



---

*Research article*

## Lagrangian decomposition for stochastic TIMES energy system optimization model

Yujun Zhu<sup>1,2</sup> and Ju Ming<sup>1,\*</sup>

<sup>1</sup> School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, China

<sup>2</sup> School of Mathematics, University of Edinburgh, Edinburgh, U.K.

\* **Correspondence:** Email: [jming@hust.edu.cn](mailto:jming@hust.edu.cn); Tel: +8602787542763.

**Abstract:** Energy system optimization models play an essential role in current decision support on topics including energy security, sustainable development and environmental protection from industrial, regional, national and even global perspective. One of the key energy system optimization models applied in international energy policy is TIMES. The article establishes two basic deterministic TIMES models which cover an energy commodity (coal or gas), a three-step supply curve and an end-use energy service demand. Then we convert the deterministic TIMES models into a stochastic optimization problem with multiple scenarios, and implement the Lagrangian decomposition approach in solving the stochastic programming models. The numerical experiment shows the feasibility of the Lagrangian decomposition algorithm to solve stochastic TIMES models with a small amount of scenarios, and analyze several reasons for non-convergence cases including the choice of step length and initial values of Lagrangian multipliers.

**Keywords:** the TIMES model; stochastic extension; two-stage stochastic programs with recourse; scenario decomposition; Lagrangian relaxation; supergradient ascent method

**Mathematics Subject Classification:** 49N15, 90C15

---

### 1. Introduction

Climate change resulting from greenhouse gas emissions is a key issue for modern society. Energy system optimization models (ESOMs), in which we construct and analyze the mathematical optimization formulation of the real-life complex energy systems, play a critical role in government decision support on industry [47], environment [30], energy [38] and economy [46]. ESOMs are intended to provide potential pathways to decarbonization, including studying the trade-offs among sustainability, affordability, and security of supply. Therefore, it is essential to consider an appropriate

level of detail in energy system modelling studies with the aim of better prediction about relevant features of the real world, which gave us an motivation to develop a stochastic extension of the current ESOMs.

According to Jia and Liu [18], current research on energy and environmental policies can be divided into “top-down” energy economic models and “bottom-up” energy technology models according to research methods and model frameworks. The essence of “top-down” model is an economic model which uses the energy prices and economic elasticity as the main economic inputs. The “top-down” models reveal the relationship between energy consumption and energy production [7]. One of the well-known representatives is the Computable General Equilibrium (CGE) model, which focuses on economic effects of the energy system and detects the supply chain including providers of raw resources and end-use consumers to maximize utility. It is widely used in energy, environmental and economic policies, especially research on energy tax [41] and environmental tax [13].

CGE energy system mainly consists of the following three aspects: Resource supply, energy demand and the equilibrium relationship between supply and demand. The total optimal objective value in CGE model is obtained by optimizing over each one of the entire system in parallel and ignores the connection among these departments. Moreover, CGE model focuses more on the economical performance of the energy decision system and less on the technical analysis inside the system. Therefore, CGE model is generally applied in energy decision support from the economical perspective [37].

The “bottom-up” energy technology mainly refers to the large-scaled dynamic energy optimization models developed by international research institutions, which are currently widely applied in energy system optimization and pollution emission issues from national, regional and industrial perspective [28]. The Integrated MARKAL-EFOM System (TIMES) model is a well-known representative of dynamic linear energy optimization models. TIMES is a technology-based bottom-up model generator for regional, national or even global energy systems, which is widely utilized to derive optimal energy, economic and environmental policies and scenarios over a medium-long time horizon [28]. Based on the inputs of users, TIMES can generate multi-period dynamic linear optimization models for multi-regions systems. The model was developed by the Energy Technology Systems Analysis Programme in the 1980s, and the TIMES family of modelling tools has been applied by approximately 177 institutions in 69 countries so far [43].

Basically TIMES is the integrated evolution of the MARKAL [40] and the EFOM [10]. The former model, MARKAL is a typical least-cost optimization model, which simulates the whole process of the energy system covering energy extraction, processing conversion, and terminal utilization. And it is driven by the demand for end-use energy services and pursues for the lowest total system cost, which provides a basic and solid modelling structure for the TIMES. In addition, the EFOM presents a detailed representation of energy flows among technologies [28]. As an integration, the TIMES model not only takes advantages from both sides, but has an analysis ability over multi-regions and multi-periods and more flexibility than its forerunners, for example, variable length periods and endogenous trade between regions [9]. Generally, the TIMES model is applied to generate vertically integrated models of the entire extended energy system, as well as to capture the main characteristics of the energy system [42]. Due to the characteristics, the TIMES model is very popular in decision support for complex energy systems with the respect to economy, environment and energy [43].

TIMES model focuses more on the technological aspect and provides a detailed description and

simulation of the energy technologies in the system [48]. It aims to minimize the total cost of the entire energy system. More parameters concerning economy, technology and climate are taken into consideration in TIMES [48]. TIMES also include numerous constraints on different sectors in the system involving resource extraction, processing, conversion, distribution and end-use demands in the energy system. Thus the main characteristics of a large-scaled energy system can be well illustrated in TIMES, as well as the complex internal connections among sectors of the energy system.

Compared with CGE energy optimization model, TIMES performs more comprehensive analysis, calculation and optimization towards several sectors in energy system [18]. However TIMES model does not perform so well as CGE on the reflection of the mutual influence between the energy system and the regional economy. The two kinds of models have different focuses, CGE attaches more importance to the economical effect while TIMES pays attention to the mutual effect between sectors and technological processes in the energy system. Our research mainly focuses on the quantitative analysis of three sections, resource supply, energy conversion and end-use demand, in the energy system with uncertainty from the technical level. Thus TIMES modelling structure is a good choice in our energy decision planning.

There have been many countries and institutions applying the TIMES model in their energy policy making from continental, national and regional perspective. Will McDowall et al. [31] assessed the impact of renewable resources and indirect emissions on European cost-optimal decarbonization pathway based on UCL's European TIMES model (ETM-UCL). Olexandr Balyk et al. [3] developed TIMES-DK for decision support in guiding Denmark towards a carbon neutral future. To achieve UK's 2050 greenhouse gas reduction targets, the UK TIMES Model (UKTM) has been widely applied in British energy strategic analysis including the 5th Carbon Budget Report [43] and residential heating decarbonization plan [25]. One of the key ESOMs used in Scottish, Westminster and international energy policy is TIMES. Scottish government has also developed policies and proposals based on the Scottish TIMES model in pursuit of decarbonization including the draft Climate Change Plan [42, 44]. Senatro Di Leo et al. [9] constructed a regional TIMES model for Basilicata which is in the south area of Italy to better realize sustainable objectives in energy planning system. J. Liu et al. [26] established C-TMS (China TIMES model system) to make policies for future low carbon development in China.

These researchers devise different scenarios, constraints and elasticity factors for TIMES to achieve the goal of sustainable development from regional, multi-regional and national perspective with perfect foresight assumptions and deterministic datasets. However, there are large uncertainties and volatility in future international energy market, climate, technology innovation and policies. Therefore, there is further work in the TIMES development—stochastic programming, and one of the popular research strands in TIMES has been to study the effect of technological innovation and diffusion in the energy sector under uncertainty [43]. Researchers have incorporate uncertainties into energy system models and made optimal decisions under stochastic values of linear programming coefficients. Kanudia and Loulou [20] first constructed a stochastic extended framework of the two-stage MARKAL model for Québec, where optimal decisions are obtained under multiple future scenarios. The stochastic mixed integer programming was then introduced in MARKAL model by Kanala and Fragniere [19]. Later the TIMES framework family was also enlarged by members to employ stochastic programming for uncertainty management in [17, 18, 32, 40].

Two-stage stochastic energy systems play an important role in TIMES family of modelling

frameworks for industrial applications [32, 40] and robustness analysis [20, 24]. Thus we mainly focus on two-step stochastic TIMES energy systems in the current stage and might extend to multi-step TIMES models in the future. In the article, we aim to extend current two-stage TIMES-like whole system modelling frameworks to stochastic formulations, based on editing the dataset for the deterministic model to create multiple scenarios, and then adding linking constraints to connect these scenarios together into a stochastic optimization problem.

The essence of the extended TIMES models in the research is a kind of stochastic linear optimization models with two planning stages. Generally, the stochastic extensions of the TIMES model are addressed by solving their deterministic equivalent optimization problems in standard deterministic approaches [28]. However, it is sometimes difficult to solve the resulting deterministic problems since they could be unreasonably large if there are a large number of scenarios or high-dimensional variables. Therefore, it is also important to develop a more effective methodology for solving the two-stage stochastic TIMES optimization problems with recourse function.

Generally the recourse functions of stochastic optimization models with two stages have the following formulation:

$$\begin{aligned}
 E_{\xi}Q(\mathbf{x}, \xi) &:= \sum_{\xi_i \in \Omega} p(\xi_i)Q(\mathbf{x}, \xi_i), \\
 \text{where } Q(\mathbf{x}, \xi_i) &:= \min_{\mathbf{y}} q(\mathbf{x}, \mathbf{y}, \xi_i), \\
 \text{s.t. } &g(\mathbf{x}, \mathbf{y}, \xi_i) \leq 0,
 \end{aligned} \tag{1.1}$$

where  $\mathbf{x}, \mathbf{y}$  are the decision variables in the first and second stages respectively,  $\xi$  is a random variable defined on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . More specifically, the sample space of the random variable  $\xi$  is  $\Omega$ , i.e. the set of all possible values of  $\xi$ ,  $\mathcal{F}$  is the  $\sigma$ -field, and  $\mathbb{P}$  is the probability function.

There is an extensive literature addressing stochastic optimization problems which mainly concerns two kind of methods. On the one hand, approximation algorithm [22] is widely applied in the stochastic programming for the random variables with a continuous distribution, i.e. there are infinite number of realizations of random variable,  $|\Omega| = +\infty$ . It is complicated to compute the recourse in (1.1) for the continuous random variable  $\xi$ . Approximation algorithms are convenient tools to deal with this type of stochastic optimization problem which approximated  $\xi$  by a discrete random variable. After representing approximately the random variable  $\xi$ , the stochastic programming problems can also be transformed into deterministic approximate optimization problems and the deterministic programs could be later solved by conventional deterministic nonlinear programming theory.

L. Qi and R. S. Womersley [34] introduced SQP algorithm into two-stage stochastic programs with recourse function. X. Feng et al. [12] studied stochastic nonlinear model based on a tailored Jacobian approximation and an adjoint-based SQP method. Approximation algorithm is also used in the stochastic programming for discrete random variables where the expected recourse function or objective function is replaced by its corresponding sample average approximation at each iteration [1, 22]. However the approximation method relies heavily on the sample size and it would result in huge computational cost.

On the other hand, decomposition algorithm [36] is popular when dealing with the stochastic programming for the random variables with a discrete distribution, i.e.  $|\Omega|$  is a finite positive integer. Through the method, the original large-scaled optimization problem is decomposed into multiple

small-scale sub-problems and then the sub-problems are solved separately. These small-scale sub-problems are generally linear, quadratic or non-linear programming and their optimal solutions could be relatively easily obtained by general deterministic programming methods or software. Horst and Pardalos also summarized the following advantages of decomposition methods [16]: (1) Decomposition methods can heavily reduce the dimension of the problem. (2) The optimization problems for each stage or scenario can be obtained through decomposition methods, which will help us to have a more comprehensive understanding of the original problem. (3) The sub-problems obtained through decomposition methods can be solved simultaneously, which can decrease time spent in computation to a large extent.

There are two popular sorts of decomposition algorithms in general: Stage decomposition and scenario decomposition. Benders decomposition is a well-known representative of stage decomposition method. J. F. Benders [4] proposed the Benders decomposition method and later, Van Slyke and Wets [45] applied the approach in two-stage linear programming and named it as the cutting plane method, in which a constraint is added to the master optimization at each iteration. However the original cutting plane method has weak effect in tightening feasible region at each iteration and could result in a large quantity of iterations. Thus there were several modified algorithms proposed to speed up the convergence of the original Benders decomposition including multi-cuts [6], sparse cutting planes [8], pseudo-valid cutting planes [35], etc. If there are many realizations of random data, then we may need a large amount of iterations to add cuts in the master problem through Benders decomposition. In addition, nested Benders decomposition cannot be generalized to multi-stage stochastic problems with integer variables. Therefore, scenario decomposition method is more widely applied in solving stochastic optimization problem with two stages.

One of the well-known representatives of scenario decomposition is the Lagrangian decomposition method [15], which is an effective tool to deal with combinatorial optimization problems. Lagrangian decomposition is also known as variable splitting, in which we replicate several variables, dualize complicating constraints, and then divide the original large-scaled problem into scenario-wise simpler sub-problems. Lagrangian decomposed sub-problems have well-understood structures which have a connection with Lagrangian relaxation formulations. It provides at least as tight lower bounds to a minimization problem as those produced from Lagrangian relaxation [23]. Additionally, one of the most important characteristics of Lagrangian decomposition methods is that they are typically simple and effective [36]. If we use subgradient algorithms to solve Lagrangian dual programs, then there are two steps at each iteration: Solving small sub-problems and updating Lagrangian multipliers by simple additions. Only a few iterations are required to achieve a high-precision solution and there is almost no need to consider the scale of the problem. Thus Lagrangian decomposition methods are widely used in stochastic linear programming and integer programming.

A. Berkelaar et al. [5] modified the dual decomposition algorithm to tackle with two-stage stochastic linear optimization problems. K. Kim and V. M. Zavala [21] studied two-stage stochastic programs with mixed-integer recourse through the Lagrangian decomposition method. I. Aravena and A. Papavasiliou [2] proposed an asynchronous Lagrangian scenario decomposition to address stochastic unit model. Since the extended TIMES energy optimization model is a stochastic linear programming problem with two stages, our research will use Lagrangian decomposition method to deal with it.

The remainder of the article is organized as follows: Section 2 outlines the energy system processes

---

through the TIMES-like modelling frameworks. We also convert two present deterministic TIMES models into stochastic optimization problems accounting for uncertainty in the background to planning and policy, and construct the database of the stochastic TIMES model as a discrete scenario tree. In Section 3 we present a theoretical overview of the Lagrangian decomposition algorithm applied to solve the stochastic extension of TIMES model. Section 4 illustrates the findings of the numerical test on the feasibility of the algorithm. In the final section, we conclude the research and provide several suggestions for future work.

## 2. Underlying models

The section provides a general description of TIMES model, as well as the mathematical specifications of two kinds of deterministic TIMES modelling frameworks. Then the deterministic TIMES systems are extended into stochastic optimization problems with non-anticipativity constraints to connect scenarios together.

### 2.1. Overview of TIMES model

TIMES is a globally popular bottom-up model energy system optimization model, which is widely applied in government decision support in energy, economic and environmental over a medium-long time horizon [28]. The inputs for TIMES model consist of the demands for end-use energy service, the storage of existing energy supply resources as well as their potential uses, the features of available energy technologies and other information depending on the goal of users. There are a set of exogenous constraints concerning resources, technology capacity, end-use demand, as well as emissions and regional policies [27]. The flexibility of data input and constraints in the TIMES enables the analysis and evaluation among different scenarios and policies.

The objective terms of the TIMES model are typically linear optimization functions and it aims to obtain the optimal investment on the resource supply, energy commodity trade, and operation expense on the energy system technologies in order to minimize the entire cost of the energy system over a medium-long term time horizon [28]. The whole planning horizon is split into multiple time periods which generally have different time length. To put it more specific, the TIMES model takes estimates of baseline energy service demands as the starting point, and then dynamically selects the strategies on the energy supply in the first stage and then energy conversion process in the following stages during the planning period [42].

As is stated above, the entire energy process mainly consists of the three sections-resource supply, conversion and end-use demand [42]. The resources in TIMES refer to the current or potential raw energy sources, such as oil, coal, and even wind or solar. The resources is obtained by general extraction approaches like drilling and mining, as well as import. The resource supply process also includes trade on the resources and we can select to export exogenously the commodities outside the model boundary. Then the resources are converted into utilizable energy services by several conversion technologies and industrial processes. Finally the end-use demand section describes the amount of end-use energy services required in each time period which is present as constraints in the linear TIMES optimization model.

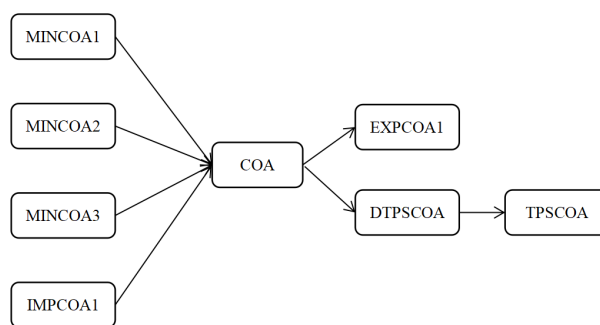
Building a TIMES model is mainly composed of two parts: setting energy database and linear programming [18]. The energy database is the basis for establishing the TIMES model, which

provides the information on the supply, conversion technology, demand of energy as well as relevant environmental or economical policies with the energy system.

In the research, we establish two kinds of TIMES systems loosely based on [14], which are illustrated as example models to demonstrate the effectiveness and validity of the scenario decomposition algorithm in Section 3. The first step TIMES model introduced in Section 2.2 (Model 1) starts with a single energy commodity, a single three-step supply curve as well as a single demand. It serves as a prototype model and a starting point for the progress in the use of TIMES features and variants [14]. The second step model in Section 2.3 (Model 2) is developed from the first one, which includes more conversion processes options.

## 2.2. A basic TIMES system–Model 1

We first present a simple TIMES model which concerns with a single region over two time periods. Coal (COA) is the only energy commodity in Model 1. We can obtain COA commodity from the following two ways. On the one hand, we can import COA from other regions by the import technology IMPCOA1. On the other hand, COA can be mined through a domestic three-step supply curve including the technologies MINCOA1, MINCOA2 and MINCOA3. Moreover, we can select a part of COA produced by the four technologies above to be exported outside of the region by the export technology EXPCOA1. The model also includes an end-use demand device called DTPSCOA, in which the input commodity is COA and the output product is a new demand commodity called TPSCOA, see Figure 1.



**Figure 1.** Schematic view of the basic deterministic TIMES model, Model 1.

There are two decision stages in Model 1 and each stage may include several years. We make decisions in the first year of each time period and implement these strategies annually, thus a time slice denotes one year in the example. The decisions we make in each time period include the output COA commodity of each energy technology, new capacity investment and the output TPSCOA commodity in the conversion technology DTPSCOA. The objective of the basic TIMES model is to minimize the total cost of the energy system. In the following subsections, we provide a full description on the basic TIMES model and its stochastic extension.

### 2.2.1. Algebraic statement

The essence of TIMES energy system optimization model is a minimization problem. In the following we present the mathematical specification of our deterministic TIMES model with two stages.

#### Indices:

- $t$  : Time period when decisions are made
- $s$  : Time slice in the time periods
- $p$  : Coal production technology

#### Sets:

- $\mathcal{T}$  : Set of time periods {1,2}
- $\mathcal{S}_t$  : Set of time slices in time period  $t$
- $\mathcal{P}$  : Set of production technologies {1,2,3 (MINCOA1 - MINCOA3), 4 (IMPCOA1), 5 (EXPCOA1)}

#### Decision variables:

- $\alpha_{p,t} \geq 0$  : The amount of energy commodity COA produced by technology  $p$  at period  $t$  [PJ]
- $\beta_t \geq 0$  : The amount of capacity addition in the conversion technology at period  $t$  [PJ]
- $\gamma_t \geq 0$  : Activity level of the conversion technology at time period  $t$  [PJ]

#### General parameters:

- $N_p$  : The cumulative amounts of resources available of technology  $p$  [PJ]
- $M_{p,t}$  : The annual cost per unit of COA production by technology  $p$  in period  $t$  [2005M€/PJ]
- $B_p$  : The maximum production limit on energy commodity COA of technology  $p$  [PJ]
- $Y_t$  : The amount of years/time slices in time period  $t$
- $E$  : The efficiency of the conversion technology DTPSCOA
- $U$  : The annual utilization factor of the conversion technology DTPSCOA
- $I$  : The investment cost per unit of capacity addition of the conversion technology [2005M€/PJ]
- $O$  : The fixed operation and maintenance cost per unit of capacity for conversion process DTPSCOA [2005M€/PJ\*Year]
- $V_t$  : The demand value for the conversion technology DTPSCOA [PJ]
- $G_p \in \{1, -1\}$  : 1 if technology  $p \in \{\text{MINCOA1}, \text{MINCOA2}, \text{MINCOA3}, \text{IMPCOA1}\}$ , and else -1
- $BY$  : The reference year for discounting
- $EY$  : The end of the whole planning horizon
- $FY_t$  : The first time slice of the time period  $t$
- $TL$  : The lifetime of the conversion technology DTPSCOA [Year]
- $DR$  : The overall discount rate for the energy system

#### Special parameters:

Apart from the general parameters, we define the following parameters which will be used in the discounting and calculation of cost components in the objective function of the example TIMES model. The parameters  $BY$  -  $DR$  are used to calculate the discount parameters  $D_t^1, D_t^2$  and fraction



parameter  $F_t$ , and the cost components are introduced in the next subsection.

$D_t^1$ : The discount term for the annual cost components at time period  $t$  including the flow cost component and the fixed operating and maintenance cost component, which is calculated as below:

$$D_t^1 := \sum_{s \in \mathcal{S}_t} (1 + DR)^{BY-s}, \quad t \in \mathcal{T}. \quad (2.1)$$

$D_t^2$ : The discount term for the cost components which happen per time period including the investment cost component on capacity addition, and salvage value of capacity and commodities at the end of the planning horizon at time period  $t$ , which is calculated as below:

$$D_t^2 := (1 + DR)^{BY-FY_t}, \quad t \in \mathcal{T}. \quad (2.2)$$

$F_t$ : The fraction of investment costs which is salvage after the end of whole planning horizon. It is applied in the computation of salvage values of investment costs, which is calculated as below [29]:

$$F_t := \frac{(1 + DR)^{TL+FY_t-EY-1} - 1}{(1 + DR)^{TL} - 1}, \quad t \in \mathcal{T}. \quad (2.3)$$

### Objective function:

TIMES models aim to minimize the total cost of the energy system and therefore they use objective terms called ‘‘Cost minimization objective’’. The total system cost is formed by the following four cost components,  $C_t^{flo}$ ,  $C_t^{fom}$ ,  $C_t^{inv}$ ,  $C_t^{sal}$ . It is also notable that the salvage value is actually a payback from the investment and therefore  $C_t^{sal}$  is defined as negative in (2.7).

- Flow cost component of resource supply processes:

$$C_t^{flo} := \sum_{p \in \mathcal{P}} M_{p,t} G_p \alpha_{p,t}, \quad t \in \mathcal{T}. \quad (2.4)$$

- Fixed operation and maintenance cost component of energy conversion process:

$$C_t^{fom} := O \sum_{k=1}^t \beta_k, \quad t \in \mathcal{T}. \quad (2.5)$$

- Investment cost component on capacity addition of energy conversion process:

$$C_t^{inv} := I\beta_t, \quad t \in \mathcal{T}. \quad (2.6)$$

- Salvage value of capacities augmentation after the end of whole planning horizon:

$$C_t^{sal} := -F_t C_t^{inv}, \quad t \in \mathcal{T}. \quad (2.7)$$

Using the notation described above, the objective function of the basic deterministic TIMES model which is to obtain the minimal value of the whole system cost reads as follows.

$$\min \sum_{t \in \mathcal{T}} D_t^1 (C_t^{flo} + C_t^{fom}) + D_t^2 (C_t^{inv} + C_t^{sal}). \quad (2.8)$$

The flow cost component in (2.4) presents the expense on the resource supply at time period  $t$ .  $C_t^{fom}$  refers to the fixed operation and maintenance (O&M) cost in (2.5) of the conversion technology DTPSCOA. The fixed O&M cost component happens in each year of the whole planning horizon, and is related with the cumulative technology capacity investment. Equation (2.6) illustrates the investment cost on the capacity addition of the conversion process, which is proportion to the decision on technology capacity investment in the current time period.

$C_t^{sal}$  denotes the salvage value of capacities after the end of whole planning horizon in (2.7). The conversion technology DTPSCOA has a finite technical lifetime  $TL$ . If the lifetime exceeds the whole horizon of the TIMES model, we will gain the salvage value of the investment cost as a return. Salvaging the investment costs is an indispensable process in the calculation of the entire system cost, since the investments on the capacity addition can still perform their function after the end of model horizon which should not be neglected. The salvage value is relevant with the investment cost, and more specifically, it can be expressed as a fraction of the investment costs in (2.3).

Considering that the equal amount of money will have less real value in the future owing to factors such as inflation, the cost components should be discounted to the reference year in the objective (2.8).

### Constraints:

In the following we list the constraints of the basic deterministic TIMES model. Cumulative resource constraints (2.9) impose an upper restriction for the three coal mining technologies in pursuit of sustainable development. The upper bounds of the coal production, import and export processes are present by annual production constraints (2.10). Export constraints (2.11) guarantee that the amount of coal exported exogenously outside the model boundary is limited. Efficiency and availability bounds (3.6) and (3.7) allow a certain degree of wastage and underutilization in the energy conversion process. End-use energy demand constraints (2.14) require energy supply and demand to be balanced at all times, while non-negativity constraints (2.15) ensure all the decision variables are no less than zero.

- Bounds on cumulative resource:

$$\sum_{t \in \mathcal{T}} Y_t \alpha_{p,t} \leq N_p, \quad p \in \mathcal{P}. \quad (2.9)$$

- Bounds on annual production:

$$\alpha_{p,t} \leq B_p, \quad p \in \mathcal{P}, t \in \mathcal{T}. \quad (2.10)$$

- Bounds on export technology:

$$\alpha_{5,t} \leq \sum_{p \in \mathcal{P} \setminus \{5\}} \alpha_{p,t}, \quad t \in \mathcal{T}. \quad (2.11)$$

- Efficiency bounds of conversion process:

$$\sum_{p \in \mathcal{P}} EG_p \alpha_{p,t} \geq \gamma_t, \quad t \in \mathcal{T}. \quad (2.12)$$

- Availability bounds of conversion process:

$$\sum_{k=1}^t U \beta_k \geq \gamma_t, \quad t \in \mathcal{T}. \quad (2.13)$$

- Bounds on end-use energy service demand:

$$\gamma_t \geq V_t, \quad t \in \mathcal{T}. \quad (2.14)$$

- Non-negativity constraints:

$$\begin{aligned} \alpha_{p,t} &\geq 0, \quad p \in \mathcal{P}, t \in \mathcal{T}, \\ \beta_t &\geq 0, \quad t \in \mathcal{T}, \\ \gamma_t &\geq 0, \quad t \in \mathcal{T}. \end{aligned} \quad (2.15)$$

### 2.2.2. Stochastic extension

There is broadly interest in the whole cycle of uncertainty treatment, from the uncertainty specification based on policymakers' judgments, through efficient optimization computation, through to communication of modelling results in the policy-making context. In this section, we extend the current deterministic TIMES system modelling framework to a stochastic formulation accounting for uncertainty in the background to planning and policy, and devise the stochastic TIMES dataset.

In the stochastic extension, we choose a scenario formulation with non-anticipativity constraints, which is also called the consensus form. In other words, the database of the stochastic TIMES model is built into a discrete scenario tree for uncertainty treatment in the energy system. Therefore, the main structure of TIMES model need not be altered but adding the scenario index  $w$ . Moreover the consensus formulation is easier to apply scenario decomposition owing to the non-anticipativity property.

Let  $\mathcal{W}$  be the set of scenarios and let  $\alpha_{p,t,w}, \beta_{t,w}, \gamma_{t,w}$  be the decision variables in the stochastic formulation for all  $w \in \mathcal{W}$ . There are several stochastic parameters associated with scenario in the formulation,  $M_{p,t,w}, B_{p,t,w}, E_{t,w}, U_{t,w}, I_{t,w}, O_{t,w}, V_{t,w}$ . In addition, the parameter  $S_w$  denoting the probability of scenario  $w$  is introduced to calculate the expected value of energy system cost. Therefore the constraints (2.9)–(2.15) are converted into the following form:

$$\begin{aligned}
\sum_{t \in \mathcal{T}} Y_t \alpha_{p,t,w} &\leq N_p, \quad p \in \mathcal{P}, w \in \mathcal{W}, \\
\alpha_{p,t,w} &\leq B_{p,t,w}, \quad p \in \mathcal{P}, t \in \mathcal{T}, w \in \mathcal{W}, \\
\alpha_{5,t,w} &\leq \sum_{p \in \mathcal{P} \setminus \{5\}} \alpha_{p,t,w}, \quad t \in \mathcal{T}, w \in \mathcal{W}, \\
\sum_{p \in \mathcal{P}} E_{t,w} G_p \alpha_{p,t,w} &\geq \gamma_{t,w}, \quad t \in \mathcal{T}, w \in \mathcal{W}, w \in \mathcal{W}, \\
\sum_{k=1}^t U_{k,w} \beta_{k,w} &\geq \gamma_{t,w}, \quad t \in \mathcal{T}, w \in \mathcal{W}, \\
\gamma_{t,w} &\geq V_{t,w}, \quad t \in \mathcal{T}, w \in \mathcal{W}, \\
\alpha_{p,t,w} &\geq 0, \quad p \in \mathcal{P}, t \in \mathcal{T}, w \in \mathcal{W}, \\
\beta_{t,w} &\geq 0, \quad t \in \mathcal{T}, w \in \mathcal{W}, \\
\gamma_{t,w} &\geq 0, \quad t \in \mathcal{T}, w \in \mathcal{W}.
\end{aligned} \tag{2.16}$$

Similarly the four cost components in the objective of TIMES are also connected with scenario, and we could extend the Eqs (2.4)–(2.7) as below:

$$\begin{aligned}
C_{t,w}^{flo} &:= \sum_{p \in \mathcal{P}} M_{p,t,w} G_p \alpha_{p,t,w}, \quad t \in \mathcal{T}, w \in \mathcal{W}, \\
C_{t,w}^{fom} &:= O_{t,w} \sum_{k=1}^t \beta_{k,w}, \quad t \in \mathcal{T}, w \in \mathcal{W}, \\
C_{t,w}^{inv} &:= I_{t,w} \beta_{t,w}, \quad t \in \mathcal{T}, w \in \mathcal{W}, \\
C_{t,w}^{sal} &:= -F_t C_{t,w}^{inv}, \quad t \in \mathcal{T}, w \in \mathcal{W}.
\end{aligned} \tag{2.17}$$

In the two-stage stochastic programming system, the decisions are fixed in the first stage, while the second stage decisions vary from scenario to scenario. Thus we use the non-anticipativity constraints (2.18) to connect all the scenarios together, which regulate that the first stage decisions for each scenario have the same values.

$$\begin{aligned}
\alpha_{p,1,|\mathcal{W}|} &= \alpha_{p,1,1}, \alpha_{p,1,w} = \alpha_{p,1,w+1}, \quad p \in \mathcal{P}, w \in \mathcal{W} \setminus \{|\mathcal{W}|\}, \\
\beta_{1,|\mathcal{W}|} &= \beta_{1,1}, \beta_{1,w} = \beta_{1,w+1}, \quad w \in \mathcal{W} \setminus \{|\mathcal{W}|\}, \\
\gamma_{1,|\mathcal{W}|} &= \gamma_{1,1}, \gamma_{1,w} = \gamma_{1,w+1}, \quad w \in \mathcal{W} \setminus \{|\mathcal{W}|\}.
\end{aligned} \tag{2.18}$$

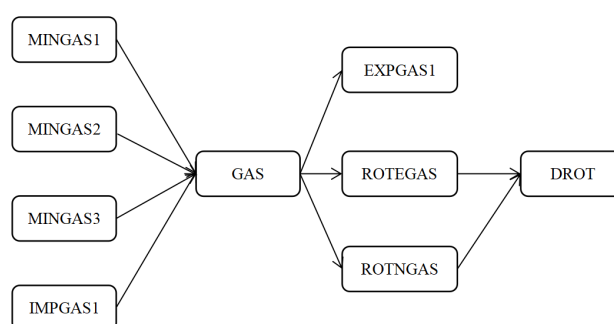
Let the constraint set be  $\Gamma_1 := \{(\alpha, \beta, \gamma) | (2.16) - (2.18), \forall p, t, w\}$ . Then the stochastic extensive formulation of the TIMES model is given by:

$$\begin{aligned}
\min \quad & \sum_{t \in \mathcal{T}, w \in \mathcal{W}} S_w (D_t^1 (C_{t,w}^{flo} + C_{t,w}^{fom}) + D_t^2 (C_{t,w}^{inv} + C_{t,w}^{sal})), \\
s.t. \quad & (\alpha, \beta, \gamma) \in \Gamma_1.
\end{aligned} \tag{2.19}$$

Contrary to the deterministic version of Model 1, the stochastic extension (2.19) is tightly associated with scenario. In the objective function,  $w$  is added into four cost components and we calculate the whole system cost based on probabilities of scenarios. The constraint set can be divided into two parts including stochastic extension of (2.9)–(2.15) and non-anticipativity constraints (2.18). The former one is scenario-wise decomposable while the latter one ensures the first stage decisions are identical for all scenarios.

### 2.3. A more complex TIMES system–Model 2

We then illustrate a more complex TIMES model, which includes a greater amount of energy conversion options. It focuses on the residential end-use energy service and gas (GAS) is the raw energy commodity in Model 2. On the supply side, there are three extraction technologies (MINGAS1-MINGAS3), an import (IMPGAS1) and export (EXPCOA) processes. The energy conversion sector provides two conversion technologies (ROTEGAS, ROTNGAS) which can transform GAS into a residential end-use energy demand DROT. Moreover, Model 2 takes emission commodities (CO<sub>2</sub>) into account in conversion processes. Figure 2 gives a schematic illustration of the TIMES energy system.



**Figure 2.** Schematic view of the more complex deterministic TIMES model, Model 2.

The second step model still concerns with a single region over two decision stages and has approximately the same initial assumptions as Model 1. The main difference between two models lies in more energy conversion processes of Model 2, and therefore it has more decision variables which also includes the gas input of energy conversion. Meanwhile, there are a greater number of parameters and constraints in Model 2, since we need to consider the storage of existing installed capacity and CO<sub>2</sub> emissions in energy conversion processes.

#### 2.3.1. Algebraic statement

The subsection gives the mathematical specification of our second step TIMES model. It has a similar modelling framework to Model 1 and thus we only present the modifications in the following. The index  $q$  and set  $Q$  are added in the model to denote two energy conversion technologies. The conversion processes are also characterised by  $R_q, C_q$  which present a base-year stock of existing installed capacity and CO<sub>2</sub> emission coefficient. We also set an upper limit  $CB$  for the total CO<sub>2</sub> emission in the energy conversion process. In addition, the coefficients of efficiency, availability, investment cost, fixed O&M cost, technology lifetime, and salvage fraction are associated with conversion process index  $q$ .

#### New index, set and parameters:

$q$  : End-use energy conversion process

$Q$  : Set of conversion technologies {1 (ROTEGAS), 2 (ROTNGAS)}

$R_q$  : The existing installed capacity of conversion technology  $q$  [PJ]

$C_q$  : The coefficient of emission commodity CO2 for conversion technology  $q$  [kt/PJ]

$CB$  : The upper bound for CO2 emission of conversion process [kt]

$E_q$  : The efficiency of conversion technology  $q$

$U_q$  : The annual utilization factor of conversion technology  $q$

$I_q$  : The investment cost per unit of capacity addition of conversion technology  $q$  [2005M€/PJ]

$O_q$  : The fixed O&M cost per unit of capacity for conversion process  $q$  [2005M€/(PJ\*Year)]

$TL_q$  : The lifetime of conversion technology  $q$  [Year]

$F_{q,t}$  : The fraction of investment costs which is salvage after the end of whole planning horizon, which is calculated as below:

$$F_{q,t} := \frac{(1 + DR)^{TL_q + FY_t - EY - 1} - 1}{(1 + DR)^{TL_q} - 1}, \quad q \in \mathcal{Q}, t \in \mathcal{T}. \quad (2.20)$$

### Decision variables:

Contrary to Model 1, there are a greater amount of decision variables including the amounts of gas production, capacity investment, input and output commodity of energy conversion:

$\alpha_{p,t} \geq 0$  : The amount of energy commodity GAS produced by technology  $p$  at period  $t$  [PJ]

$\beta_{q,t} \geq 0$  : The amount of capacity addition in conversion technology  $q$  at period  $t$  [PJ]

$\gamma_{q,t} \geq 0$  : The amount of GAS consumed in conversion technology  $q$  at time period  $t$  [PJ]

$\sigma_{q,t} \geq 0$  : Activity level of conversion technology  $q$  at time period  $t$  [PJ]

### Objective function:

Due to the addition of  $q$  in energy conversion process, the calculations of the fixed O&M, investment and salvage cost components are all modified into (2.21)–(2.24). Then we can obtain the objective function of Model 2 as (2.8).

- Fixed operation and maintenance cost component of energy conversion process:

$$C_t^{fom} := \sum_{q \in \mathcal{Q}} O_q (R_q + \sum_{k=1}^t \beta_{q,k}), \quad t \in \mathcal{T}. \quad (2.21)$$

- Investment cost component on capacity addition of energy conversion process:

$$C_{q,t}^{inv} := I_q \beta_{q,t}, \quad q \in \mathcal{Q}, t \in \mathcal{T}. \quad (2.22)$$

- Salvage value of capacities augmentation after the end of whole planning horizon:

$$C_{q,t}^{sal} := -F_{q,t} C_{q,t}^{inv}, \quad q \in \mathcal{Q}, t \in \mathcal{T}. \quad (2.23)$$

- Addition of investment and salvage value on capacity addition of energy conversion process:

$$C_t^{inv} + C_t^{sal} := \sum_{q \in \mathcal{Q}} (1 - F_{q,t}) I_q \beta_{q,t}, \quad t \in \mathcal{T}. \quad (2.24)$$

### Constraints:

Compared with Model 1, there are several differences in the constraint set. Model 2 allows the reservation of raw energy commodity GAS and energy balance constraints (2.25) ensure gas input in the conversion process does not exceed energy supply. In pursuit of sustainable development, we impose an upper limit for the total amount of CO<sub>2</sub> emission in the energy system in (2.26). The inequalities (2.27) guarantee that conversion process ROTEGAS cannot increase its capacity endogenously through new investment, while the initial assumption that conversion process ROTEGAS is available only in the second stage is given by (2.28). Moreover, we make modifications to several bounds such as efficiency and availability bounds because of new index, decisions and parameters.

- Bounds on energy balance:

$$\sum_{k=1}^t \left( \sum_{p \in \mathcal{P}} G_p \alpha_{p,k} - \sum_{q \in \mathcal{Q}} \gamma_{q,k} \right) \geq 0, \quad t \in \mathcal{T}. \quad (2.25)$$

- Bound on emission commodity in conversion technology:

$$\sum_{q \in \mathcal{Q}} C_q \sigma_{q,t} \leq CB, \quad t \in \mathcal{T}. \quad (2.26)$$

- Bounds on capacity increase of conversion technology ROTEGAS:

$$\beta_{1,t} = 0, \quad t \in \mathcal{T}. \quad (2.27)$$

- Bound on start year of conversion technology ROTNGAS:

$$\beta_{2,1} = 0. \quad (2.28)$$

- Efficiency bounds of conversion process:

$$E_q \gamma_{q,t} \geq \sigma_{q,t}, \quad q \in \mathcal{Q}, t \in \mathcal{T}. \quad (2.29)$$

- Availability bounds of conversion process:

$$U_q (R_q + \sum_{k=1}^t \beta_{q,t}) \geq \sigma_{q,t}, \quad q \in \mathcal{Q}, t \in \mathcal{T}. \quad (2.30)$$

- Bounds on end-use energy service demand:

$$\sum_{q \in \mathcal{Q}} \sigma_{q,t} \geq V_t, \quad t \in \mathcal{T}. \quad (2.31)$$

- Non-negativity constraints:

$$\begin{aligned} \alpha_{p,t} &\geq 0, & p \in \mathcal{P}, t \in \mathcal{T}, \\ \beta_{q,t} &\geq 0, & q \in \mathcal{Q}, t \in \mathcal{T}, \\ \gamma_{q,t} &\geq 0, & q \in \mathcal{Q}, t \in \mathcal{T}, \\ \sigma_{q,t} &\geq 0, & q \in \mathcal{Q}, t \in \mathcal{T}. \end{aligned} \quad (2.32)$$

### 2.3.2. Stochastic extensions

In the following we extend deterministic Model 2 into a stochastic optimization energy system similarly as is introduced in Section 2.2.2. Inequalities (2.33) are stochastic extensions of the initial deterministic constraint set and they can be decomposed by scenario.

$$\begin{aligned}
 C_{t,w}^{flo} &:= \sum_{p \in \mathcal{P}} M_{p,t,w} G_p \alpha_{p,t,w}, \quad t \in \mathcal{T}, w \in \mathcal{W}, \\
 C_{t,w}^{fom} &:= \sum_{q \in \mathcal{Q}} O_{q,t,w} (R_q + \sum_{k=1}^t \beta_{q,k,w}), \quad t \in \mathcal{T}, w \in \mathcal{W}, \\
 C_{t,w}^{inv} + C_{t,w}^{sal} &:= \sum_{q \in \mathcal{Q}} (1 - F_{q,t}) I_{q,t,w} \beta_{q,t,w}, \quad t \in \mathcal{T}, w \in \mathcal{W}, \\
 \alpha_{p,t,w} &\leq B_{p,t,w}, \quad p \in \mathcal{P}, t \in \mathcal{T}, w \in \mathcal{W}, \\
 \alpha_{5,t,w} &\leq \sum_{p \in \mathcal{P} \setminus \{5\}} \alpha_{p,t,w}, \quad t \in \mathcal{T}, w \in \mathcal{W}, \\
 \sum_{k=1}^t (\sum_{p \in \mathcal{P}} G_p \alpha_{p,k,w} - \sum_{q \in \mathcal{Q}} \gamma_{q,k,w}) &\geq 0, \quad t \in \mathcal{T}, w \in \mathcal{W}, \\
 E_{q,t,w} \gamma_{q,t,w} &\geq \sigma_{q,t,w}, \quad q \in \mathcal{Q}, t \in \mathcal{T}, w \in \mathcal{W}, \\
 U_{q,t,w} (R_q + \sum_{k=1}^t \beta_{q,k,w}) &\geq \sigma_{q,t,w}, \quad q \in \mathcal{Q}, t \in \mathcal{T}, w \in \mathcal{W}, \\
 \sum_{q \in \mathcal{Q}} \sigma_{q,t,w} &\geq V_{t,w}, \quad t \in \mathcal{T}, w \in \mathcal{W}, \\
 \beta_{2,1,w} = 0, \beta_{1,t,w} = 0, &\quad t \in \mathcal{T}, w \in \mathcal{W}, \\
 \alpha_{p,t,w}, \beta_{q,t,w}, \gamma_{q,t,w}, \sigma_{q,t,w} &\geq 0, \quad p \in \mathcal{P}, q \in \mathcal{Q}, t \in \mathcal{T}, w \in \mathcal{W}.
 \end{aligned} \tag{2.33}$$

The non-anticipativity property is guaranteed by the following constraints.

$$\begin{aligned}
 \alpha_{p,1,|\mathcal{W}|} &= \alpha_{p,1,1}, \alpha_{p,1,w} = \alpha_{p,1,w+1}, \quad p \in \mathcal{P}, w \in \mathcal{W} \setminus \{|\mathcal{W}|\}, \\
 \beta_{1,q,|\mathcal{W}|} &= \beta_{1,q,1}, \beta_{1,q,w} = \beta_{1,q,w+1}, \quad q \in \mathcal{Q}, w \in \mathcal{W} \setminus \{|\mathcal{W}|\}, \\
 \gamma_{1,q,|\mathcal{W}|} &= \gamma_{1,q,1}, \gamma_{1,q,w} = \gamma_{1,q,w+1}, \quad q \in \mathcal{Q}, w \in \mathcal{W} \setminus \{|\mathcal{W}|\}, \\
 \sigma_{1,q,|\mathcal{W}|} &= \sigma_{1,q,1}, \sigma_{1,q,w} = \sigma_{1,q,w+1}, \quad q \in \mathcal{Q}, w \in \mathcal{W} \setminus \{|\mathcal{W}|\}.
 \end{aligned} \tag{2.34}$$

We denote  $\Gamma_2 := \{(\alpha, \beta, \gamma, \sigma) | (2.33) - (2.34), \forall p, q, t, w\}$  as the constraint set of stochastic Model 2. Then we obtain the following stochastic extensive TIMES specification. Different from deterministic Model 2, the objective function is the expected value of energy system cost and non-anticipativity bounds (2.34) are included in the constraint set. The stochastic optimization TIMES model (2.35) has a greater amount of decision variables and constraints and it is more complex than stochastic Model 1.

$$\begin{aligned}
 \min \quad & \sum_{t \in \mathcal{T}, w \in \mathcal{W}} S_w (D_t^1 (C_{t,w}^{flo} + C_{t,w}^{fom}) + D_t^2 (C_{t,w}^{inv} + C_{t,w}^{sal})), \\
 \text{s.t.} \quad & (\alpha, \beta, \gamma, \sigma) \in \Gamma_2.
 \end{aligned} \tag{2.35}$$



### 3. Algorithm design

In this section, we develop a scenario decomposition algorithm to solve the stochastic TIMES models proposed in Sections 2.2.2 and 2.3.2. We first decompose the stochastic TIMES models by scenario using Lagrangian relaxation method, and then solve the multiple sub-problems respectively by the supergradient ascent algorithm.

To simplify the notation system, let  $\mathbf{x}_w = (\alpha_{1,1,w}, \dots, \alpha_{|\mathcal{P}|,1,w}, \beta_{1,w}, \gamma_{1,w})$  and  $\mathbf{y}_w = (\alpha_{1,2,w}, \dots, \alpha_{|\mathcal{P}|,2,w}, \beta_{2,w}, \gamma_{2,w})$  be the first and second stage decisions for scenario  $w \in \mathcal{W}$  in the stochastic TIMES models, where  $|\mathcal{P}|$ ,  $|\mathcal{W}|$  are the amounts of all technologies and scenarios. Let  $\mathbf{c}, \mathbf{q}$  be the vectors composed by certain parameters in the first stage, and uncertain parameters in the second stage respectively. The constraint set associated with scenario is  $\mathcal{X}_w$ , and  $\mathcal{X}_1 \cup \dots \cup \mathcal{X}_{|\mathcal{W}|} = \mathcal{W}$ . Then the stochastic TIMES models (2.19) and (2.35) are both equivalent to the following specification:

$$\min \sum_{w \in \mathcal{W}} S_w(\mathbf{c}^T \mathbf{x}_w + \mathbf{q}_w^T \mathbf{y}_w), \quad (3.1)$$

$$s.t. (\mathbf{x}_w, \mathbf{y}_w) \in \mathcal{X}_w, \quad w \in \mathcal{W}, \quad (3.2)$$

$$\mathbf{x}_w = \mathbf{x}_{w+1}, \quad w \in \mathcal{W} \setminus \{|\mathcal{W}|\}, \quad (3.3)$$

$$\mathbf{x}_{|\mathcal{W}|} = \mathbf{x}_1. \quad (3.4)$$

#### 3.1. Lagrangian decomposition

In scenario decomposition process, we make a copy of the first-stage decisions for each scenario and use the non-anticipativity constraints. The non-anticipativity constraints (3.3) and (3.4) are expressed as equality constraints. We transform them into the inequality constraints (3.5) to avoid the non-signed vectors of Lagrangian multipliers in the Lagrangian dual problem [11]:

$$\begin{aligned} \mathbf{x}_w &\leq \mathbf{x}_{w+1}, \quad w \in \mathcal{W} \setminus \{|\mathcal{W}|\}, \\ \mathbf{x}_{|\mathcal{W}|} &\leq \mathbf{x}_1. \end{aligned} \quad (3.5)$$

Then we dualize the non-anticipativity constraints (3.5) and obtain the corresponding Lagrangian relaxation problem in the following:

$$\begin{aligned} D(\lambda) := \min \quad & \sum_{w \in \mathcal{W}} S_w(\mathbf{c}^T \mathbf{x}_w + \mathbf{q}_w^T \mathbf{y}_w) + \sum_{w \in \mathcal{W} \setminus \{|\mathcal{W}|\}} \lambda_w^T (\mathbf{x}_w - \mathbf{x}_{w+1}) + \lambda_{|\mathcal{W}|}^T (\mathbf{x}_{|\mathcal{W}|} - \mathbf{x}_1), \\ s.t. \quad & (\mathbf{x}_w, \mathbf{y}_w) \in \mathcal{X}_w, \quad w \in \mathcal{W}, \end{aligned} \quad (3.6)$$

where  $\lambda_w$  is a non-negative vector with the same dimensions as  $\mathbf{x}_w$  for  $w \in \mathcal{W}$ , and  $\lambda = (\lambda_1, \dots, \lambda_{|\mathcal{W}|})$  which concatenates all the vectors  $\lambda_w$  together. To find the best lower bound for the primal problem (3.6), we need to solve the Lagrangian dual problem in the following:

$$\max_{\lambda \geq 0} D(\lambda). \quad (3.7)$$

Such dualization process is correct based on Proposition 1 regarding feasibility which follows [15]:

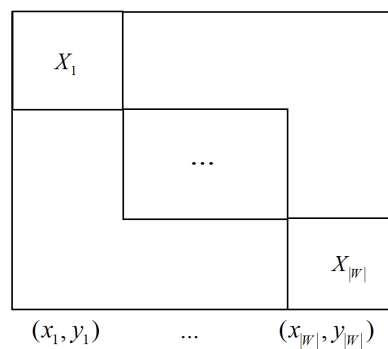
**Proposition 1.** (1) *The optimal value of the Lagrangian dual problem (3.7) is a lower bound for the optimal value of the primal optimization problem (3.1).*

(2) If the optimal solution of the Lagrangian dual problem (3.7)  $(x^*, y^*, \lambda^*)$  is feasible for the original optimization problem (3.1), then (3.1) and (3.7) have the same optimal objective value.

*Proof.* (1) We define  $z_{LR}(\lambda), z_{LD}, z_P$  as the optimal objective value for (3.6), (3.7) and (3.1) respectively. According to the definition of relaxations,  $z_{LR}(\lambda) \leq z_P$  for any  $\lambda \geq 0$ . Since  $z_{LD} = \max_{\lambda \geq 0} z_{LR}(\lambda)$ , we obtain that  $z_{LD} \leq z_P$ , i.e.  $z_{LD}$  is a lower bound for  $z_P$ .

(2) If  $\mathbf{x}^*$  is feasible for (3.1), then we obtain that  $\mathbf{x}_1^* = \mathbf{x}_2^* = \dots = \mathbf{x}_w^*$ . Therefore, the penalty term for the Lagrangian relaxation has the value of 0, and we obtain that  $z_{LD} = z_P$ .

To level up the computation efficiency, we can decompose the Lagrangian dual problem (3.7) into  $|\mathcal{W}|$  sub-problems which correspond to scenarios. Then the simpler and smaller separate problems can be solved in parallel. Figure 3 presents a schematic view of the reformulated optimization problem (3.6), which is the basis of the Lagrangian decomposition.



**Figure 3.** Schematic view of optimization problems with non-anticipativity constraints.

We make advantage of the structure of the Lagrangian relaxation problem (3.6) that it can be split into sub-problems by scenario. The penalty term in the objective function is shown as below:

$$\begin{aligned}
 & \lambda_1(\mathbf{x}_1 - \mathbf{x}_2) + \lambda_2(\mathbf{x}_2 - \mathbf{x}_3) + \dots + \lambda_{|\mathcal{W}|-1}(\mathbf{x}_{|\mathcal{W}|-1} - \mathbf{x}_{|\mathcal{W}|}) + \lambda_{|\mathcal{W}|}(\mathbf{x}_{|\mathcal{W}|} - \mathbf{x}_1) \\
 &= (\lambda_1 - \lambda_{|\mathcal{W}|})\mathbf{x}_1 + (\lambda_2 - \lambda_1)\mathbf{x}_2 + (\lambda_3 - \lambda_2)\mathbf{x}_3 + \dots + (\lambda_{|\mathcal{W}|} - \lambda_{|\mathcal{W}|-1})\mathbf{x}_{|\mathcal{W}|} \\
 &= (\lambda_1 - \lambda_{|\mathcal{W}|})\mathbf{x}_1 + \sum_{w=2}^{|\mathcal{W}|} (\lambda_w - \lambda_{w-1})\mathbf{x}_w.
 \end{aligned} \tag{3.8}$$

Thus the optimization problem (3.6) is expressed as a sum of smaller and simpler sub-problems as follows. All of these sub-problems are small-scaled optimization programs corresponding to each scenario. And the objective functions include a small part of the relaxation term of non-anticipativity constraints (3.5).

$$D(\lambda) = \sum_{i=1}^{|\mathcal{W}|} D_i(\lambda), \tag{3.9}$$

where  $D_1(\lambda)$  is given as:

$$D_1(\lambda) := \min_{\mathbf{x}_1, \mathbf{y}_1} S_1(\mathbf{c}^T \mathbf{x}_1 + \mathbf{q}_1^T \mathbf{y}_1) + (\lambda_1 - \lambda_{|\mathcal{W}|}) \mathbf{x}_1, \quad (3.10)$$

*s.t.*  $(\mathbf{x}_1, \mathbf{y}_1) \in \mathcal{X}_1,$

and for other  $w = 2, 3, \dots, |\mathcal{W}|$  we obtain the following optimization problem:

$$D_w(\lambda) := \min_{\mathbf{x}_w, \mathbf{y}_w} S_w(\mathbf{c}^T \mathbf{x}_w + \mathbf{q}_w^T \mathbf{y}_w) + (\lambda_w - \lambda_{w-1}) \mathbf{x}_w, \quad (3.11)$$

*s.t.*  $(\mathbf{x}_w, \mathbf{y}_w) \in \mathcal{X}_w, \quad w = 2, 3, \dots, |\mathcal{W}|.$

### 3.2. Supergradient ascent method

The Lagrangian dual problem (3.7) is a non-smooth optimization program with a concave objective function. We can use a variety of non-smooth optimization methods to obtain the optimal solutions. Here we use the supergradient method, because it's widely applied and quite simple to use and explain. Since the Lagrangian dual problem (3.7) is a maximization problem, we present a brief introduction to ascent method instead of descent.

In the supergradient method, we require the supergradient of the objective function of the Lagrangian dual problem (3.7) and the updated Lagrangian multipliers. Thus we should first compute the supergradient of  $D(\lambda)$  at  $\lambda$ . According to the basic supergradient property, the supergradient  $\mathbf{g}$  at  $\lambda$  needs to fulfill is shown in the following:

$$D(\mu) \leq D(\lambda) + \mathbf{g}^T(\mu - \lambda), \quad \forall \mu. \quad (3.12)$$

For a given Lagrangian multiplier  $\lambda$ , if  $\tilde{\mathbf{x}} = (\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_{|\mathcal{W}|})^T, \tilde{\mathbf{y}} = (\tilde{\mathbf{y}}_1, \dots, \tilde{\mathbf{y}}_{|\mathcal{W}|})^T$  are the optimal solutions of the Lagrangian relaxation problem (3.6),  $\tilde{\mathbf{r}} = (\tilde{\mathbf{r}}_1, \dots, \tilde{\mathbf{r}}_{|\mathcal{W}|})^T$  is set to be the residual vector of the solution  $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$  in the following:

$$\begin{aligned} \tilde{\mathbf{r}}_w &:= \tilde{\mathbf{x}}_{w+1} - \tilde{\mathbf{x}}_w, \quad w = 1, 2, \dots, |\mathcal{W}| - 1, \\ \tilde{\mathbf{r}}_{|\mathcal{W}|} &:= \tilde{\mathbf{x}}_1 - \tilde{\mathbf{x}}_{|\mathcal{W}|}. \end{aligned} \quad (3.13)$$

Then the supergradient for  $D(\cdot)$  at  $\lambda$  is  $\mathbf{g}(\lambda) = -\tilde{\mathbf{r}}$  and its proof is as follows. We rearrange the terms of the supergradient inequality (3.12) by introducing  $\mathbf{g}(\lambda) = -\tilde{\mathbf{r}}$  and we obtain the following inequality:

$$D(\mu) \leq D(\lambda) + \tilde{\mathbf{r}}^T(\lambda - \mu) = D(\lambda) + \lambda^T \tilde{\mathbf{r}} - \mu^T \tilde{\mathbf{r}}. \quad (3.14)$$

It is obvious that the value of the right side of (3.14) is just the objective function of the Lagrangian relaxation problem (3.6) at  $\mu$ , where  $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$  are optimal solutions of  $D(\lambda)$ . Since the Lagrangian relaxation problem (3.6) is a minimization problem,  $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$  are feasible solutions for  $D(\mu)$ , and therefore we can obtain the following relation:

$$\begin{aligned} & D(\lambda) + \lambda^T \tilde{\mathbf{r}} - \mu^T \tilde{\mathbf{r}} \\ & \geq \min(D(\lambda) + \lambda^T \tilde{\mathbf{r}} - \mu^T \tilde{\mathbf{r}}) \\ & = D(\mu). \end{aligned} \quad (3.15)$$

Thus the supergradient property (3.12) always holds, and the supergradient for  $D(\cdot)$  at  $\lambda$  is  $\mathbf{g}(\lambda) = -\tilde{\mathbf{r}}$ . Therefore, in our stochastic extended TIMES model, the supergradient for  $D(\cdot)$  at  $\lambda$  is the weighted residual  $\mathbf{r}$  of the optimal solution for  $D(\lambda)$ .

To construct the consensus form, we make a copy of the first stage decisions and parameters and the non-anticipativity constraints restrict that the first stage decision variables have identical value for all scenarios. Thus we use the following stopping criterion [11] which guarantees that the gap between the first stage decision for each scenario is sufficiently small. In inequalities (3.16),  $|\mathcal{W}|n_x$  denotes the number of non-anticipativity constraints for the  $\mathbf{x}$  vector of variables, and  $\epsilon$  is the given tolerance rate. In the numerical experiment in Section 4, we have considered  $\epsilon = 0.01$ .

$$RS := \sum_{w \in \mathcal{W}} \|\mathbf{r}\|_2^2, \quad \frac{RS}{|\mathcal{W}|n_x} \leq (\epsilon L_x)^2. \quad (3.16)$$

We assume that  $\lambda^*$ ,  $\lambda_k$  are an optimal solution and feasible solution of the Lagrangian dual problem (3.7) respectively. So if we take a sufficiently small step length  $\beta > 0$ , then we can obtain the following inequality:

$$\begin{aligned} & \|\lambda_k + \beta g(\lambda_k) - \lambda^*\|_2^2 \\ &= \|\lambda_k - \beta \mathbf{r}_k - \lambda^*\|_2^2 \\ &= \|\lambda_k - \lambda^*\|_2^2 - 2\beta \mathbf{r}^T (\lambda_k - \lambda^*) + \beta^2 \|\mathbf{r}\|_2^2 \\ &< \|\lambda_k - \lambda^*\|_2^2. \end{aligned} \quad (3.17)$$

The relation (3.17) shows that  $\lambda_k + \beta g(\lambda_k)$  is strictly closer to the optimal solution of the Lagrangian dual problem (3.7)  $\lambda^*$  than  $\lambda_k$ , which proves the convergence property of the supergradient ascent algorithm below to solve the Lagrangian dual problem (3.7) where the step length for the iteration count  $k$  is denoted as  $\beta_k$ .

Then the scenario decomposition algorithm can provide an approximation to the optimal Lagrangian dual objective value, and the speed of convergence depends on the choice of step length  $\beta_k$ , which follows the theorem [39] as below:

**Theorem 3.1** (Choice of stepsize in supergradient algorithm). *If the stepsizes applied in the supergradient descent method satisfy the following conditions:*

$$\beta_K = \frac{\lambda_K}{\|g(\lambda_K)\|}, \lambda_K \downarrow 0, \sum_{K=1}^{\infty} \lambda_K = +\infty. \quad (3.18)$$

*Then the sequence of best function-values  $\max D(\lambda)$  tends to the maximal value over  $\mathbb{R}^w$ .*

A common choice of  $\beta$  is Polyak step length, which has the following updating formulation [33]:

$$\beta_k := \frac{z_{LD} - z_{LR}(\lambda^k)}{RS}. \quad (3.19)$$

However the optimal Lagrangian dual objective value  $z_{LD}$  is unknown, thus we can find an upper bound  $\overline{z_{LD}}$  which can be any feasible solution of the original stochastic optimization problem. We also introduce an auxiliary parameter  $\alpha_k \in (0, 2]$  to control the convergence rate. If the objective function value does not increase in an given amount of iterations, the auxiliary parameter  $\alpha_k$  will be decreased. Then the step length updating formulation is as follows. According to [11], convergence will be reached in a finite amount of iterations using Polyak step length.

$$\begin{aligned}\beta_k &:= \alpha_k * \frac{\overline{z_{LD}} - z_{LR}(\lambda^k)}{RS} \\ &= \alpha_k * \frac{\overline{D(\lambda)} - D(\lambda^k)}{RS}.\end{aligned}\tag{3.20}$$

### 3.3. Lagrangian decomposition algorithm

Based on the theoretical content in Sections 3.1 and 3.2, we propose the following scenario decomposition algorithm to solve the stochastic TIMES model:

**Algorithm 3.1** (Lagrangian decomposition with supergradient algorithm). *Set the iteration count  $k = 0$ , the initial Lagrangian multiplier  $\lambda^0 = (\lambda_1^0, \dots, \lambda_{|\mathcal{W}|}^0)^T$ , and the tolerance rate  $\epsilon$ .*

*Step 1. Set  $k = k + 1$ . Solve the separate optimization problems (3.10) and (3.11) in parallel and let  $(\mathbf{x}_w^k, \mathbf{y}_w^k)$  be an optimal solution of  $D_w(\lambda^k)$ . Calculate the residual vector  $\mathbf{r}^k$  and the stopping term  $RS$  as (3.11) and (3.16):*

*(a) If (3.16) is satisfied, then we stop the process and the optimal solutions are as follows:*

$$\begin{aligned}\mathbf{x}^* &= \bar{\mathbf{x}}^k, \\ \mathbf{y}_w^* &= \mathbf{y}_w^k, \quad w \in \mathcal{W}.\end{aligned}\tag{3.21}$$

*(b) If (3.16) is unsatisfied, then go to Step 2.*

*Step 2. Update the Lagrangian multipliers  $\lambda^{k+1}$  by:*

$$\lambda_w^{k+1} = \lambda_w^k + \beta_k \mathbf{g}_w^k, \quad w \in \mathcal{W}.\tag{3.22}$$

## 4. Numerical test

In the following we conduct a numerical test to evaluate the feasibility and efficiency of Algorithm 3.1. The two deterministic TIMES energy system optimization models (Model 1 and Model 2), their stochastic extensions and the scenario decomposition algorithm are all implemented in Mosel. The deterministic TIMES datasets refer from [14]. We make several minor adjustments to the deterministic datasets and generate the stochastic scenarios artificially. The stochastic parameters in each scenario have the same orders of magnitude as their corresponding deterministic inputs, and their values are devised at random.

### 4.1. Test on the first step TIMES model–Model 1

Model 1 is a starting model for a series of TIMES development and variants in [14], which concerns with a single raw energy commodity, a single conversion process and a single end-use demand service. First we check the correctness of Model 1 by using Deterministic Test Case 1, and then use Stochastic Test Case 1 to assess the efficiency of Algorithm 3.1.

## Deterministic Test Case 1

The deterministic TIMES modelling framework, which is described in Section 2.2.1, lays a basis for the processes of stochastic extension and scenario decomposition. Therefore, it is essential to elaborate on the validity of Model 1.

There are two time periods in the test case, 2005, 2006 and 2007 respectively, i.e.  $BY = 2005, EY = 2007, Y_1 = 1, Y_2 = 2$ . The whole energy system consists of three components, resource supply, conversion process and end-use demand. The resource supply section includes a three-step supply curve (MINCOA1, MINCOA2 and MINCOA3), an import approach (IMPCOA1) and an export option (EXPCOA1), which describes the production and trade of the energy commodity COA. The second component illustrates the conversion technology DTPSCOA in which COA is transformed into an end-use energy service TPSCOA. Finally the end-use demand section illustrates our demand for TPSCOA.

In the following we present a brief introduction to Deterministic Test Case 1. There are three mining technologies and their corresponding cumulative amounts of resources available are 8000PJ, 16000PJ and 32000PJ. The mining costs per unit of COA are 2, 2.5 and 3 respectively, and the prices for the export and import option are both 2.75. The unit for costs and prices are all 2005M€/PJ. We also have the maximum production limits on technology MINCOA1, MINCOA2 and EXPCOA1 and they are 6074PJ, 2025PJ and 1147PJ separately. As for the conversion technology DTPSCOA, the efficiency is 100% and the utilization factor is 95%, while the technology life is 20 years. The investment cost and fixed O&M cost each are  $10 \cdot 2005\text{M€/PJ}$  and  $0.2 \cdot 2005\text{M€/PJ}$ . We require 13414PJ end-use energy service TPSCOA annually. The overall discount for the energy system is 5%. An overview of the test case statistics is provided in Table 1.

**Table 1.** Overview of statistics in Deterministic Test Case 1.

$N_1$	$N_2$	$N_3$	$M_1$	$M_2$	$M_3$
80000	160000	320000	2	2.5	3
$M_4$	$M_5$	$B_1$	$B_2$	$B_5$	$E$
2.75	2.75	6074	2025	1147	1
$U$	$I$	$O$	$V$	$TL$	$DR$
0.95	10	0.2	13414	20	0.05

We input the test case data for deterministic Model 1 and obtain the following optimal solutions shown in Table 2. It is guaranteed that we construct a correct TIMES modelling framework as the optimal solution is so close to the results in [14]. There is only one difference between the two results—our results present that 5315PJ coal is imported each year while the annual import amount is 6462PJ and the export amount is 1147PJ in [14]. Since the test case statistics show that the prices for import and export are the same, both two versions of outputs are optimal for the energy system and therefore the feasibility of our deterministic TIMES model is validated.

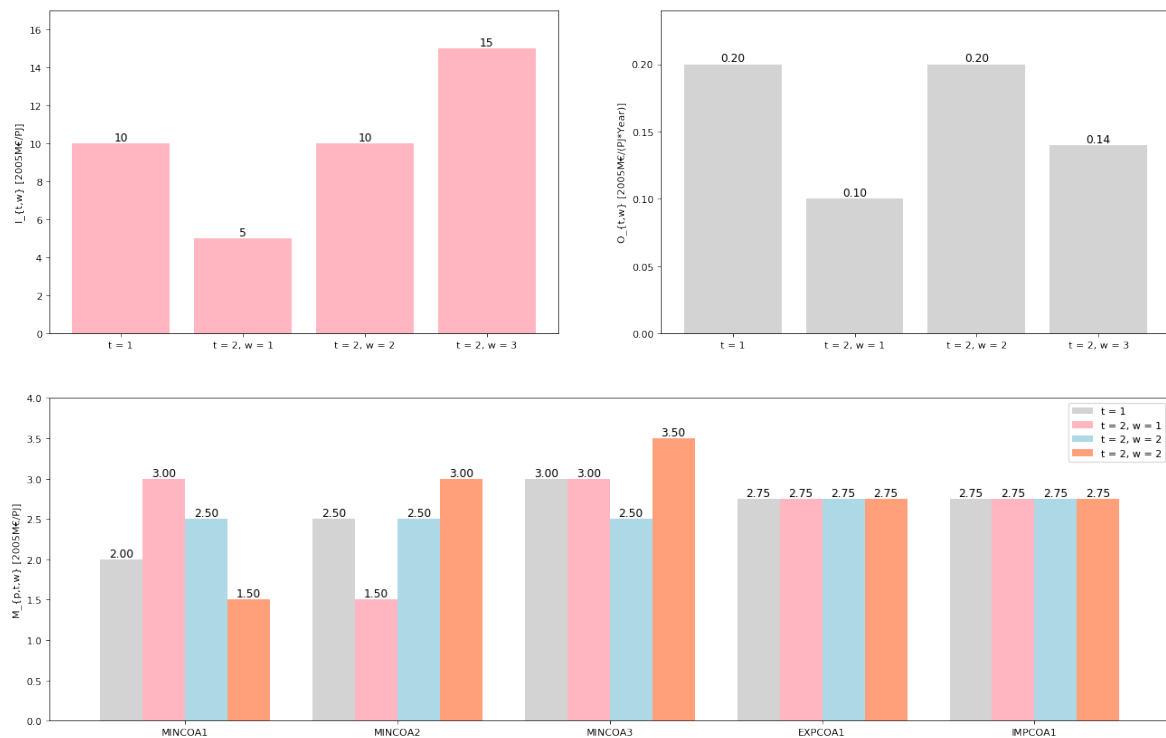
**Table 2.** Overview of outputs in Deterministic Test Case 1.

	$\alpha_{1,t}$	$\alpha_{2,t}$	$\alpha_{3,t}$	$\alpha_{4,t}$	$\alpha_{5,t}$	$\beta_t$	$\gamma_t$
$t = 1$	6074	2025	0	5315	0	14120	13414
$t = 2$	6074	2025	0	5315	0	14120	13414

### Stochastic Test Case 1

In the following, we evaluate the efficiency and feasibility of the Lagrangian decomposition algorithm in solving the stochastic extended TIMES model illustrated in Section 2.2.2. We apply the following Stochastic Test Case 1 in the numerical experiment, which are stochastic extensions to the deterministic dataset in Table 1. Therefore, the basic setting of the energy system keeps the same.

There are 3 scenarios with the probabilities of 0.3, 0.3 and 0.4 respectively in Stochastic Test Case 1. Compared with Deterministic Test Case 1, there are 3 stochastic parameters  $I_{t,w}$ ,  $O_{t,w}$ ,  $M_{p,t,w}$  varying between scenarios in the test case, which are described in Figure 4. Moreover the second-stage demands for the end-use energy device TPSCOA are modified to 26828PJ for two scenarios. In the stochastic test case, We set the initial Lagrangian multipliers to be all zeroes, and  $\overline{z_{LD}} = 226631$  in step length formulation (3.20).

**Figure 4.** Overview of stochastic parameters  $I_{t,w}$ ,  $V_{t,w}$  and  $M_{p,t,w}$  in Stochastic Test Case 1.

We input Stochastic Test Case 1 data in Mosel and obtain the following optimal solutions present in Table 3 after two iterations. The optimal objective value for Stochastic Test Case 1 is  $226630.1765 \cdot 2005\text{M€}$ . To assess the correctness of the results, we also construct the stochastic extension with non-anticipativity constraints of Model 1 (Validation Model 1) in Mosel and input the stochastic dataset

to the validation model. The outputs are identical and it confirms the scenario decomposition and Lagrangian multipliers updating processes in Algorithm 3.1 is valid.

**Table 3.** Overview of outputs in Stochastic Test Case 1.

	$\alpha_{1,t,w}$	$\alpha_{2,t,w}$	$\alpha_{3,t,w}$	$\alpha_{4,t,w}$	$\alpha_{5,t,w}$	$\beta_{t,w}$	$\gamma_{t,w}$
$t = 1, w \in \mathcal{W}$	6074	2025	0	5315	0	28240	13414
$t = 2, w = 1$	0	2025	0	24803	0	0	26828
$t = 2, w = 2$	6074	2025	19876	0	1147	0	26828
$t = 2, w = 3$	6074	0	0	20754	0	0	26828

We can make a comparison between the deterministic and stochastic outputs of Model 1. Due to different values of  $M_{p,2,w}$  in the scenarios, we will make different decisions in the second stage on the amounts of coal production for each technology  $\alpha_{p,2,w}$ . Additionally, the capacity investment on conversion technology DTPSCOA  $\beta_{1,w}$  increases compared with deterministic outputs, since the expected value of investment costs  $I_{2,w}$  becomes higher than its corresponding deterministic value  $I = 10$ .

#### 4.2. Test on the second step TIMES model–Model 2

Model 1 depicts a quite simple and basic TIMES energy system, and we need to use a more complex TIMES model to evaluate the feasibility and efficiency of the scenario decomposition algorithm. Model 2 illustrated in Section 2.3 is a development from the first step TIMES model, which concerns with more conversion technology options. We use the following deterministic and stochastic test case statistics in the simulation, and Deterministic Test Case 2 also refers from [14].

#### Deterministic Test Case 2

We first illustrates Deterministic Test Case 2 briefly. As a development of Deterministic Test Case 1, the test case holds similar assumptions. There are still two decision stages and five gas production technologies in the resource supply section. Then raw energy commodity GAS is converted into residential end-use energy demand DROT by two conversion processes, ROTEGAS and ROTNGAS. The conversion processes have different existing installed capacity and start years, and therefore there are a larger quantity of parameters, decisions and constraints in Deterministic Test Case 2. Table 4 gives an overview of the deterministic test case inputs.

**Table 4.** Overview of statistics in Deterministic Test Case 2.

$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	$B_1$	$B_5$
3.6	5.4	4.14	4.5	4.5	3950	2516
$E_1$	$E_2$	$U_1$	$U_2$	$I_2$	$O_1$	$O_2$
1	1.2	0.95	0.95	12	0.24	0.24
$R_1$	$C_1$	$C_2$	$CB$	$V$	$TL_1$	$TL_2$
5486	56.1	46.8	482660	5126	10	20



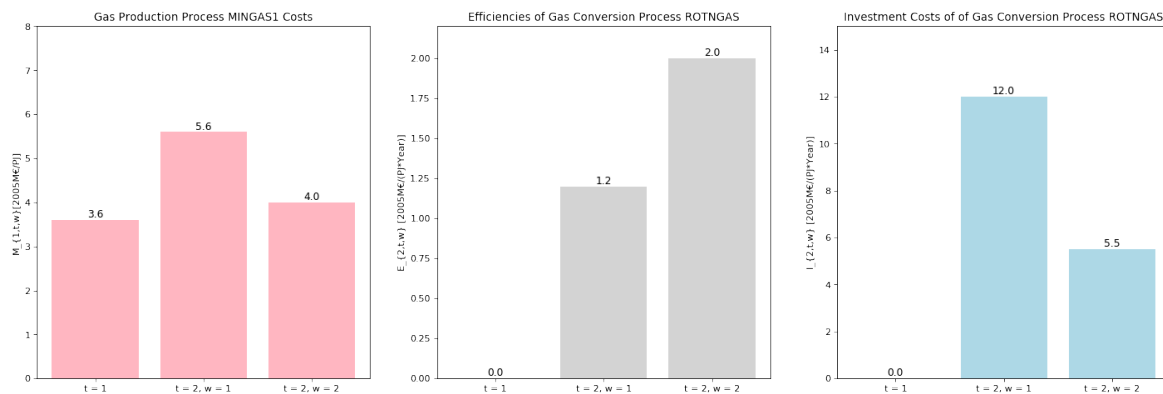
Table 5 shows the optimal solutions of Deterministic Test Case 2 and the objective value of the energy system is  $65757.2075 \cdot 2005\text{M€}$ , which are in common with the outputs given by [14]. Therefore it is confirmed that our deterministic TIMES modelling framework of Model 2 is feasible and valid.

**Table 5.** Overview of outputs in Deterministic Test Case 2.

	$\alpha_{1,t}$	$\alpha_{2,t}$	$\alpha_{3,t}$	$\alpha_{4,t}$	$\alpha_{5,t}$	$\beta_{1,t}$	$\beta_{2,t}$	$\gamma_{1,t}$	$\gamma_{2,t}$	$\sigma_{1,t}$	$\sigma_{1,t}$
$t = 1$	3950	0	1210	0	0	0	0	5160	0	5160	0
$t = 2$	3950	0	1132	0	0	0	412	4690	392	4690	470

### Stochastic Test Case 2

To evaluate the numerical behaviour of Algorithm 3.1, we apply the following stochastic test case. Stochastic Test Case 2 includes two scenarios with the same probabilities. The first-stage statistics keep the same. However, in the second stage, there are three stochastic parameters  $M_{p,t,w}$ ,  $E_{q,t,w}$ ,  $I_{q,t,w}$ , which appear in the objective function and constraints respectively. We present these stochastic statistics in Figure 5. The initial Lagrangian multipliers used in the test case are all zeroes, and  $Z_{LD}$  has the value of 65894.



**Figure 5.** Overview of stochastic parameters  $M_{p,t,w}$ ,  $E_{q,t,w}$ ,  $I_{q,t,w}$  in Stochastic Test Case 2.

In Stochastic Test Case 2, the optimal value for the objective function is  $58939.6712 \cdot 2005\text{M€}$ , which is obtained after five iterations. We outline the optimal decisions for the stochastic test case in Table 6. Likewise, a validation TIMES model is implemented in Mosel, which is the stochastic extended Model 1 with non-anticipativity constraints. We input the test case data in the validation model and obtain the same TIMES outputs as shown in Table 6. The simulation results reveal the feasibility of Algorithm 3.1 and its relatively good numerical performance in solving stochastic TIMES extensions with a small amount of scenarios.

Similarly, the simulation results of deterministic and stochastic models are compared with each other. In the first time period, energy conversion process ROTNGAS is unavailable and we could only use ROTNGAS process. However, due to higher efficiency and lower investment cost on capacity enhancement, we will make different decisions on the utilization of gas conversion technology ROTNGAS. Moreover, different choices of gas extraction technologies are made in the first scenario, since we could have a lower investment on gas production by technology MINGAS3.

**Table 6.** Overview of outputs in Stochastic Test Case 2.

	$\alpha_{1,t,w}$	$\alpha_{2,t,w}$	$\alpha_{3,t,w}$	$\alpha_{4,t,w}$	$\alpha_{5,t,w}$	$\beta_{1,t,w}$	$\beta_{2,t,w}$	$\gamma_{1,t,w}$	$\gamma_{2,t,w}$	$\sigma_{1,t,w}$	$\sigma_{1,t,w}$
$t = 1, w \in \mathcal{W}$	3950	0	1210	0	0	0	0	5160	0	5160	0
$t = 2, w = 1$	0	0	5082	0	0	0	412	4690	392	4690	470
$t = 2, w = 2$	2580	0	0	0	0	0	5432	0	2580	0	5160

### 4.3. Non-convergence analysis

In the subsections above, We have validated the efficiency and feasibility of the Lagrangian decomposition algorithm applied in stochastic extended TIMES systems. In the following, we make a further investigation on the numerical results and summarize several reasons for non-convergence cases.

#### Remark on choices of step length

The subgradient and supergradient algorithm do not strictly belong to descent/ascent algorithms. Thus the choice of step length is so important for the supergradient method in order to ensure the convergence of the algorithm. There are several commonly used stepsize criteria for supergradient algorithms, including the constant stepsize  $\beta_k = 2$ , the diminishing stepsize  $\beta_k = 1/k$  and Polyak stepsize as (3.19) presents.

In our research, we launch a numerical experiment with the three kinds of step length and present the simulation results in Table 7.  $k$  denotes the number of iterations for stochastic test cases, and  $n+$  means that the program does not converge after  $n$  iterations. For Stochastic Test Case 1, there are no upper bounds on the capacity investment  $\beta_{1,w}$  and activity level  $\gamma_{1,w}$  for the coal conversion technology DTPSCOA in the first stage. If the step lengths in the first several iterations are too large, then the objective functions of the sub-problems  $D_1(\lambda)$  and  $D_i(\lambda)$  might include components like  $k_1 * \beta_{1,w} + k_2 * \gamma_{1,w}$  with negative reals  $k_1$  and  $k_2$ . Therefore unsuitable step lengths such as the constant and diminishing stepsizes will lead to unbounded sub-problems. As for Stochastic Test Case 2, if we choose the constant stepsize, we will obtain a solution, where the gas production for processes MINGAS1 and MINGAS3  $\alpha_{1,1,w}, \alpha_{3,1,w}$  for two scenarios are different. It actually contradicts the non-anticipativity property. However, the program stops after 51 iterations, since the stopping criterion allows a certain margin of error. As a result, we decide to use Polyak stepsize in the supergradient method to solve the Lagrangian dual problem of stochastic TIMES model.

Table 7 clearly demonstrates that Polyak stepsize performs the best among the three choices of step length in our numerical experiment. However, Polyak stepsize relies heavily on the optimal objective value of the Lagrangian dual problem  $z_{LD}$ . Although we use  $\overline{z_{LD}}$  to replace the optimal objective value, it is still not an easy work to find a suitable feasible objective value of the original problem, which is one of the shortcomings of the current choice of stepsize. Currently we randomly select several feasible solutions and calculate their objective values given by Table 7. Their simulation results demonstrate that we will reach the unboundedness of sub-problems in Stochastic Test Case 1 as well if selecting a too large feasible objective value  $\overline{z_{LD}} = 397628$ . Although the objective value increases by less than 5% in the second example, we will need over three times the number of iterations to obtain the optimal solutions than the first example.

**Table 7.** Numerical performance of step lengths.

Step Length	Stochastic Test Case 1		Stochastic Test Case 2			
	$k$	If Converge	$k$	If Converge		
$\beta_k = 2$	196+	Non-convergence	51	Non-convergence		
$\beta_k = 1/k$	196+	Non-convergence	29	Convergence		
Polyak stepsize	$\overline{z_{LD}}$	$k$	If Converge	$\overline{z_{LD}}$	$k$	If Converge
	226631	2	Convergence	65894	5	Convergence
	235894	7	Convergence	59376	12	Convergence
	397628	196+	Non-convergence	75238	33	Convergence

For Stochastic Test Case 2, the optimal feasible objective value we have found in the current stage is  $\overline{z_{LD}} = 65894$ . Table 7 also reveals that a greater number of iterations are needed if we select a relatively high or low step length. However, we cannot confirm that  $\overline{z_{LD}} = 65894$  is the most suitable feasible objective value for the stochastic test case. Therefore, we will continue the research on the choices of stepsize of our algorithm in the future and one of the potential directions might be to adopt the dynamic stepsize criterion to cope with the problem.

### Remark on initial values of Lagrangian multipliers

Apart from the selection of step length criterion, we also need to focus on the initial choices of the Lagrangian multipliers. Unsuitable initial Lagrangian multipliers also result in unbounded optimization systems, and thus the program will become so time-consuming and need a huge amount of iterations. Table 8 compares the system performance under different initial choices of Lagrangian multipliers. The results illustrate that our supergradient method in solving the Lagrangian dual of the stochastic TIMES model is sensitive to the initial values of Lagrangian multipliers. If  $\lambda_{\beta,w+1} - \lambda_{\beta,w}$  or  $\lambda_{\gamma,w+1} - \lambda_{\gamma,w}$  is relatively large at the beginning like the third example, we will encounter unbounded sub-problems and take more than 150 iterations in the program. Therefore, how to select suitable initial value for the Lagrangian multipliers  $\lambda^0$  just like the first and second options is also an essential topic in our future research to improve our algorithm.

**Table 8.** Numerical performance of initial Lagrangian multipliers.

$\lambda_{\beta,\gamma}^0 = (\lambda_{\beta,1}^0, \lambda_{\beta,2}^0, \lambda_{\beta,3}^0, \lambda_{\gamma,1}^0, \lambda_{\gamma,2}^0, \lambda_{\gamma,3}^0)$	Number of Iterations	If Converge
(0, 0, 0, 0, 0, 0)	2	Convergence
(0.0005, 0, 0.0005, 0, 0.0005, 0)	2	Convergence
(0.1, 0.3, 0.2, 0.3, 0.2, 0.4)	157+	Non-convergence

### Remark on a special non-convergence case

Except for Stochastic Test Case 1, we also generate a stochastic dataset with four scenarios for Model 1. However, the test case is non-converged and the reason does not lie in unboundedness of sub-problems. For the first stage decision variables,  $\beta_{1,w}$  is different according to the scenario index  $w$  at the beginning. After two iterations  $\beta_{1,2}$  and  $\beta_{1,3}$  become 17594, and we obtain  $\beta_{1,4} = 17594$  after eight iterations. However  $\beta_{1,1}$  gets fluctuated in the following steps with the value of either 14120 or

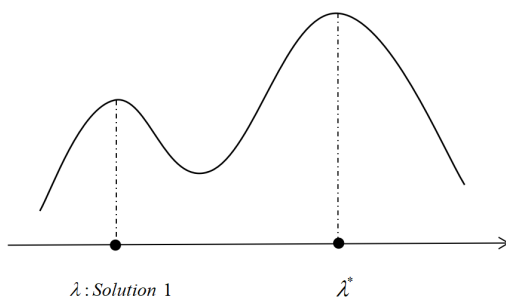
26828. The objective value of the Lagrangian dual problem is slowly increasing in general from the initial 186273.1125 to 186857.9304 after 112 iterations. In some iterations the objective value get decreased a little which is acceptable as the supergradient method is non-monotone. We present the different outputs  $\beta_{t,w}$  of the test case in Table 9, while the values of optimal decision  $\alpha_{p,t,w}, \gamma_{t,w}$  are identical in all scenarios.

**Table 9.** Outputs of Test Case 4 for stochastic TIMES model.

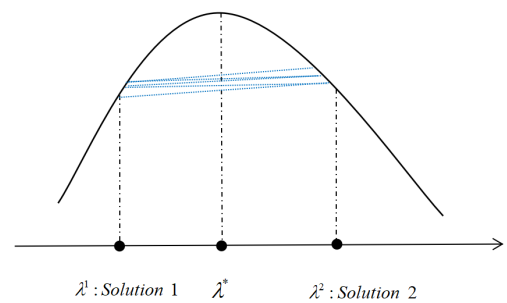
$\beta_{t,w}$	$t = 1$		$t = 2$			
	$w = 1$	$w = 2, 3, 4$	$w = 1$	$w = 2$	$w = 3$	$w = 4$
Solution 1	14120	17594	12780	0	10646	10646
Solution 2	26828	17594	0	0	10646	10646

Currently we summarise two main reasons for the non-converged test case. The first reason is that we reach a local optimal solution, Solution 1 in Table 9, since the numerical test illustrates that the iterations with the outputs of Solution 1 have higher objective values of approximately 186858 than those returning Solution 2 of 186850. The program does not stop as we use a strict stopping criterion that accuracy rate is no less than 1%. The schematic view of the situation is shown in Figure 6.  $\lambda$  is a local optimal solution while  $\lambda^*$  is the global optimal solution which we actually want to obtain. Actually it is one of the shortcomings of supergradient method that it is easily get stuck in a local optimal solution rather than a global one. To deal with the problem, we could try to use other non-smooth optimization methods, like bundle method.

The non-convergence may also result from unsuitable selection of step lengths. Figure 7 presents the schematic view of the second reason. The optimal Lagrangian multipliers  $\lambda^*$  are likely to be located between  $\lambda^1$  and  $\lambda^2$ . However due to too large stepsize, the optimal solution  $\lambda^*$  is reaches after a large quantity of steps which exceeds the upper bound for Mosel with community license.



**Figure 6.** Schematic view of reason 1.



**Figure 7.** Schematic view of reason 2.

## 5. Conclusions

TIMES is one of widely used energy system optimization models in energy and environmental policy research. The variability of future development in regional energy market makes it necessary to take uncertainty into consideration in energy decision support. In the article, we have extended the current deterministic TIMES modelling framework into a stochastic optimization problem with

multiple scenarios, and implemented the Lagrangian decomposition approach in solving the stochastic programming model.

We have first established two basic deterministic TIMES models based on [14] which deals with a single energy commodity, a single supply curve, a single demand option and different options of energy conversion process. The models mainly illustrate the following three sectors in energy system with two planning stages—resource supply and trade, conversion process and end-use demand. We have implemented the deterministic TIMES models in Mosel and validated the accuracy of the models by a numerical experiment.

We have also developed a method for converting the present deterministic TIMES model into a stochastic optimization problem. We edit the database of the deterministic TIMES model into a discrete scenario tree for uncertainty treatment in the policy-making context of the whole energy system. The first stage decision variables are copied for each scenario, the non-anticipativity constraints are added in the stochastic optimization problem to connect these scenarios together.

Finally we have applied Lagrangian decomposition method in solving the stochastic TIMES optimization models. We dualize the non-anticipativity constraints and decompose the Lagrangian dual problem into simpler and smaller sub-problems by scenario. Proposition 1 provides a solid theoretical foundation for the Lagrangian decomposition process. Then the Lagrangian dual problems are solved by supergradient ascent method. We have launched a numerical test on the performance of the scenario decomposition algorithm to solve two stochastic extended TIMES models by two stochastic test cases as well. We have found that the algorithm solves the test cases with  $\leq 3$  scenarios in no more than five iterations, which indicates that it has an overall good performance in solving stochastic TIMES with a small amount of scenarios. In addition, findings from the test show that unsuitable choices of step length and initial values of Lagrangian multipliers will lead to non-convergence of the algorithm.

There are following avenues for future research based on the findings in the paper:

First it is worth further research on detecting why the results of test case in Section 4.3 do not converge. Currently we summarise two main reasons. The first reason is that we reach a local optimal solution. It is one of the shortcomings of supergradient method that it converges to a local optima rather than a global one. In this case further insights can be gained by using other non-smooth optimization methods, like bundle method. In addition, unsuitable stepsizes can also lead to the fluctuation in the iterations. If the step length is too large, we will use much more iterations (over the maximum limit of Mosel with community license) to obtain the optimal solution. In this case we need to take further studies on step length as the next paragraph illustrates.

Further research may also aim at selecting a more suitable step length. We apply Polyak step length in the current algorithm in the thesis, however preliminary numerical results have shown that this kind of stepsize depends greatly on the choice of upper bounds for the Lagrangian dual problem. If we select a relatively large upper bound, then the sub-problems may be unbounded. Therefore, other types of step length such as dynamic stepsize criterion need to be investigated in future studies. In addition, there is another potential direction to deal with the problem. To avoid the unboundness of the sub-problems, the projected supergradient method is a good alternative, which makes a projection on the supergradient direction to ensure the feasibility of the Lagrangian multipliers.

It is also recommendable to research the numerical performance of the algorithm on more complex test models and cases. Currently we consider the simple two-stage TIMES models with a single region

and stochastic test cases with no greater than four scenarios. In real-life applications, the situation will be more complicated than the examples and will cover more time periods, regions and uncertain inputs. Therefore, the feasibility of the algorithm in solving stochastic extensions with more complex TIMES frameworks, a larger quantity of scenarios and decision stages should be confirmed by future research.

In addition, we might make a further investigation on how to generate the scenarios properly for our stochastic extended TIMES models. Currently the statistics are generated artificially and randomly and our aim is to validate the Lagrangian decomposition algorithm. However, in industrial applications, it is necessary to obtain TIMES inputs that meet future patterns. As a result, a data-driven manner to generate scenarios is a potential research topic.

## Acknowledgments

The authors thank Dr. Chris Dent (University of Edinburgh) and Zeynep Suvak (University of Edinburgh) for their valuable suggestions that greatly assisted our work. We thank Dr. Lars Schewe (University of Edinburgh) for useful discussions.

## Conflict of interest

All authors declare that there are no conflicts of interest regarding the publication of this paper.

## References

1. A. Agra, C. Requejo, F. Rodrigues, An adjustable sample average approximation algorithm for the stochastic production-inventory-routing problem, *Networks*, **72** (2018), 5–24. <https://doi.org/10.1002/net.21796>
2. I. Aravena, A. Papavasiliou, Asynchronous Lagrangian scenario decomposition, *Math. Program. Comput.*, **13** (2021), 1–50. <https://doi.org/10.1007/s12532-020-00185-4>
3. O. Balyk, K. S. Andersen, S. Dockweiler, M. Gargiulo, K. Karlsson, R. Næraa, et al., TIMES-DK: technology-rich multi-sectoral optimisation model of the Danish energy system, *Energy Strategy Rev.*, **23** (2019), 13–22. <https://doi.org/10.1016/j.esr.2018.11.003>
4. J. F. Benders, Partitioning procedures for solving mixed-variables programming problems, *Numer. Math.*, **4** (1962), 238–252.
5. A. Berkelaar, C. Dert, B. Oldenkamp, S. Zhang, A primal-dual decomposition-based interior point approach to two-stage stochastic linear programming, *Oper. Res.*, **50** (2002), 904–915. <https://doi.org/10.1287/opre.50.5.904.360>
6. J. R. Birge, F. V. Louveaux, A multicut algorithm for two-stage stochastic linear programs, *Eur. J. Oper. Res.*, **34** (1988), 384–392. [https://doi.org/10.1016/0377-2217\(88\)90159-2](https://doi.org/10.1016/0377-2217(88)90159-2)
7. C. Böhringer, The synthesis of bottom-up and top-down in energy policy modeling, *Energy Econ.*, **20** (1998), 233–248. [https://doi.org/10.1016/S0140-9883\(97\)00015-7](https://doi.org/10.1016/S0140-9883(97)00015-7)
8. S. S. Dey, M. Molinaro, Q. Wang, Analysis of sparse cutting planes for sparse MILPs with applications to stochastic MILPs, *Math. Oper. Res.*, **43** (2018), 304–332. <https://doi.org/10.1287/moor.2017.0866>

9. S. Di Leo, F. Pietrapertosa, S. Loperte, M. Salvia, C. Cosmi, Energy systems modelling to support key strategic decisions in energy and climate change at regional scale, *Renew. Sust. Energy Rev.*, **42** (2015), 394–414. <https://doi.org/10.1016/j.rser.2014.10.031>
10. R. K. Dutkiewicz, *International symposium on electricity, health and the environment*, Cap Town: Energy Research Institute, University of Cape Town, 1995.
11. L. F. Escudero, M. Garin, G. Perez, A. Unzueta, Lagrangian decomposition for large-scale two-stage stochastic mixed 0-1 problems, *Top*, **20** (2012), 347–374. <https://doi.org/10.1007/s11750-011-0237-1>
12. X. Feng, S. Di Cairano, R. Quirynen, Inexact adjoint-based SQP algorithm for real-time stochastic nonlinear MPC, *IFAC-PapersOnLine*, **53** (2020), 6529–6535. <https://doi.org/10.1016/j.ifacol.2020.12.068>
13. Y. Fu, G. Huang, L. Liu, M. Zhai, A factorial CGE model for analyzing the impacts of stepped carbon tax on Chinese economy and carbon emission, *Sci. Total Environ.*, **759** (2021), 143512. <https://doi.org/10.1016/j.scitotenv.2020.143512>
14. Documentation for the TIMES Model Part IV, Energy Technology Systems Analysis Programme, 2016. Available from: [https://iea-etsap.org/docs/Documentation\\_for\\_the\\_TIMES\\_Model-Part-IV\\_October-2016.pdf](https://iea-etsap.org/docs/Documentation_for_the_TIMES_Model-Part-IV_October-2016.pdf).
15. M. Guignard, S. Kim, Lagrangian decomposition: A model yielding stronger lagrangian relaxation bounds, *Math. Program.*, **39** (1993), 215–228. <https://doi.org/10.1007/BF02592954>
16. R. Horst, P. M. Pardalos, N. Van Thoai, *Introduction to global optimization*, Springer Science & Business Media, 2000.
17. C. S. Ioakimidis, K. N. Genikomsakis, Integration of seawater pumped-storage in the energy system of the Island of Sao Miguel (Azores), *Sustainability*, **10** (2018), 3438. <https://doi.org/10.3390/su10103438>
18. Y. Jia, R. Liu, Analysis of Beijing energy saving and emission reduction strategy based on times model, *J. Basic Sci. Eng.*, **21** (2013), 857–865.
19. R. Kanala, E. Fragniere, Mixing behavioral and technological data in mathematical programming framework, In: *Operations Research Proceedings 2012*, Springer, 2014, 163–168. [https://doi.org/10.1007/978-3-319-00795-3\\_24](https://doi.org/10.1007/978-3-319-00795-3_24)
20. A. Kanudia, R. Loulou, Robust responses to climate change via stochastic MARKAL: The case of Quebec, *Eur. J. Oper. Res.*, **106** (1998), 15–30. [https://doi.org/10.1016/S0377-2217\(98\)00356-7](https://doi.org/10.1016/S0377-2217(98)00356-7)
21. K. Kim, V. M. Zavala, Algorithmic innovations and software for the dual decomposition method applied to stochastic mixed-integer programs, *Math. Program. Comput.*, **10** (2018), 225–266. <https://doi.org/10.1007/s12532-017-0128-z>
22. A. J. Kleywegt, A. Shapiro, T. Homem-de Mello, The sample average approximation method for stochastic discrete optimization, *SIAM J. Optim.*, **12** (2002), 479–502. <https://doi.org/10.1137/S1052623499363220>

23. D. Kuhn, P. M. Esfahani, V. A. Nguyen, S. Shafieezadeh-Abadeh, *Wasserstein distributionally robust optimization: Theory and applications in machine learning*, Operations Research & Management Science in the Age of Analytics, 2019, 130–166. <https://doi.org/10.1287/educ.2019.0198>
24. T. Larsson, C. O. Wene, Developing strategies for robust energy systems. I: Methodology, *Int. J. Energy Res.*, **17** (1993), 503–513. <https://doi.org/10.1002/er.4440170606>
25. P. Li, I. Keppo, N. Strachan, Incorporating homeowners' preferences of heating technologies in the UK TIMES model, *Energy*, **148** (2018), 716–727. <https://doi.org/10.1016/j.energy.2018.01.150>
26. J. Liu, W. Chen, D. Liu, China's low carbon energy development strategy based on china times model system, *J. Tsinghua Univ.*, **51** (2011), 525–529.
27. R. Loulou, M. Labriet, ETSAP-TIAM: The TIMES integrated assessment model Part I: Model structure, *Comput. Manag. Sci.*, **5** (2008), 7–40. <https://doi.org/10.1007/s10287-007-0046-z>
28. Documentation for the TIMES Model Part I, Energy Technology Systems Analysis Programme, 2016. Available from: [https://iea-etsap.org/docs/Documentation\\_for\\_the\\_TIMES\\_Model-Part-I\\_July-2016.pdf](https://iea-etsap.org/docs/Documentation_for_the_TIMES_Model-Part-I_July-2016.pdf).
29. Documentation for the TIMES Model Part II, Energy Technology Systems Analysis Programme, 2016. Available from: [https://iea-etsap.org/docs/Documentation\\_for\\_the\\_TIMES\\_Model-Part-II\\_July-2016.pdf](https://iea-etsap.org/docs/Documentation_for_the_TIMES_Model-Part-II_July-2016.pdf).
30. M. J. Mayer, A. Szilágyi, G. Gróf, Environmental and economic multi-objective optimization of a household level hybrid renewable energy system by genetic algorithm, *Appl. Energy*, **269** (2020), 115058. <https://doi.org/10.1016/j.apenergy.2020.115058>
31. W. McDowall, B. S. Rodriguez, A. Usubiaga, J. A. Fernández, Is the optimal decarbonization pathway influenced by indirect emissions? Incorporating indirect life-cycle carbon dioxide emissions into a European TIMES model, *J. Clean. Prod.*, **170** (2018), 260–268. <https://doi.org/10.1016/j.jclepro.2017.09.132>
32. S. Pfenninger, A. Hawkes, J. Keirstead, Energy systems modeling for twenty-first century energy challenges, *Renew. Sust. Energy Rev.*, **33** (2014), 74–86. <https://doi.org/10.1016/j.rser.2014.02.003>
33. B. T. Polyak, *Introduction to optimization*, New York: Optimization Software Inc., Publications Division, 1987.
34. L. Qi, R. S. Womersley, An SQP algorithm for extended linear-quadratic problems in stochastic programming, *Ann. Oper. Res.*, **56** (1995), 251–285. <https://doi.org/10.1007/BF02031711>
35. W. Romeijnders, N. van der Laan, Pseudo-valid cutting planes for two-stage mixed-integer stochastic programs with right-hand-side uncertainty, *Oper. Res.*, **68** (2020), 1199–1217. <https://doi.org/10.1287/opre.2019.1905>
36. A. M. Rush, M. J. Collins, A tutorial on dual decomposition and Lagrangian relaxation for inference in natural language processing, *J. Artif. Intell. Res.*, **45** (2012), 305–362. <https://doi.org/10.1613/jair.3680>
37. A. Schaefer, H. D. Jacoby, Technology detail in a multisector CGE model: Transport under climate policy, *Energy Econ.*, **27** (2005), 1–24. <https://doi.org/10.1016/j.eneco.2004.10.005>



38. G. S. Seck, V. Krakowski, E. Assoumou, N. Maïzi, V. Mazauric, Embedding power system's reliability within a long-term Energy System Optimization Model: Linking high renewable energy integration and future grid stability for France by 2050, *Appl. Energy*, **257** (2020), 114037. <https://doi.org/10.1016/j.apenergy.2019.114037>
39. T. Schulze, *Stochastic programming for hydro-thermal unit commitment*, PhD's Thesis, The University of Edinburgh, 2015.
40. A. J. Seebregts, G. A. Goldstein, K. Smekens, Energy/environmental modeling with the MARKAL family of models, In: *Operations Research Proceedings 2001*, Springer, 2002, 75–82. [https://doi.org/10.1007/978-3-642-50282-8\\_10](https://doi.org/10.1007/978-3-642-50282-8_10)
41. G. Tian, Z. Wu, Y. Hu, Calculation of optimal tax rate of water resources and analysis of social welfare based on CGE model: A case study in Hebei Province, China, *Water Policy*, **23** (2021), 96–113. <https://doi.org/10.2166/wp.2020.118>
42. A Scottish TIMES Model: An Overview, The Scottish Government for the Environment, Climate Change and Land Reform Committee, 2016. Available from: [https://archive2021.parliament.scot/S5\\_Environment/General%20Documents/20160922\\_Scottish\\_TIMES\\_model\\_-\\_an\\_overview.pdf](https://archive2021.parliament.scot/S5_Environment/General%20Documents/20160922_Scottish_TIMES_model_-_an_overview.pdf).
43. UK TIMES Model Overview, UCL Energy Institute, 2014. Available from: [https://www.ucl.ac.uk/drupal/site\\_energy-models/sites/energy-models/files/uk-times-overview.pdf](https://www.ucl.ac.uk/drupal/site_energy-models/sites/energy-models/files/uk-times-overview.pdf).
44. Review of the Scottish TIMES Energy System Model, University College London, 2021. Available from: <https://era.ed.ac.uk/handle/1842/37509>.
45. R. M. Van Slyke, R. Wets, L-shaped linear programs with applications to optimal control and stochastic programming, *SIAM J. Appl. Math.*, **17** (1969), 638–663. <https://doi.org/10.1137/0117061>
46. Y. Wang, Y. Wang, Y. Huang, F. Li, M. Zeng, J. Li, et al., Planning and operation method of the regional integrated energy system considering economy and environment, *Energy*, **171** (2019), 731–750. <https://doi.org/10.1016/j.energy.2019.01.036>
47. Y. Zhang, M. Hong, J. Li, J. Ren, Y. Man, Energy system optimization model for tissue papermaking process, *Comput. Chem. Eng.*, **146** (2021), 107220. <https://doi.org/10.1016/j.compchemeng.2020.107220>
48. Z. X. Zhang, Can china afford to commit itself an emissions cap? An economic and political analysis, *Energy Econ.*, **22** (2000), 587–614. [https://doi.org/10.1016/S0140-9883\(00\)00048-7](https://doi.org/10.1016/S0140-9883(00)00048-7)



AIMS Press

©2022 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)