Modeling long-term energy transition pathways

Algorithmic approaches and their properties

Stefan Strömer[1,2] Kenneth Bruninx[2] Stefan Pfenninger-Lee[2]

[1] AIT Austrian Institute of Technology

[2] Delft University of Technology

2024-02-07

Overview

Why?

We (may) know where we want (or need) to be in 25 years: Models, plans, studies, reports, and opinions exist. What is often lost in the discussion are not only concrete first, but all steps along the way, towards reaching a specific goal.

Content

- Challenges & different model types
- Overview: Approaches
  - Simulation models
  - Reducing model resolution
  - Myopic paths
  - Decomposition algorithms
  - Stochastic Dual Dynamic Programming
- A simplified example

Challenges

General

- *"Are assets assets or burdens?"* (salvage/residual values)
- *"Is there a world after 2050?"* (end of horizon effects)
- Depiction of investments via annuities or lump-sum payments
- Seasonal storages (or budget constraints)
- Balancing contradictory levels of detail: long-term vs. hourly

Interpretation, comparing results, and benchmarking

Keep in mind, that ***optimal*** does not infer ***unique***. There may highly diverse optimal solutions, differing structurally, and even more near-optimal ones.

Snapshot models

Definition

- operate on consecutive *snapshots* ("= time steps") of a system
- for each, all input data can be specified
- often make use of (multi-) hourly snapshots
- often span a single, full year

Properties

- are "easy" to describe mathematically
- provide insights into the energy system for individual years
- lack the ability to depict dynamics in investment strategies throughout a transition period

Transition (pathway) models

Definition

- can be seen as extension to snapshot models
- represent the entire transition pathway to a long term target
- trade off short- and long-term goals
- can apply discounting for (far) future

Properties

- can be “hard” to describe mathematically
- provide insights into the energy system over a period of many years
- quickly become computationally intractable
- are not fully supported by many tools

Agent-based models (ABMs)

Simulation models, especially agent-based ones, can be

- used to simulate large periods of time, while keeping computational complexity (somewhat) under control;
- used to study investment decisions that profit-seeking investors might choose, in contrast to often-assumed central planning;
- “easily” linked to other topics, like risk aversion, social acceptance, learning curves, ...

Soft coupling: Coupling ABMs with optimization models can not only be done by considering agents that solve optimization problems, but also via (iterative) soft coupling.

Deissenroth, M., et al. (2017) – elib.dlr.de/117348/

Reducing model resolution - (I)

Direct reduction of model resolution

Examples: Either reduce the temporal resolution from n- to m-hourly ($m > n$) blocks, or reduce the spatial resolution (e.g., via clustering).

> Frysztacki M., et al. (2021) – doi: 10.1016/j.apenergy.2021.116726

Variable resolutions

Examples: Extend the simple reduction to account for higher details (temporal and spatial) during important periods (e.g., extreme events).

> Poncelet, K., et al. (2016) – doi: 10.1016/j.apenergy.2015.10.100

Reducing model resolution - (II)

Representative periods

Instead of modeling each snapshot, a *representative period* presumes that some periods throughout the year are highly similar to others, and therefore need not be modelled individually. Periods are often *days or weeks*.

Interdependencies

While many approaches target a reduction of interdependencies between distinct periods - myopic paths are a high-level example of that - representative periods initially increase the temporal coupling of a model.

Without any inter-temporal constraints, models immediately decompose into their representative periods, that can then be recombined (e.g., using weights).

Reducing model resolution - (III)

Fixing cyclic states

For small scale storages, it can be argued that this cyclic behaviour occurs more often, e.g., each week, or even day. Assuming that each cyclic period is constrained to the same initial state of charge (e.g., 50%) immediately leads to decoupled periods.

Foregoing per-snapshot decisions

Assuming that demand and generation behave "identical" for time periods that are similar, all decision variables - except those related to *states* (e.g., storages) - can be dropped and replaced by their representative.

github.com/spine-tools/SpineOpt.jl

Reducing model resolution - (IV)

Mixed temporal resolutions

With a similar argument, temporal granularity of decision variables can be reduced, e.g., by modeling renewables in hourly resolution, and nuclear plants in congruent 8-hour blocks. *This may not be studied enough to actually gauge whether it's actually an improvement.*

Complex inter-period behaviour

Many modern frameworks apply a formulation of seasonal storages, that summarizes intra-period states, properly respecting upper/lower storage bounds, while dropping decision variables based on representative periods.

Kotzur, L., et al. (2018) – doi: 10.1016/j.apenergy.2018.01.023
Gonzato, S., et al. (2021) – doi: 10.1016/j.apenergy.2021.117168
Gabrielli, P., et al. (2018) – doi: 10.1016/j.apenergy.2017.07.142

Myopic Paths

Approach

- picks a number of modeling years
- intializes the model for the first year, and optimizes it
- uses the results of that run as initialization of the next year's model

Limpens, G., et al. (2024) – doi: 10.1016/j.apenergy.2023.122501
Victoria, M., et al. (2020) – doi: 10.1038/s41467-020-20015-4

Advantages

- allow decommissioning between years
- are "easy" to configure with input data
- can overlap years (similar to MPC)

Disadvantages

- prone to lock-in effects
- assume intra-year perfect foresight
- "budget" constraints must be provided

Motivating model decompositions

$$
\begin{aligned}
\min_{\mathbf{x} \in X \subset \mathbb{R}^m} \quad & c' \cdot x \\
s.t. \quad & \\
& A \cdot \mathbf{x} \leq b
\end{aligned}
$$

Figure 1: Starting from an abstract form of a monolithic model ...

Motivating model decompositions

$$\min_{\mathbf{x} \in X \subset \mathbb{R}^m} \quad c' \cdot x$$
$$s.t.$$
$$A \cdot \mathbf{x} \leq b$$

$$\min_{(\mathbf{x},\mathbf{y}) \in X \subset \mathbb{R}^m} \quad c'_{design} \cdot y + c'_{operation} \cdot x$$
$$s.t.$$
$$A_1 \cdot \mathbf{x} \leq b_1$$
$$A_2 \cdot \mathbf{y} \leq b_2$$
$$A_3 \cdot \mathbf{x} \leq T \cdot \mathbf{y}$$

separation #1:

design vs.
operational variables

Figure 2: ... separating design and operational decisions ...

Motivating model decompositions

$$\begin{aligned} \min_{\mathbf{x} \in X \subset \mathbb{R}^m} \quad & c' \cdot x \\ s.t. \quad & \\ & A \cdot \mathbf{x} \leq b \end{aligned}$$

$$\begin{aligned} \min_{(\mathbf{x},\mathbf{y}) \in X \subset \mathbb{R}^m} \quad & c'_{design} \cdot y + c'_{operation} \cdot x \\ s.t. \quad & \\ & A_1 \cdot \mathbf{x} \leq b_1 \\ & A_2 \cdot \mathbf{y} \leq b_2 \\ & A_3 \cdot \mathbf{x} \leq T \cdot \mathbf{y} \end{aligned}$$

$$\begin{aligned} \min_{\mathbf{x} \in X \subset \mathbb{R}^?} \quad & c'_{operation} \cdot x \\ s.t. \quad & \\ & A_1 \cdot \mathbf{x} \leq b_1 \\ & A_3 \cdot \mathbf{x} \leq b_3 \quad (:= T \cdot y) \end{aligned}$$

separation #1:

design vs. operational variables

separation #2:

imagine that $\mathbf{y}$ is a fixed input

Figure 3: ... considering what could happen if we knew about the design beforehand.

The "Benders way to think"

$$\underbrace{\begin{aligned} \min_{\mathbf{y} \in X \subset \mathbb{R}^?} \quad & c'_{design} \cdot \mathbf{y} \\ s.t. \quad & \\ & A \cdot \mathbf{y} \leq b \end{aligned}}_{\textbf{main / outer problem}}$$

$$\underbrace{\begin{aligned} \min_{\mathbf{x} \in X \subset \mathbb{R}^?} \quad & c'_{operation} \cdot \mathbf{x} \\ s.t. \quad & \\ & A_{i,\cdot} \cdot \mathbf{x} \leq b_i && \forall i \in \{k+1, \dots, n\} \\ & A_{i,\cdot} \cdot \mathbf{x} \leq y_i && \forall i \in \{1, \dots, k\} \end{aligned}}_{\textbf{sub / inner problem(s)}}$$

Figure 4: Two separate tasks, two separate optimization problems ...

The "Benders way to think"

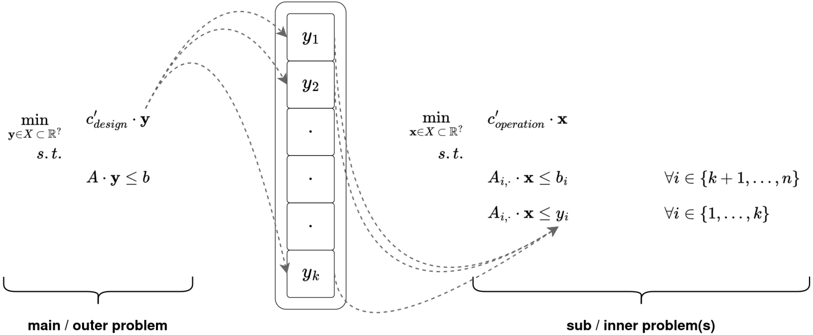
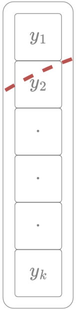


Figure 5: ... including a way to communicate a chosen design ...

The “Benders way to think”

$$
\begin{aligned}
\min_{\mathbf{y} \in X \subset \mathbb{R}^?} \quad & c'_{design} \cdot \mathbf{y} + \theta \\
s.t. \quad & \\
& A \cdot \mathbf{y} \leq b \\
& \theta \geq \dots
\end{aligned}
$$

main / outer problem



$$
\begin{aligned}
\min_{\mathbf{x} \in X \subset \mathbb{R}^?,\, \mathbf{b}} \quad & c'_{operation} \cdot \mathbf{x} \\
s.t. \quad & \\
& A_{i,\cdot} \cdot \mathbf{x} \leq b_i & & \forall i \in \{k+1, \dots, n\} \\
& A_{i,\cdot} \cdot \mathbf{x} \leq \mathbf{b}_i & & \forall i \in \{1, \dots, k\} \\
& \mathbf{b}_i = y_i & (\pi_i) & \forall i \in \{1, \dots, k\}
\end{aligned}
$$

sub / inner problem(s)

Figure 6: … including a “feedback” loop to improve our decisions.

Motivating SDDP

Used to solve

- **multistage** *[a sequence of decisions over time]*,
- **stochastic** *[potentially existing uncertainty that is gradually revealed over time]*

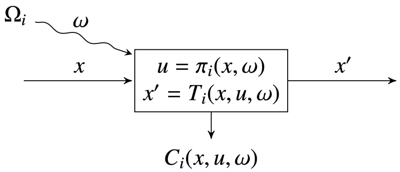
optimization problems. More or less formal relations can be seen to

- dynamic programming
- Kelley's cutting plane algorithm
- Benders decomposition
- reinforcement (Q) learning
- backpropagation

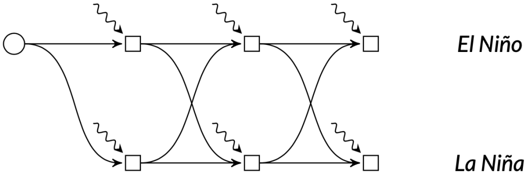
Dowson, O. and Kapelevich, L. (2021) – doi: 10.1287/ijoc.2020.0987

Motivating SDDP

Going from a single node ...



... to policy graphs with transition probabilities



... or even infinite horizons!

Dowson, O. (2020) – doi: 10.1002/net.21932

Motivating SDDP

The i-th value function $V_i(x, \omega)$

$$\begin{aligned} V_i(x, \omega) = \min_{\bar{\mathbf{x}}, \mathbf{x}', \mathbf{u}} \quad & C_i(\bar{\mathbf{x}}, \mathbf{u}, \omega) + \mathbb{E}_{j \in i^+, \varphi \in \Omega_j}[V_j(\mathbf{x}', \varphi)] \\ & \mathbf{x}' = T_i(\bar{\mathbf{x}}, \mathbf{u}, \omega), \quad \mathbf{u} \in U_i(\bar{\mathbf{x}}, \omega) \\ & \bar{\mathbf{x}} = x, \end{aligned}$$

... and it's approximation after k iterations

$$\begin{aligned} V_i^{(k)}(x, \omega) = \min_{\bar{\mathbf{x}}, \mathbf{x}', \mathbf{u}} \quad & C_i(\bar{\mathbf{x}}, u, \omega) + \theta \\ & \mathbf{x}' = T_i(\bar{\mathbf{x}}, \mathbf{u}, \omega), \quad \mathbf{u} \in U_i(\bar{\mathbf{x}}, \omega) \\ & \bar{\mathbf{x}} = x \\ & \theta \geq \ldots \end{aligned}$$

Properties

The beauty of SDDP lies in its natural relation to familiar modelling approaches: A path is defined by specific periods that are modelled in detail, with a transition function between them. Periods can, e.g., be full years or shorter periods with sophisticated linking.

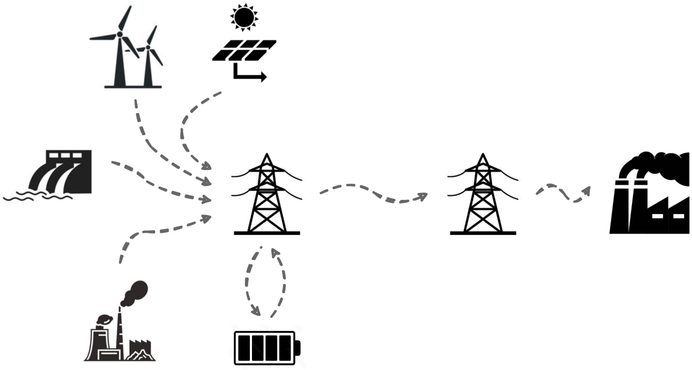
Some advantages

- approach can be used for deterministic problems
- alternative risk measures, e.g., CVaR can directly be used
- allows *hazard-decision* ("wait-and-see") or *decision-hazard* ("here-and-now") modes

Some challenges

- requires *relatively complete recource* (c.f. always feasible optimality-cut Benders)
- cut generation may be subject to instabilities (c.f. single- vs. multi-cut),
- which puts more emphasize on proper scaling of the model during formulation

Overview



github: sstroemer/euro2024-pathway-sddp-example

Results: Course of objective “bounds”

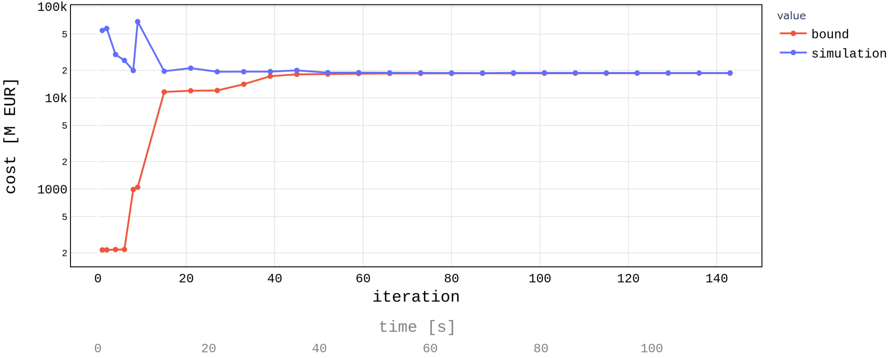


Figure 7: Due to first-class support for stochastic models, no explicit upper bound (compared to Benders) is tracked. Instead, an estimation is given by a forward-pass simulation. The iterative nature allows early stopping, which may otherwise not be a possibility.

Results: Investments

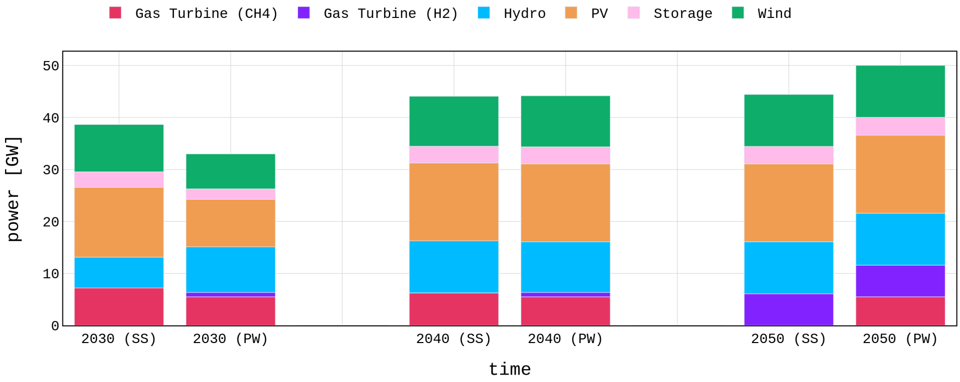


Figure 8: Investments into generation and storage assets, from 2030 to 2050, comparing a single-shot (SS) and pathway (PW) approach.

Results: Energy mix

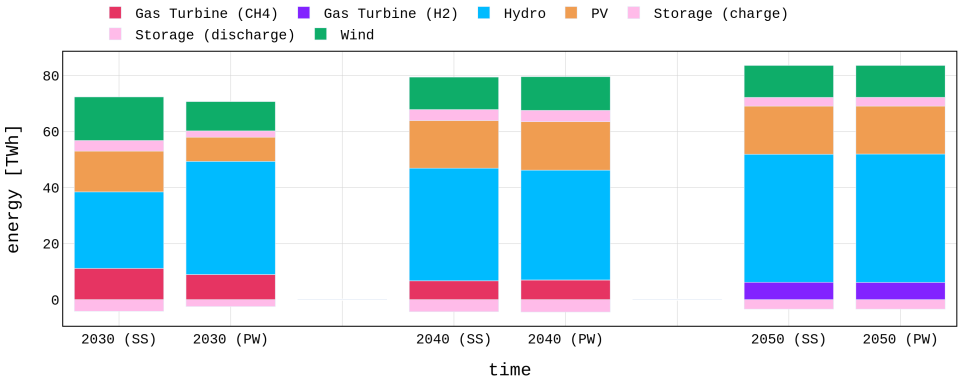
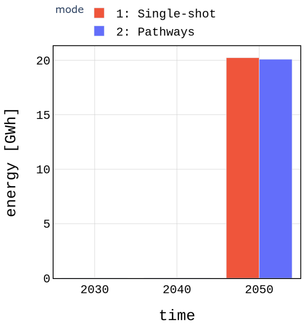


Figure 9: Contribution of different technologies to the total annual energy mix, from 2030 to 2050, comparing a single-shot (SS) and pathway (PW) approach.

Results: System KPIs



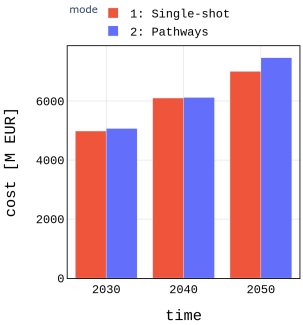


Figure 10: Comparison of high-level results: Load shedding results are almost identical, and only occur in later years (assumption: reduced costs due to the introduction of DSR products). The total annual system cost is slightly higher for PW (expected!), which even leads to a solution that could be considered near-optimal. Overheads in 2050 are caused by no decommissioning assumption in the simplified example.