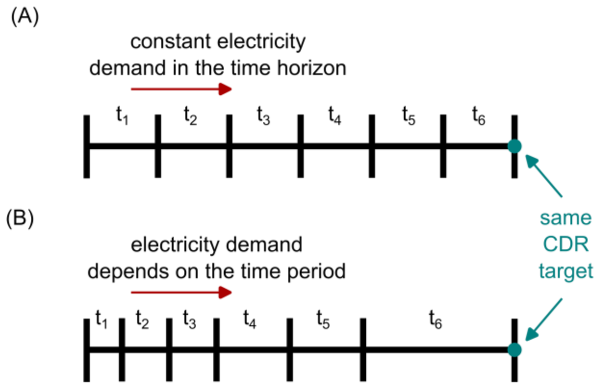

Figure 7. Discretization of the time periods. In the analysis presented in Section 4.1 (subplot A), the time horizon is divided into six time periods of equal length corresponding to five years. In this Section 4.2 we present the results for a non-homogenous discretization of the time horizon: six total time periods of length 2-4-4-5-5-10 years.

We report the results of the inhomogeneous time horizon discretization in Table 5 for the minimization of the expected total cost at a net zero target. First, we highlight that in the full space problem we were able to reach a solution within the specified optimality gap although infeasibilities on the original MIP problem were not resolved. On the other hand, our decomposition algorithm can reach a solution within the given tolerance without any numerical challenges highlighting the robustness of the proposed solution method. The total expected cost, calculated as the probability-weighted average of the last nodes, is 2% higher than the full space objective function, within the optimality gap. Lastly, we mention that the total expected cost in this case study is higher than the previous (10.7 trillion Eur, see Table 3) due to the duration of the periods being different.

The underlying assumption in the original model (Galán-Martín et al., 2021) is that the demand is maintained constant within the years of the first time period, and only updated at the start of the following period according to the following equation $d_t = d_{t-1} * \mathrm{F}$, where F is a constant factor greater than one. Therefore, it is updated more frequently when there are shorter periods. For example, since in the case of identical time periods the first time period comprises five years while here only two, the demand at the end of the time horizon is higher in the case presented in this section. We provide a clarifying example in the time horizon 2020 – 2025 with reference to Figure 7. In case A, it is assumed that there is no increase in the first five years, meaning that the cumulative final demand is equal to five times the one in 2020. In the inhomogeneous time period case (B), the first time period of two years follows the same assumption and therefore the cumulative demand in 2022 is two times the one in 2020, while the remaining years are updated according to the given correlation.

Table 5. Comparison of full space and decomposed multistage stochastic models with inhomogeneous time horizon discretization for the minimization of the expected cost to meet the net zero target in 2050. Reduction in computational time given by the decomposition algorithm: 43%. * probability-weighted sum at final nodes. ** GAMS output: Fixed MIP status (5): *optimal with unscaled infeasibilities*

	MIP - MSS 64 scenarios	Decomposed MIP - MSS
Number variables [Millions]	7.9	7.9
Number binary variables [Thousands]	300.7	300.7
Number equations [Millions]	4.2	4.2
Resource usage (generation + solution time) [hr]	~3.2	~2
Solver	Cplex 22.1	Cplex 22.1
Termination criteria: Opt. gap [%]	5	5
Objective function value [Trillion Eur]	12.3**	12.5*

Next, we look into the technology deployment to meet the energy demand (Table A3 and Table A4). In this case, since the first time period is shorter, a reduced set of technologies is selected in all the scenarios, comprising wind onshore (Germany), solar PV open (Luxemburg), and forest residues with CCS (Poland). Additional capacity of solar PV open is installed in Belgium at t_3 and t_5. The capacity of solar PV open and forest residues with CCS is expanded at the beginning of t_4 in 50% of the countries included in our model. This happens because the length of time periods t_1 and t_2 are 2 and 4 years, respectively, while in p_4 it is five years and more capacity is needed to cover the increase in electricity demand. Overall, the net zero target is achieved by deploying the same set of technologies shown in Figure 5 B.

5. Conclusions

In this study, we have highlighted the importance of considering uncertainty in energy system planning, particularly in relation to the implications on the total cost and technology capacity installed arising from the variability in electricity demand. The consequent complexity of the multistage stochastic programming model required the development of a tailored decomposition approach, which reduces the computational time significantly up to 90%. We observed that in the case of inhomogeneous time periods, which lead to numerical challenges to solve the full-space model, our algorithm provides a numerically robust solution within the optimality tolerance while still providing a significant speedup with respect to the monolithic model.

We found that among all the countries in the European Union, France, Germany, Italy, Spain, and the United Kingdom, provide the majority of the electricity with the greatest variability across all the scenarios solutions. The technologies with the highest capacity deployed for electricity production include natural gas, nuclear, wind onshore, solar PV and BECCS. Biomass emerged as a crucial component within the model, where its deployment was not only selected to achieve negative emission targets but also in scenarios where overshooting was allowed with capacity expansion at each time period for selected feedstocks.

Overall, effectively managing uncertainties is paramount for the successful implementation of projects, particularly in the context of climate action. Further research and attention to these uncertainties will contribute to the advancement and realization of sustainable energy systems, especially focusing on the learning curves of carbon removal options. While our analysis focused on the technical aspects, it is important to acknowledge that successful implementation of BECCS and DACCS technologies requires addressing additional considerations such as public acceptance, community engagement, policy incentives, and economic viability, which are beyond the scope of this study.

CRediT authorship contribution statement

Valentina Negri: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Daniel Vázquez**: Methodology, Software, Writing – original draft, Writing – review & editing. **Ignacio Grossmann**: Conceptualization, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. **Gonzalo Guillén-Gosálbez**: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Apap, R.M., Grossmann, I.E., 2017. Models and computational strategies for multistage stochastic programming under endogenous and exogenous uncertainties. Computers & Chemical Engineering 103, 233–274. https://doi.org/10.1016/J.COMPCHEMENG.2016.11.011

Balasubramanian, J., Grossmann, I.E., 2004. Approximation to Multistage Stochastic Optimization in Multiperiod Batch Plant Scheduling under Demand Uncertainty. Industrial & Engineering Chemistry Research 43, 3695–3713. https://doi.org/10.1021/IE030308

Barbaro, A., Bagajewicz, M.J., 2004. Managing Financial Risk in Planning under Uncertainty. AIChE Journal 50, 963–989. https://doi.org/10.1002/aic.10094

Birge, J.R., Louveaux, F., 2011. Introduction to Stochastic Programming. Springer Science & Business Media.

Bistline, J.E.T., Blanford, G.J., 2021. Impact of carbon dioxide removal technologies on deep decarbonization of the electric power sector. Nature Communications 2021 12:1 12, 1–12. https://doi.org/10.1038/s41467-021-23554-6

Cobo, S., Negri, V., Valente, A., Reiner, D.M., Hamelin, L., Mac Dowell, N., Guillén-Gosálbez, G., 2023. Sustainable scale-up of negative emissions technologies and practices: where to focus. Environ. Res. Lett. 18, 023001. https://doi.org/10.1088/1748-9326/acacb3

Creutzig, F., Breyer, C., Hilaire, J., Minx, J., Peters, G.P., Socolow, R., 2019. The mutual dependence of negative emission technologies and energy systems. Energy and Environmental Science 12, 1805–1817. https://doi.org/10.1039/c8ee03682a

Daggash, H.A., Heuberger, C.F., Mac Dowell, N., 2019. The role and value of negative emissions technologies in decarbonising the UK energy system. International Journal of Greenhouse Gas Control 81, 181–198. https://doi.org/10.1016/j.ijggc.2018.12.019

Daggash, H.A., Mac Dowell, N., 2019. Higher Carbon Prices on Emissions Alone Will Not Deliver the Paris Agreement. Joule 3, 2120–2133. https://doi.org/10.1016/j.joule.2019.08.008

Doukas, H., Nikas, A., González-Eguino, M., Arto, I., Anger-Kraavi, A., 2018. From Integrated to Integrative: Delivering on the Paris Agreement. Sustainability 10, 2299. https://doi.org/10.3390/su10072299

Ehrenstein, M., Wang, C.H., Guillén-Gosálbez, G., 2019. Strategic planning of supply chains considering extreme events: Novel heuristic and application to the petrochemical industry. Computers & Chemical Engineering 125, 306–323. https://doi.org/10.1016/J.COMPCHEMENG.2019.03.020

Escudero, L.F., Garín, A., Merino, M., Pérez, G., 2007. The value of the stochastic solution in multistage problems. TOP 15, 48–64. https://doi.org/10.1007/s11750-007-0005-4

European Commission, 2019. The European Green Deal. European Commission.

Fajardy, M., Patrizio, P., Daggash, H.A., Mac Dowell, N., 2019. Negative Emissions: Priorities for Research and Policy Design. Frontiers in Climate 1, 6. https://doi.org/10.3389/fclim.2019.00006

Fusco, A., Gioffrè, D., Francesco Castelli, A., Bovo, C., Martelli, E., 2023. A multi-stage stochastic programming model for the unit commitment of conventional and virtual power plants bidding in the day-ahead and ancillary services markets. Applied Energy 336, 120739. https://doi.org/10.1016/j.apenergy.2023.120739

Galán-Martín, Á., Vázquez, D., Cobo, S., Dowell, N.M., Caballero, J.A., Guillén-Gosálbez, G., 2021. Delaying carbon dioxide removal in the European Union puts climate targets at risk. Nature Communications 12, 6490. https://doi.org/10.1038/s41467-021-26680-3
Grant, N., Hawkes, A., Mittal, S., Gambhir, A., 2021. The policy implications of an uncertain carbon dioxide removal potential. Joule 5, 2593–2605. https://doi.org/10.1016/J.JOULE.2021.09.004
Grossmann, I.E., Apap, R.M., Calfa, B.A., García-Herreros, P., Zhang, Q., 2016. Recent advances in mathematical programming techniques for the optimization of process systems under uncertainty. Computers & Chemical Engineering, 12th International Symposium on Process Systems Engineering & 25th European Symposium of Computer Aided Process Engineering (PSE-2015/ESCAPE-25), 31 May - 4 June 2015, Copenhagen, Denmark 91, 3–14. https://doi.org/10.1016/j.compchemeng.2016.03.002
Guillén-Gosálbez, G., Mele, F.D., Grossmann, I.E., 2010. A bi-criterion optimization approach for the design and planning of hydrogen supply chains for vehicle use. AIChE Journal 56, 650–667. https://doi.org/10.1002/AIC.12024
Gupta, V., Grossmann, I.E., 2011. Solution strategies for multistage stochastic programming with endogenous uncertainties. Computers & Chemical Engineering 35, 2235–2247. https://doi.org/10.1016/J.COMPCHEMENG.2010.11.013
Hasan, M.M.F., Zantye, M.S., Kazi, M.K., 2022. Challenges and opportunities in carbon capture, utilization and storage: A process systems engineering perspective. Computers & Chemical Engineering 166, 107925. https://doi.org/10.1016/j.compchemeng.2022.107925
IPCC, 2022. Climate Change 2022. Mitigation of Climate Change. Working Group III contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC.
Lappas, N.H., Gounaris, C.E., 2016. Multi-stage adjustable robust optimization for process scheduling under uncertainty. AIChE Journal 62, 1646–1667. https://doi.org/10.1002/aic.15183
Lee, J., Bae, S., Kim, W.C., Lee, Y., 2023. Value function gradient learning for large-scale multistage stochastic programming problems. European Journal of Operational Research 308, 321–335. https://doi.org/10.1016/j.ejor.2022.10.011
Li, C., Grossmann, I.E., 2021. A Review of Stochastic Programming Methods for Optimization of Process Systems Under Uncertainty. Frontiers in Chemical Engineering 0, 34. https://doi.org/10.3389/FCENG.2020.622241
Li, P., Arellano-Garcia, H., Wozny, G., 2008. Chance constrained programming approach to process optimization under uncertainty. Computers and Chemical Engineering 32, 25–45. https://doi.org/10.1016/j.compchemeng.2007.05.009
Meckling, J., Sterner, T., Wagner, G., 2017. Policy sequencing toward decarbonization. Nat Energy 2, 918–922. https://doi.org/10.1038/s41560-017-0025-8
Meersman, T., Maenhout, B., Van Herck, K., 2023. A nested Benders decomposition-based algorithm to solve the three-stage stochastic optimisation problem modeling population-based breast cancer screening. European Journal of Operational Research 310, 1273–1293. https://doi.org/10.1016/j.ejor.2023.04.027
Negri, V., Galán-Martín, Á., Pozo, C., Fajardy, M., Reiner, D.M., Mac Dowell, N., Guillén-Gosálbez, G., 2021. Life cycle optimization of BECCS supply chains in the European Union. Applied Energy 298, 117252. https://doi.org/10.1016/J.APENERGY.2021.117252

Oliveira, F., Gupta, V., Hamacher, S., Grossmann, I.E., 2013. A Lagrangean decomposition approach for oil supply chain investment planning under uncertainty with risk considerations. Computers & Chemical Engineering 50, 184–195. https://doi.org/10.1016/J.COMPCHEMENG.2012.10.012

Rathi, T., Zhang, Q., 2022. Capacity planning with uncertain endogenous technology learning. Computers & Chemical Engineering 164, 107868. https://doi.org/10.1016/J.COMPCHEMENG.2022.107868

Realmonte, G., Drouet, L., Gambhir, A., Glynn, J., Hawkes, A., Köberle, A.C., Tavoni, M., 2019. An inter-model assessment of the role of direct air capture in deep mitigation pathways. Nat Commun 10, 3277. https://doi.org/10.1038/s41467-019-10842-5

Ritchie, H., Roser, M., Rosado, P., 2020. CO_2 and Greenhouse Gas Emissions. Our World in Data.

Ruszczyński, A., 1997. Decomposition methods in stochastic programming. Mathematical Programming 79, 333–353. https://doi.org/10.1007/BF02614323

Sagues, W.J., Park, S., Jameel, H., Sanchez, D.L., 2019. Enhanced carbon dioxide removal from coupled direct air capture-bioenergy systems. Sustainable Energy and Fuels 3, 3135–3146. https://doi.org/10.1039/c9se00384c

Sahinidis, N.V., Grossmann, I.E., Fornari, R.E., Chathrathi, M., 1989. Optimization model for long range planning in the chemical industry. Computers & Chemical Engineering 13, 1049–1063. https://doi.org/10.1016/0098-1354(89)87046-2

Schenuit, F., Colvin, R., Fridahl, M., McMullin, B., Reisinger, A., Sanchez, D.L., Smith, S.M., Torvanger, A., Wreford, A., Geden, O., 2021. Carbon Dioxide Removal Policy in the Making: Assessing Developments in 9 OECD Cases. Frontiers in Climate 3.

Shapiro, A., 2021. Tutorial on risk neutral, distributionally robust and risk averse multistage stochastic programming. European Journal of Operational Research 288, 1–13. https://doi.org/10.1016/j.ejor.2020.03.065

Solano Rodriguez, B., Drummond, P., Ekins, P., 2017. Decarbonizing the EU energy system by 2050: an important role for BECCS. Climate Policy 17, S93–S110. https://doi.org/10.1080/14693062.2016.1242058

Statista, 2021. EU-27: energy sector GHG emission shares [WWW Document]. Statista. URL https://www.statista.com/statistics/1000061/ghg-emissions-sources-energy-sector-european-union-eu/ (accessed 7.5.23).

Uribe-Rodríguez, A., Castro, P.M., Guillén-Gosálbez, G., Chachuat, B., 2023. Assessment of Lagrangean decomposition for short-term planning of integrated refinery-petrochemical operations. Computers & Chemical Engineering 174, 108229. https://doi.org/10.1016/j.compchemeng.2023.108229

Vespucci, M.T., Bertocchi, M., Zigrino, S., Escudero, L.F., 2013. Stochastic optimization models for power generation capacity expansion with risk management, in: International Conference on the European Energy Market. IEEE. https://doi.org/10.1109/EEM.2013.6607352

Victoria, M., Zhu, K., Brown, T., Andresen, G.B., Greiner, M., 2020. Early decarbonisation of the European energy system pays off. Nat Commun 11, 6223. https://doi.org/10.1038/s41467-020-20015-4

Appendix A

We report in this appendix additional results and analyses related to the studies presented in the main manuscript. In Section A1 the insights obtained from the decomposition algorithm are given, while in Section A2 we discuss the risk-averse model.

A.1 Additional results

The following tables include detailed information referring to time period p_1 on the solution obtained from the multistage stochastic model in its full-space form (Table A1) and the decomposed one for different CDR targets (Table A2). Table A3 includes the technologies not selected at any time period and Table A4 the technologies selected for capacity expansion at p_1 for the case study with inhomogeneous time periods.

Table A1. Technology expansion and respective location at t1 in the solution of MSS-1.

Technology	Country
Forest residues	Estonia, Finland
Forest residues with CCS	Bulgaria, Croatia, Germany, Hungary, Luxembourg, Poland, Sweden
Geothermal	Germany, Portugal
Natural gas	Denmark, Germany, Luxembourg, Poland
Natural gas with CCS	Denmark
Solar PV open	Hungary, Luxembourg, Malta, Poland, Slovakia

Table A2. Technology information obtained from the decomposition algorithm step 1, and computational time of step 1.

CDR target [$GtCO_2$]	Computational time [min]	Technologies not selected in any scenario for capacity expansion at any time period	Technologies selected in all the scenarios in (country) for capacity expansion at t1
–9	21	•Coal •Hydropower •Nuclear •Solar PV roof •Switchgrass •Switchgrass with CCS	•Forest residues with CCS (Poland) •Solar PV open (Luxembourg)
–6	16	•Wind offshore •Hydropower •Hydropower reservoir •Solar PV roof •Coal •Nuclear •Switchgrass	•Solar PV open(Luxembourg, Malta) •Natural gas (Luxembourg) •Forest residues with CCS (Poland)
–3	15	•Wind offshore •Hydropower •Hydropower reservoir	•Geothermal (Germany) •Natural gas (Denmark, Luxembourg) •Solar PV open(Luxembourg, Malta)

		•Solar PV roof •Concentrated solar power •Coal •Nuclear •Coal with CCS •Switchgrass	•Forest residues with CCS (Poland)
+3	14	•Wind offshore •Hydropower •Hydropower reservoir •Solar PV roof •Concentrated solar power •Coal •Nuclear •Coal with CCS •Natural gas with CCS •Miscanthus •Switchgrass •Miscanthus with CCS •Switchgrass with CCS	•Geothermal (Germany) •Natural gas (Denmark) •Woody residues with CCS (Greece) •Solar PV open(Luxembourg) •Forest residues with CCS (Poland)
+6	14	•Wind offshore •Hydropower •Hydropower reservoir •Solar PV roof •Concentrated solar power •Coal •Nuclear •Coal with CCS •Natural gas with CCS •Miscanthus •Switchgrass •Miscanthus with CCS •Switchgrass with CCS	•Geothermal (Germany) •Woody residues with CCS (Greece) •Solar PV open(Luxembourg) •Forest residues with CCS (Poland)

Table A3. Technologies not selected for capacity expansion in any county, any scenario and any time period for the minimization of the total cost with inhomogeneous time periods under a net-zero target.

Technologies not selected
Wind offshore
Hydropower
Hydropower reservoir
Solar PV roof
Concentrated solar power
Coal
Nuclear
Coal with CCS
Natural gas with CCS
Miscanthus
Switchgrass
Willow
Straw residues
Woody residues
Forest residues
Miscanthus with CCS
Switchgrass with CCS

Table A4. Technology selection at time period p_1 in step 1 using a time horizon with inhomogeneous time periods. No expansion occurs at p_6.

Time period	Technology expanded	Country
t_1	Wind onshore	Germany
	Solar	Luxembourg
	Forest residues with CCS	Poland
t_2		
t_3	Solar PV open	Belgium
t_4	Solar PV open	Austria, Bulgaria, Spain, Estonia, Finland, Hungary, Italy, Lithuania, Luxembourg, Poland, Portugal, Slovakia
	Forest residues with CCS	Denmark, Sweden
t_5	Solar PV open	Belgium
t_6		

Lastly, we report in Table A5 the precise resource usage time in seconds with reference to Table 1, Table 3 and Table 5.

Table A5. Resource usage of model generation and solution time in seconds for the case studies investigated, i.e., with homogeneous and inhomogeneous time periods. The resource usage reported does not include

Case study/ resource usage time [s]	Full-space	Decomposed
Homogeneous time horizon discretization	46597	4444
Inhomogeneous time horizon discretization	11490	7815

A.2 Risk management

Among the possible frameworks to include uncertainty in optimization problems, stochastic programming is a risk-neutral approach, because it optimizes the expectation of the objective function by neglecting that some of the scenarios might incur high costs (Oliveira et al., 2013).

Risk management is most widely explored in two-stage stochastic problems (Barbaro and Bagajewicz, 2004; Oliveira et al., 2013; Vespucci et al., 2013), although applications to multistage have also been explored (Shapiro, 2021). The reason for richer literature in two-stage stochastic models with risk metrics is that including in the model a term that represents the risk leads to a considerable increase in the model complexity. Different metrics to manage risk have been defined in the literature: downside risk, value at risk, and conditional value at risk.

Following Oliveira et al. (2013), we calculate the expected shortage risk (ES). The latter has the advantage of adding only one constraint and one continuous variable in the model. Nonetheless, given the complexity of our problem, we decide to explore the risk-averse case by replacing our model with a two-stage stochastic model where we include the formulation of ES. The mathematical formulation is given in Eq. (A*8*) and (A*9*), which adopts the same nomenclature as Oliveira et al. (2013).

$$\min_{x,y,\delta}\left(cx+\sum_u P^u q y^u+\mathrm{PEN}\sum_u P^u\delta^u\right) \tag{A8}$$

$$\begin{gathered} s.t.\, Ax+b\le 0 \\ Tx+Wy^u\le h^u \quad \forall\, u\in U \\ y^u\in Y \\ cx+qy^u-\omega\ \le \delta^u \\ \delta^u\ge 0 \quad \forall\, u\in U \end{gathered} \tag{A9}$$

ES is calculated as in Eq.(A*10*).

$$ES(\omega,\, x)=\frac{1}{\sum_{u|\delta^u\ge 0}P^u}\ \sum_u P^u\delta^u \tag{A10}$$

We observe that a small trade-off of cost vs. risk can be observed because the model is highly constrained. Indeed, the electricity demand has to be met as an equality constraint. In the case of adding a slack variable, increasing the demand uncertainty, e.g., ±50% instead of 20%, and a longer time horizon would allow us to observe more significant trade-offs. However, this would lead to a computationally intractable model with the evaluated solution method.