Handling of long-term storage in multi-horizon stochastic programs

Michal Kaut

November 27, 2023

This paper shows how to implement long-term storage in the multi-horizon modelling paradigm, expanding the range of problems this approach is applicable to. The presented implementation is based on the HyOpt optimization model, but the ideas should be transferable also to other models implementing the multi-horizon approach.

We illustrate the effects of several different formulations on a simple case of electrifying an offshore installation using wind turbines and a hydrogen-based energy-storage system. The results show that the formulations provide reasonable modelling of the storage capacity, without sacrificing the advantages of the multi-horizon approach.

Multi-horizon stochastic programming (Kaut et al., 2019), is a modelling paradigm that enables the combination of several time scales in a single optimization model. Typically, this means a combination of long-term (strategic) decisions with a short-term (tactical or operational) model, used to evaluate the quality of the strategic decisions. For example, we want to build a new infrastructure (strategic time scale), which requires modelling its performance under varying operational conditions (operational time scale). An example of a multi-horizon scenario tree is in Fig. 1, where '●' denotes a *strategic node*, where strategic decisions are made. Each strategic node includes four *operational scenarios*, i.e., sequences of *operational nodes*, denoted by '▪'.

The main advantage of this approach is that it drastically reduces the size of the problem, especially if the operational scenarios are significantly shorter than the strategic periods. For instance, we can use a set of representative weeks or even days to represent strategic periods spanning several years.

The price we pay for this simplification is that there is no link between operational scenarios in consecutive strategic periods, as seen in Fig. 1. Consequently, the approach does not seemingly allow modelling of storages spanning multiples strategic periods. Strømholm and Rolfsen (2021) partially addressed this issue by providing formulas for the inventory level at the end of strategic periods, but their approach does not consider how this affects the required storage capacities.

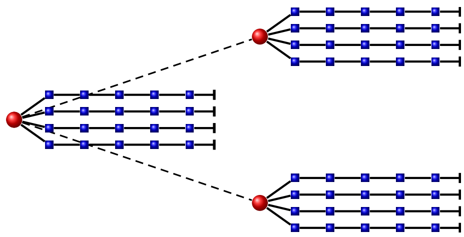

Figure 1: Scenario tree with two strategic periods and four operational scenarios in each strategic node.

In this paper, we build upon Strømholm and Rolfsen (2021) and show how to approximate storage capacities and inventory levels for different types of storages and scenarios. For this, we use the HyOpt optimization model developed at SINTEF (Kaut et al., 2019). HyOpt is limited to two-horizon trees, so we will from now on limit the presentation to this. More precisely, we assume scenario trees with structure as in Fig. 1, i.e., a tree consisting of strategic-decisions nodes, each with an attached set of operational scenarios. However, most of the presented logic should generalize to multi-horizon trees as well.

In the rest of the paper, we first present the most important parts of the HyOpt model in Section 1, before we proceed to the storage modelling in Section 2. Finally, we illustrate and test the presented approach in Section 3.

1 Relevant aspects of the HyOpt model

In this section, we present parts of the HyOpt model relevant to our paper. For the complete model formulation, see Kaut et al. (2019).

1.1 Notation

Indices

sn	strategic node, i.e., one node of the strategic scenario tree
sc	operational scenario
op	operational period

Parameters

$W^{\mathrm{SC}}_{sn,sc}$	weight of oper. scenario sc in strategic node sn
$T^{\mathrm{M}}_{sn,sc}$	time-scaling multiplier for scenario sc in strat. node sn
$M^{\mathrm{SC}}_{sn,sc}$	overall multiplier for scenario sc in strat. node sn
$\Delta T^{\mathrm{SP}}_{sn}$	duration of strategic period in strat. node sn
$\Delta T^{\mathrm{SC}}_{sc}$	duration of operational scenario sc
$\Delta T^{\mathrm{OP}}_{sc,op}$	duration of oper. period op in operational scenario sc

Table 1: Weights and multipliers of 3 operational scenarios. Here, $\Delta T^{\mathrm{d}}_{sc}$ is the duration of oper. scenario sc in days, and the last column shows how many days each scenario represents: days = $\Delta T^{\mathrm{SC,d}}_{sc} \times M^{\mathrm{SC}}_{sn,sc}$.

oper. sc.	$W^{\mathrm{SC}}_{sn,sc}$	$\Delta T^{\mathrm{d}}_{sc}$	$T^{\mathrm{M}}_{sn,sc}$	$M^{\mathrm{SC}}_{sn,sc}$	days
scen. 1	182/365	7	365/7	26	182
scen. 2	182/365	7	365/7	26	182
scen. 3	1/365	1	365	1	1
sum	1.0				365

1.2 Scenario-tree structure in HyOpt

In HyOpt, the strategic scenario tree consists of *strategic nodes*, each belonging to one *strategic period.* A *strategic scenario* is a path from the *root* of the strategic tree to one of the leaves.

Each of the strategic nodes has one or more *operational scenarios* attached to it. These scenarios can have different length and time resolution. Each scenario has also assigned *weight* $W^{\mathrm{SC}}_{sn,sc}$, denoting the time spent in the scenario as a fraction of the strategic period – implying $\sum_{sc} W^{\mathrm{SC}}_{sn,sc} = 1$ for all strategic nodes sn.

If the duration of an operational scenario does not equal the duration of the strategic period of the node it is attached to, we have to scale all effects of the scenario using a multiplier

$$M^{\mathrm{SC}}_{sn,sc} = W^{\mathrm{SC}}_{sn,sc} \times T^{\mathrm{M}}_{sn,sc} \tag{1}$$

where

$$T^{\mathrm{M}}_{sn,sc} = \Delta T^{\mathrm{SP}}_{sn} / \Delta T^{\mathrm{SC}}_{sc} \tag{2}$$

For example, Table 1 shows a situation where a strat. node in a 1-year strat. period has 3 operational scenarios, two weekly for normal situations and one daily representing an extreme day.

There, the last column shows how may days of the year are represented by each scenario. The table also shows that the *weights* are *not* the same as probabilities if the scenarios have different lengths: since the extreme day happens (on average) once per year while each of the normal scenarios happens 26 times, the probability of the scenarios happening are 1/53 and 26/53, respectively.

2 Storage handling

This section presents formulas for handling of storages in the HyOpt model. In particular, we demonstrate how to adjust the storage capacity, dependent on the duration of the operational scenarios and the time-scale of the storages. We also introduce a new notion of scenario *clusters* and show how this affects the presented formulas.

2.1 Relevant notation

Indices i storage node, i.e., node in the network modelling the system

Variables

$\mathbf{inv}_{i,sn,sc,op}$	inventory level of storage i at the end of oper. period (sn, sc, op)
$\mathbf{inv}^{\text{init}}_{i,sn,sc}$	inventory level at the start of oper. scenario (sn, sc)
$\mathbf{inv}^{\text{end}}_{i,sn,sc}$	inventory level at the end of oper. scenario (sn, sc)
$\mathbf{inv}^{\text{init}}_{i,sn}$	initial inventory level of storage i at strat. node sn
$\mathbf{inv}^{\text{end}}_{i,sn}$	final inventory level of storage i at strat. node sn

All the inventory variables are non-negative and limited by $\mathbf{cap}_{s,sn}$, the capacity available at strat. node sn:

$$0 \leq \mathbf{inv}_{i,sn,sc,op} \leq \mathbf{cap}_{s,sn}\,, \tag{3}$$

and correspondingly for the other variables.

2.2 Scaling storages from operational scenarios

While scaling with $M^{\text{SC}}_{sn,sc}$ works for costs, storage handling requires extra attention. There are (at least) two issues that have to be addressed: inventory level at the end of a strategic period, and the minimal storage capacity required for the presented scenarios.

If we denote by $\mathbf{inv}^{\Delta}_{i,sn,sc}$ the change of the inventory during oper. scenario sc,

$$\mathbf{inv}^{\Delta}_{i,sn,sc} = \mathbf{inv}^{\text{end}}_{i,sn,sc} - \mathbf{inv}^{\text{init}}_{i,sn,sc}\,,$$

then the overall change in the inventory at strategic node sn is

$$\mathbf{inv}^{\Delta}_{i,sn} = \sum_{sc} M^{\text{SC}}_{sn,sc} \times \mathbf{inv}^{\Delta}_{i,sn,sc}\,.$$

If we further assume that all operational scenarios start with the same inventory level,

$$\mathbf{inv}^{\text{init}}_{i,sn,sc} = \mathbf{inv}^{\text{init}}_{i,sn}\,,$$

then the final inventory is given by (Strømholm and Rolfsen, 2021)

$$\mathbf{inv}^{\text{end}}_{i,sn} = \mathbf{inv}^{\text{init}}_{i,sn} + \mathbf{inv}^{\Delta}_{i,sn}\,. \tag{4}$$

The above requirement that all operational scenarios within one strategic node start with the same inventory level corresponds to an interpretation of operational scenarios as a list of possibilities that occur randomly: since we do not know which scenarios will happen, we cannot adjust inventory levels in preparation for them.

This requirement has the additional benefit of aiding in the dimensioning of storage. To understand why, let us consider again the three scenarios from Table 1 and assume that the inventory changes of storage i in the three scenarios are (10, -9, -25), respectively. As long as there is a cost associated with the installed storage capacity, and assuming no additional constraints, the model will select the minimal capacity that can satisfy

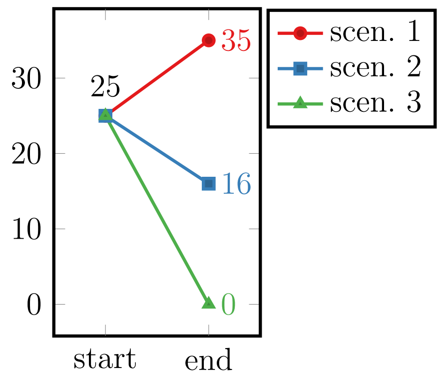

Figure 2: Storage level changes with common starting point

Eq. (3) in all scenarios. In our case, this results in the inventory levels presented in Fig. 2, requiring storage capacity of 35[1].

However, it is important to realize that this storage capacity is an underestimate. Since both 'normal' scenarios (scen. 1 and scen. 2) occur 26 times during the strategic period, we can expect to see them happening multiple times in a row, assuming they occur randomly. This increases the required capacity significantly.

For example, if we start with two consecutive occurrences of scen. 1, we end up with an inventory level of $25 + 2 \times 10 = 45$, requiring $\mathbf{cap}_{i,sn} \geq 45$. At the same time, a single occurrence of scen. 2 preceded or followed by scen. 3 would result in an infeasible inventory level $25 - 25 - 9 = -9$. To avoid the infeasibility, the initial inventory level would have to increase to $25+9 = 34$, requiring $\mathbf{cap}_{i,sn} \geq 44 \ldots$ and $\mathbf{cap}_{i,sn} \geq 45+9 = 54$ if we wanted to consider the first sequence as well. In both cases, allowing more repetitions would increase the required capacity even further.

On the other hand, if we for some reason knew that the two normal scenarios alternate, then none of the repetitions could happen and $\mathbf{cap}_{i,sn} = 44$ would be sufficient to cover all allowed permutations. This shows that the magnitude of the capacity underestimation is necessarily case-dependent.

2.3 Operational scenarios in sequence

In the preceding sections, operational scenarios were assumed to occur randomly. However, this assumption does not hold when the scenarios represent events that occur in sequence, such as seasons within a year. This is a common approach in long-term models, as demonstrated in Skar et al. (2016) or Strømholm and Rolfsen (2021).

In such cases, we do *not* expect the scenarios to occur randomly throughout the strategic period. Instead, if we have one scenario per season, its probability is 100% during that season and 0% elsewhere. It follows that the scenarios occur multiple times in a row – 13 times in case of weekly scenarios representing seasons. This has a significant impact on the required storage capacity.

[1] Note that this assumes that the inventory levels change monotonously from the start to the end of each operational period, which is normally not the case. We will address this issue later.

Table 2: Weights and multipliers of 5 operational scenarios. Again, $\Delta T^{\mathrm{d}}_{sc}$ is the duration of oper. scenario sc in days and column 'days' shows how many days each scenario represents. In addition, the last column shows the inventory changes scaled to the whole strategic period, $\mathbf{inv}^{\Delta\mathrm{SP}}_{i,sn,sc} = T^{\mathrm{M}}_{sn,sc} \times \mathbf{inv}^{\Delta}_{i,sn,sc}$.

oper. sc.	$W^{\mathrm{SC}}_{sn,sc}$	$\Delta T^{\mathrm{d}}_{sc}$	$T^{\mathrm{M}}_{sn,sc}$	$M^{\mathrm{SC}}_{sn,sc}$	days	$\mathbf{inv}^{\Delta}_{i,sn,sc}$	$\mathbf{inv}^{\Delta\mathrm{SP}}_{i,sn,sc}$
winter	91/365	7	365/7	13	91	-10	-130
spring	91/365	7	365/7	13	91	15	195
summer	91/365	7	365/7	13	91	-5	-65
autumn	91/365	7	365/7	13	91	0	0
bad day	1/365	1	365	1	1	-5	-5
sum	1.0				365		-5

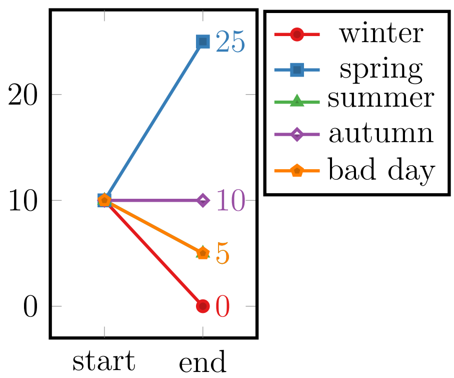

Figure 3: Storage level changes with oper. scenarios from Table 2 interpreted as random events

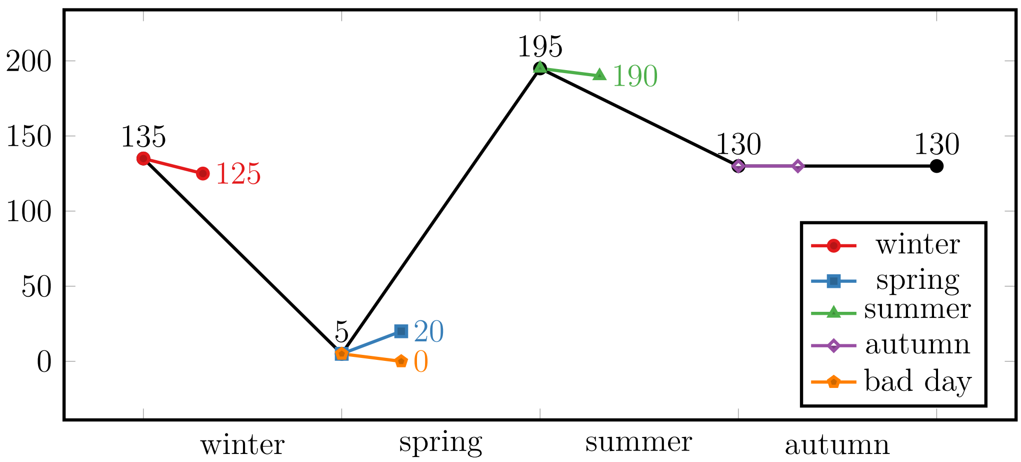

Figure 4: Storage level changes with oper. scenarios from Table 2, interpreted as four seasons in sequence with the 'bad day' scenario assigned to spring.

To illustrate this, consider scenarios from Table 2. If we treat them as random events, then the overall storage change is -5, and the required capacity is 25, as shown in Fig. 3. To let the scenarios represent the sequence of seasons instead, we first have to decide what to the with the 'bad day' scenario. We have assigned it to spring, which is therefore represented by two scenarios and its overall inventory change is 190. The inventory-level changes for each scenario are shown in the last column of Table 2 and the whole annual profile in Fig. 4. The figure shows that the minimal storage capacity required in this case is 195 – almost eight times more than in the original approach. Note that the total inventory-level change is not affected by the grouping and remains at -5.

2.3.1 Extra model notation

To implement the scenario groups in the model, we need to introduce extra notation:

Parameters and sets

$\boldsymbol{G}^{\mathrm{SC}}_{sn}$	ordered list of scenario groups at strategic node sn
$\boldsymbol{S}^{\mathrm{G}}_{sn,g}$	set of operational scenarios in group $g \in \boldsymbol{G}^{\mathrm{SC}}_{sn}$
$\Delta T^{\mathrm{G}}_{sn,g}$	duration of scenario group $g \in \boldsymbol{G}^{\mathrm{SC}}_{sn}$
$W^{\mathrm{SC,G}}_{sn,g,sc}$	relative weight of scenario sc within scen. group $g \in \boldsymbol{G}^{\mathrm{SC}}_{sn}$
$T^{\mathrm{M,G}}_{sn,g,sc}$	time scaling factor for scenario sc within scen. group $g \in \boldsymbol{G}^{\mathrm{SC}}_{sn}$
$M^{\mathrm{SC,G}}_{sn,g,sc}$	multiplier for scenario sc within scen. group $g \in \boldsymbol{G}^{\mathrm{SC}}_{sn}$
$\boldsymbol{g}^{\mathrm{first}}_{sn}$	first scen. group in strat. node sn
$\boldsymbol{g}^{\mathrm{last}}_{sn}$	last scen. group in strat. node sn

Variables

$\mathbf{inv}^{\mathrm{G,init}}_{i,sn,g}$	inventory level at the start of scenario group $g \in \boldsymbol{G}^{\mathrm{SC}}_{sn}$
$\mathbf{inv}^{\mathrm{G,end}}_{i,sn,g}$	inventory level at the end of scenario group $g \in \boldsymbol{G}^{\mathrm{SC}}_{sn}$

The inventory-level variables are again non-negative and limited by $\mathbf{cap}_{i,sn}$. In addition, the new entities are connected by the following relations:

$$
\begin{aligned}
W^{\mathrm{SC,G}}_{sn,g,sc} &= W^{\mathrm{SC}}_{sn,sc} \;/ \sum_{\tilde{sc}\in g} W^{\mathrm{SC}}_{sn,\tilde{sc}} \\
T^{\mathrm{M,G}}_{sn,g,sc} &= \Delta T^{\mathrm{G}}_{sn,g} / \Delta T^{\mathrm{OP}}_{sc,op} \\
M^{\mathrm{SC,G}}_{sn,g,sc} &= W^{\mathrm{SC}}_{sn,sc} \times T^{\mathrm{M,G}}_{sn,g,sc} \\
\mathbf{inv}^{\mathrm{G,init}}_{i,sn,\boldsymbol{g}^{\mathrm{first}}_{sn}} &= \mathbf{inv}^{\mathrm{init}}_{i,sn} \\
\mathbf{inv}^{\mathrm{init}}_{i,sn,sc} &= \mathbf{inv}^{\mathrm{G,init}}_{i,sn,g} \\
\mathbf{inv}^{\mathrm{G,end}}_{i,sn,g} &= \mathbf{inv}^{\mathrm{G,init}}_{i,sn,g} + \sum_{sc\in \boldsymbol{S}^{\mathrm{G}}_{sn,g} g} M^{\mathrm{SC,G}}_{sn,g,sc} \times \mathbf{inv}^{\Delta}_{i,sn,sc} \\
\mathbf{inv}^{\mathrm{G,init}}_{i,sn,g} &= \mathbf{inv}^{\mathrm{G,end}}_{i,sn,g-1} \text{ if } g-1 \in \boldsymbol{G}^{\mathrm{SC}}_{sn} \text{ else } \mathbf{inv}^{\mathrm{init}}_{i,sn,sn} \\
\mathbf{inv}^{\mathrm{end}}_{i,sn} &= \mathbf{inv}^{\mathrm{G,end}}_{i,sn,\boldsymbol{g}^{\mathrm{last}}_{sn}}
\end{aligned}
$$

Table 3: Alternative scenarios for the summer group. Analogously to Table 2, $\mathbf{inv}^{\Delta\mathrm{G}}_{i,sn,g,sc} = T^{\mathrm{M,G}}_{sn,g,sc} \times \mathbf{inv}^{\Delta}_{i,sn,sc}$ is the inventory-level change during the scenario, scaled to the duration of the group. The last column shows the maximum inventory level reached within each oper. scenario, relative to its starting level, $\mathbf{inv}^{\max\Delta}_{i,sn,sc} = \max_{op} \mathbf{inv}_{i,sn,sc,op} - \mathbf{inv}^{\mathrm{init}}_{i,sn,sc}$.

oper. scen.	$W^{\mathrm{SC}}_{sn,sc}$	$W^{\mathrm{SC,G}}_{sn,g,sc}$	$T^{\mathrm{M,G}}_{sn,g,sc}$	$M^{\mathrm{SC,G}}_{sn,g,sc}$	$\mathbf{inv}^{\Delta}_{i,sn,sc}$	$\mathbf{inv}^{\Delta\mathrm{G}}_{i,sn,g,sc}$	$\mathbf{inv}^{\max\Delta}_{i,sn,sc}$
summer-1	42/365	6/13	13	6	5	30	10
summer-2	42/365	6/13	13	6	−12	−72	0
summer-3	7/365	1/13	13	1	−23	−23	0

where $g \in \boldsymbol{G}^{\mathrm{SC}}_{sn}$ and $sc \in \boldsymbol{S}^{\mathrm{G}}_{sn,g}g$ in all constraints where they appear. Since we treat scenarios within each group as random events, we force them to start from a common initial level. The last constraints ensures that the installed capacity is big enough to handle also the final inventory levels in the sequence.

2.3.2 Repeated scenarios

Now consider what would happen if we replaced the current summer scenario by the three weekly scenarios presented in Table 3. There, the last column shows the maximum inventory level during the scenario, relative to the start of the scenario: it means that inventory in scenario 'summer-1' first increases by 10 units and then drops, so the overall change is +5 units.

Since $30 - 72 - 23 = 65$, the total inventory change during summer is unchanged, so the overall inventory profile would remain the same as in Fig. 4 – except that scenario 'summer-1' would increase the required storage capacity from 195 to 205, to handle the maximum level attained during the scenario.

But what if this scenario repeated two or more times in row? This would require even higher capacity. This means that we have to decide how many repetitions do we want to take into account. The maximum number of repetitions is $M^{\mathrm{SC,G}}_{sn,g,sc}$; for 'summer-1', this means 6 occurrences in a row – but this happens with probability of only $(W^{\mathrm{SC,G}}_{sn,g,sc})^6 = (6/13)^6 = 0.97\%$, so considering it might be too conservative. Instead, we have chosen to set a limit on the probability we want to consider, denoted by P^{R}. This limits the number of consecutive occurrences to

$$seq^{\mathrm{G}}_{sn,g,sc} = \lfloor \ln(P^{\mathrm{R}}) / \ln(W_{sn,s}/W^{\mathrm{g}}_{sn,g}) \rfloor \,, \tag{5}$$

bounded by 1 from below and $M^{\mathrm{SC}}_{sn,sc}$ from above.

In our case, we use $P^{\mathrm{R}} = 5\%$. For 'summer-1', this means considering only $seq^{\mathrm{G}}_{sn,g,sc} = \lfloor \ln(0.05)/\ln(6/13) \rfloor = 3$ occurrences. In the initial one, the inventory level goes from 195 to 200, with a maximum at 205. The first repetition starts with inventory level of $195 + 5 = 200$ and the second with $195 + 2 \times 5 = 205$, with a maximum reaching $205 + 2 \times 5 = 215$. This would thus become the new required capacity.

To add this capacity requirement to the model, we just need to handle the last repetition, since the inventory levels change linearly throughout the sequence and the initial occurrence is already considered in the model and therefore satisfies Eq. (3). The last occurrence has inventory levels shifted by $(seq^{\mathrm{G}}_{sn,g,sc}-1)\,\mathbf{inv}^{\Delta}_{i,sn,sc}$, so its version of Eq. (3) is

$$0 \le \mathbf{inv}_{i,sn,sc,op} + (seq^{\mathrm{G}}_{sn,g,sc} - 1)\,\mathbf{inv}^{\Delta}_{i,sn,sc} \le \mathbf{cap}_{i,sn}\,, \tag{6}$$

for all operational scenarios sc with $seq^{\mathrm{G}}_{sn,g,sc} > 1$.

2.3.3 Multi-year scenarios

Until now, we have implicitly assumed annual strategic periods where the sequence of scenario groups (seasons) covers the whole strategic period. If the duration increases to two years, then the overall storage-level change doubles to -10 (-5 each year), but this is clearly not the case for the inventory levels. Indeed, if the overall storage-level change were zero, then the second year would be a copy of the first, and the required storage capacity would not change at all.

This can be addressed analogously to the scenario repetition, using

$$0 \le \mathbf{inv}_{i,sn,sc,op} + (\lfloor(\Delta T^{\mathrm{SP}}_{sn} / \sum_{g\in \boldsymbol{G}^{\mathrm{SC}}_{sn}} \Delta T^{\mathrm{G}}_{sn,g}\rfloor - 1)\,\mathbf{inv}^{\Delta\mathrm{G}}_{i,sn} \le \mathbf{cap}_{i,sn}\,, \tag{7}$$

where $\mathbf{inv}^{\Delta\mathrm{G}}_{i,sn}$ is the total inventory-level change in all the scenario groups,

$$\mathbf{inv}^{\Delta\mathrm{G}}_{i,sn} = \sum_{g\in \boldsymbol{G}^{\mathrm{SC}}_{sn}} \sum_{sc\in \boldsymbol{S}^{\mathrm{G}}_{sn,g}} \mathbf{inv}^{\Delta\mathrm{G}}_{i,sn,g,sc}\,.$$

We also need constraints that combine Eqs. (6) and (7), to accommodate repeated scenarios within repeated scenarios groups:

$$0 \le \mathbf{inv}_{i,sn,sc,op} + (seq^{\mathrm{G}}_{sn,g,sc} - 1)\,\mathbf{inv}^{\Delta}_{i,sn,sc} + (\lfloor \Delta T^{\mathrm{SP}}_{sn} / \Delta T^{\mathrm{G}}_{sn,g}\rfloor - 1)\,\mathbf{inv}^{\Delta}_{i,sn} \le \mathbf{cap}_{i,sn} \tag{8}$$

Note that Eq. (8) alone is insufficient, and we require all Eqs. (6) to (8). To illustrate this, let us assume that we had a scenario with a storage-level change of 10 and $seq^{\mathrm{G}}_{sn,g,sc}=4$, in a two-year strategic period with an overall inventory-level change of -30. Then

$$(seq^{\mathrm{G}}_{sn,g,sc} - 1)\,\mathbf{inv}^{\Delta}_{i,sn,sc} + (\lfloor \Delta T^{\mathrm{SP}}_{sn} / \Delta T^{\mathrm{G}}_{sn,g}\rfloor - 1)\,\mathbf{inv}^{\Delta}_{i,sn} = 3\times 10 + 1 \times -30 = 0\,,$$

so Eq. (8) would not have any effect, but Eqs. (6) and (7) would still be needed.

2.4 Initial and final inventory levels

Without any limitations on initial and final inventory levels, the model will likely choose to start each operational scenario with full storage and end it with an empty one, effectively obtaining full storage for free. There are two common approaches to address this issue: assigning a value to the inventory or requiring that the storage ends with the same level as it started. In HyOpt, we use the latter approach with several *time scopes* for the storage-level looping, to distinguish between short- and long-term storages:

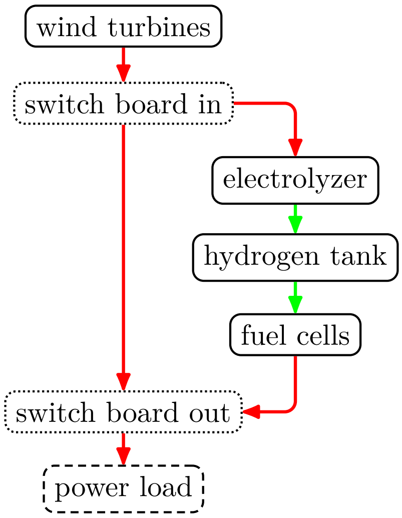

Figure 5: HyOpt representation of of the test case from Section 3. Red arrows represent flow of power, green arrows flow of hydrogen. Nodes whose capacity is being optimized are denoted by solid-line borders.

oper-scenario, where we require that the inventory level loops back in every operational scenario, i.e., $\mathbf{inv}^{\Delta}_{i,sn,sc} = 0$. This is typically used for batteries, where each scenario constitutes a daily or weekly 'schedule' with charging overnight.

scen-group, where we require that the inventory level loops back within each scenario group, i.e., $\mathbf{inv}^{\Delta G}_{i,sn,g} = 0$. If the scenario groups represent seasons, this would fit storages that cycle in weeks or a few months, such as medium-size hydrogen-storage systems or smaller hydro reservoirs.

strat-node, where we require that the inventory-level at the end of a strategic node is equal to the initial level, i.e., $\mathbf{inv}^{\Delta}_{i,sn} = 0$. With yearly strategic time periods, this corresponds to inter-seasonal storage like large hydrogen storage or mid-sized hydro reservoirs.

overall, where we require that the final inventory level at the end of all strategic nodes in the last strat. periods finish on a level that is equal to the initial level in the first strat. period. This corresponds to large hydro reservoirs.

Note that a storage node with scope 'strat-node' or shorter would not, by definition, need Eqs. (7) and (8) and with 'oper-scenario' scope would not need Eq. (6) either.

3 Test case

We consider a simple case inspired by the LowEmission project[2], involving electrification of an offshore installation using wind turbines combined with energy storage, in our case a hydrogen-base system consisting of an electrolyzer, hydrogen tank, and fuel cells. In

[2]See https://www.sintef.no/projectweb/lowemission/

Table 4: Model time structures considered in the test case. $T^{\text{hor,y}}$ is the overall time horizon, in years, $|\boldsymbol{\mathcal{T}}^{\text{SP}}|$ is the number of strategic periods, and $\Delta T^{\text{SP,y}}_{sn}$ shows the duration of the strat. periods, in years.

struct.	$T^{\text{hor,y}}$	$\lvert\boldsymbol{\mathcal{T}}^{\text{SP}}\rvert$	$\Delta T^{\text{SP,y}}_{sn}$	wind-data year
$\mathbf{1 \times 1}$	1	1	1	2018
$\mathbf{1 \times 5}$	5	1	5	2018–2022
$\mathbf{5 \times 1}$	5	5	$(1, 1, 1, 1, 1)$	2018–2022

HyOpt, this gives a network structure as in Fig. 5, with 'switch board' nodes added for modelling convenience.

The wind-production data are based on actual wind-speed measurements from the Ekofisk field in the North Sea, covering the period from 2018 to 2022. These measurements are converted to wind-production capacity factor using a production profile for Vestas *V164/8000* wind turbine, obtained from the Open Energy Platform[3]. We assume a fixed power load of 20 MW. All costs and performance data are taken from the open HyOpt test case available at `https://gitlab.sintef.no/open-hyopt/test-case-1`.

3.1 Case variants

We test the model with three time structures, 1×1, 1×5, and 5×1, presented in Table 4. They all use hourly time resolution. We use astronomical definitions of seasons, where each season covers three months and winter starts on December 1. Consequently, our 'year' goes from December to November in order to make it a sequence of four complete seasons. In other words, when we say, for example, 2018, the actual interval is 2017-12-01 to 2018-11-30.

For each time structure, we test the effect of the following operational variants:

full, with one operational scenario per strat. period, covering the full length of the period.

mean, with four operational scenarios per strat. period, with length of one week. Each scenario represents one season and is selected as the week in the data with its average capacity factor closest to the average for the season (in the given time interval).

mean+min, with eight operational scenarios per strat. period, with length of one week. For each season, we select the week with the smallest above-average capacity factor, and the week with the smallest capacity factor as the worst case. The scenario weights are selected so that the weighted average is equal to the average capacity factor of the season.

Note that this simple selection approach is applicable only to one-dimensional randomness (in our case, the wind-production capacity factor). With multidimensional

[3] See `https://openenergy-platform.org/`.

Table 5: Number of weeks in operational scenarios and the size reduction compared to the full model, for the tested time structures and model variants

	1×1		1×5		5×1	
case variant	weeks	reduct.	weeks	reduct.	weeks	reduct.
full	52	–	260	–	260	–
mean	4	13	4	65	20	13
mean+min	8	6.5	8	32.5	40	6.5

randomness, or a need for more scenarios, a different approach would be needed. A popular option is to use some clustering approach, for example k-medoids (Kaufman and Rousseeuw, 1990, Chapter 2) – a method similar to k-means, but using actual data points as cluster centres. For more methods see, for example, Bounitsis et al. (2022) or Kaut (2021).

Moreover, we use two versions of the multi-scenario variants:

fan, where all scenarios are interpreted as random events, so we have a single 'fan' of scenarios.

groups, with scenarios grouped by season.

The scenario-tree sizes of the variants are summarized in Table 5. There, we can see that the size reduction from using the weekly scenarios varies from 6.5 times to 32.5 times, compared to the case with a full time sequence.

3.2 Results

All test variants we solved using FICO™ Xpress Solver v 8.14, on a laptop with Intel® Core™ i7-7600U CPU at 2.80 GHz and 16 GB of RAM. Results for the one-year case are shown in Figs. 6 to 8. The left panel show that the solution times of the scenario-based versions are significantly smaller than the 'full' model, with a speed-up matching the problem-size reduction presented in Table 5. In particular, the speed-up ranges from 5–6 in the worst case of variant '5×1' to over 100 for '1×5'.

Regarding the objective function, i.e., the total costs, we can make the following observations:

- 'mean+min' variants are more costly than 'mean'. This is natural, as they have to take into account the extreme scenarios.
- 'groups' variants are more costly than 'fan'. This shows that the dynamics introduced by the grouping works, forcing the model to tackle multiple consequent occurrences of scenarios representing one season.
- the 'mean+min_groups' variants can be more costly than 'full'. This could be because the scenario variants force the model to handle both the 'min' scenario and multiple occurrences of the 'mean' scenarios at the same time (at the start of

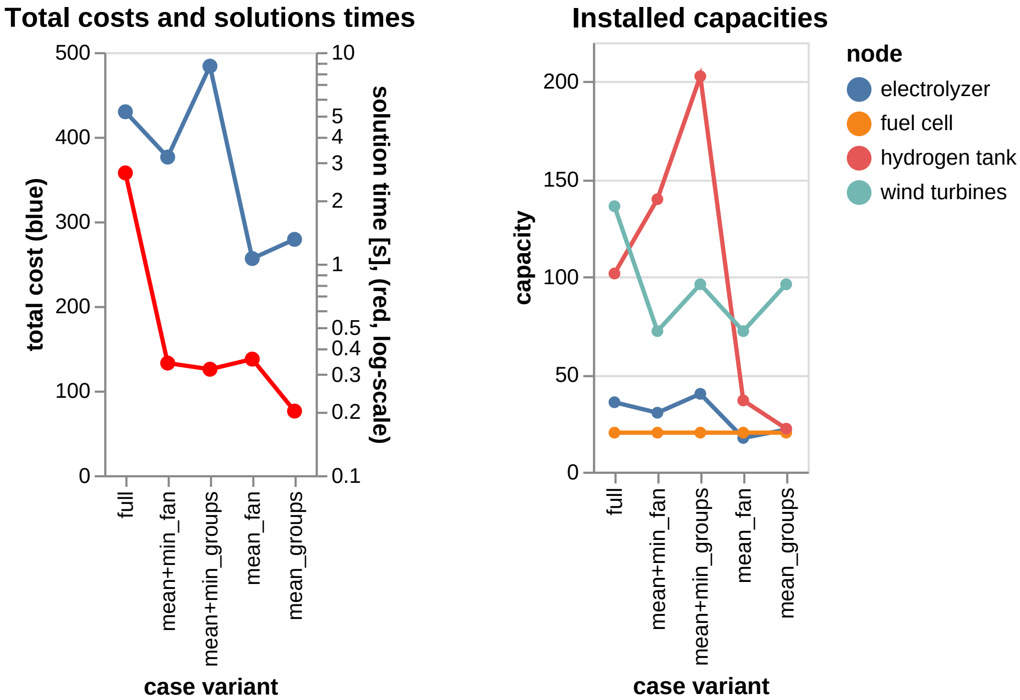

Figure 6: Results for time struct. '1 × 1': one strat. period one year long

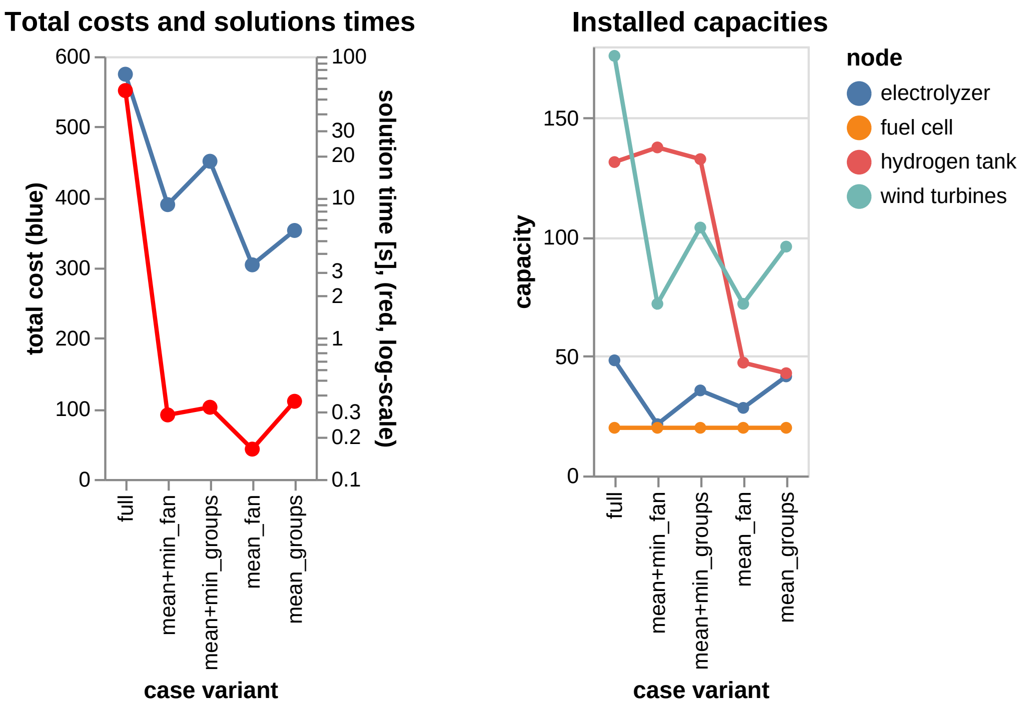

Figure 7: Results for time struct. '1 × 5': one strat. period 5 years long

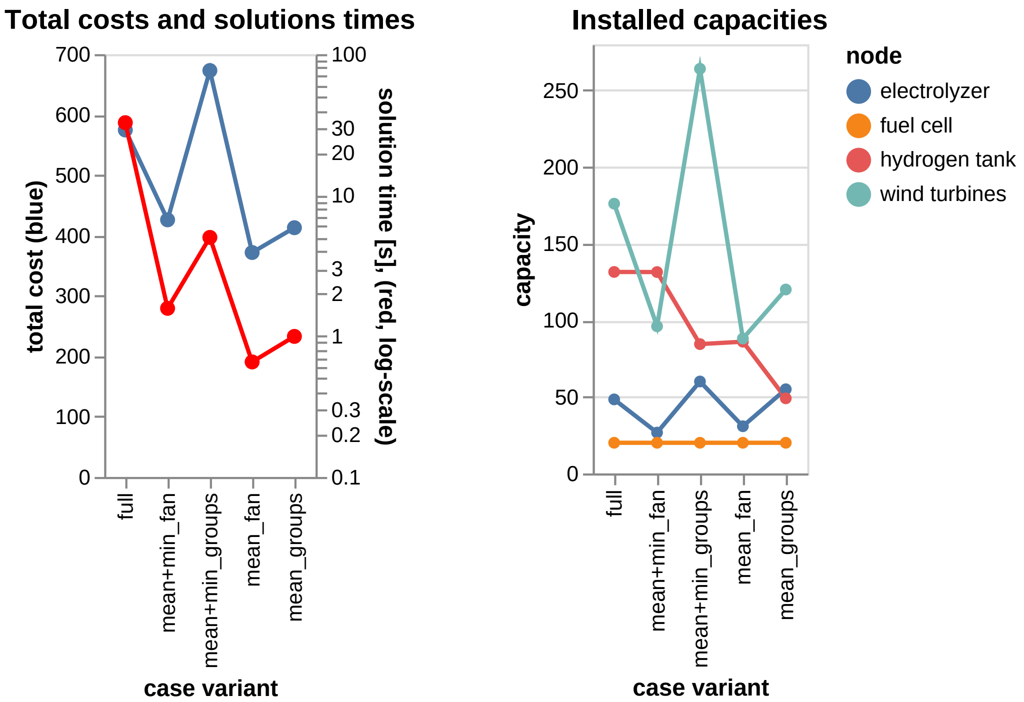

Figure 8: Results for time struct. '5 × 1': 5 yearly strat. periods

each season), while the actual sequence in the 'full' variant might turn out to be easier to handle.

The last point is important: Eq. (6) forces the model to take into account multiple repetitions of the involved scenarios, up to the specified probability. Therefore, the scenarios are designed to make the model tackle many different permutations of the involved scenarios. The 'full' variant, on the other hand, allows the model to adapt to the one scenario (historical data) included in the tree. Consequently, the scenario selection might produce more expensive solutions.

Note that it does not mean that scenarios produce *better* solutions. Consider, for example, a 12-week season represented by 2 weekly scenarios with equal weight, where the inventory level increases by 10 in the first scenario and decreases by 10 in the second one. With 5% limit on scenario repetitions, we get $\text{rep}_{s,sn,g} = 4$, so the model will have to tackle 4 consequent occurrences of either scenarios. This will force the storage capacity to be at least 80, to allow the inventory to go both 40 units up and 40 down, assuming it starts at 40. If the two scenarios in reality tend to alternate and never occur more than twice in a row, the actual storage-capacity requirement would be lower and so would be the total costs.

The right panels of Figs. 6 to 8 show where the costs come from. We can see that increased wind variability is being addressed by a combination of increased hydrogen storage and increased wind-production capacity. Specifically, in '1 × 1', the most expensive variant is the one with largest storage, while in '1 × 5' and '5 × 1' it is the one with the largest wind park. This interplay between the two components makes it difficult to draw conclusions about the effects of the scenario structure on either of them alone. This is a well-known feature of stochastic-programming models, where the objective functions tend to be rather flat, with several different strategies having close to optimal objective value.

4 Conclusions

This paper presents a way to model long-term storages within the multi-horizon modelling paradigm. This removes one of the major obstacles for using this modelling approach and makes it applicable to a wider selection of problems.

The presented formulation is implemented in the HyOpt optimization model, freely available at `https://gitlab.sintef.no/open-hyopt/`. The test case from Section 3 is available at `https://gitlab.sintef.no/open-hyopt/test-case-2` and includes scripts for solving it using PyHyOpt, a python interface to HyOpt, also available from the HyOpt page.

Acknowledgements

The work presented in this paper has been supported by the LowEmission project, the Research Council of Norway project number 296207, `https://prosjektbanken.forskningsradet.no/project/FORISS/296207`.

References

Georgios L. Bounitsis, Lazaros G. Papageorgiou, and Vassilis M. Charitopoulos. Data-driven scenario generation for two-stage stochastic programming. *Chemical Engineering Research and Design*, 187:206–224, nov 2022. doi: 10.1016/j.cherd.2022.08.014.

Leonard Kaufman and Peter J. Rousseeuw. *Finding Groups in Data*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., March 1990. doi: 10.1002/9780470316801.

Michal Kaut. Scenario generation by selection from historical data. *Computational Management Science*, 18(3):411–429, jun 2021. doi: 10.1007/s10287-021-00399-4.

Michal Kaut, Truls Flatberg, and Miguel M. Ortiz. The HyOpt model: Input data and mathematical formulation. techreport 2019:01439, SINTEF, 2019. URL `https://hdl.handle.net/11250/2643389`.

Christian Skar, Gerard Doorman, Gerardo Pérez-Valdés, and Asgeir Tomasgard. A multi-horizon stochastic programming model for the European power system. techreport 2/2016, CenSES, 2016. URL `https://www.ntnu.no/censes/working-papers`.

Lars Skaugen Strømholm and Raag August Sandal Rolfsen. Flexible hydrogen production : a comprehensive study on optimizing cost-efficient combinations of production and storage capacity to exploit electricity price fluctuations. mathesis, Norwegian School of Economics (NHH), 2021. URL https://hdl.handle.net/11250/2770501.