Accurately modeling long-term storage with minimum representative hours in large-scale renewable energy systems

Energy System Optimization Workshop @TU Graz

Jacob Mannhardt, Lukas Kunz, Giovanni Sansavini

28.11.2025

ETH zürich

RRE Reliability & Risk Engineering

Accurately modeling long-term storage with minimum representative hours in large-scale renewable energy systems

Make TSA unsexy again!

Energy System Optimization Workshop @TU Graz

Jacob Mannhardt, Lukas Kunz, Giovanni Sansavini

28.11.2025

The curse of high dimensionality

Our models have high dimensionality:

Many considered years [$\mathcal{O}(10)$]
and $\boxed{\text{time steps } [\mathcal{O}(1000)]}$

Many nodes [$\mathcal{O}(100)$]

Many energy carriers [$\mathcal{O}(10)$]
and technologies [$\mathcal{O}(100)$]

Complex technology descriptions
(binary operations, nonlinear relationships)

Scenarios and uncertainty handling

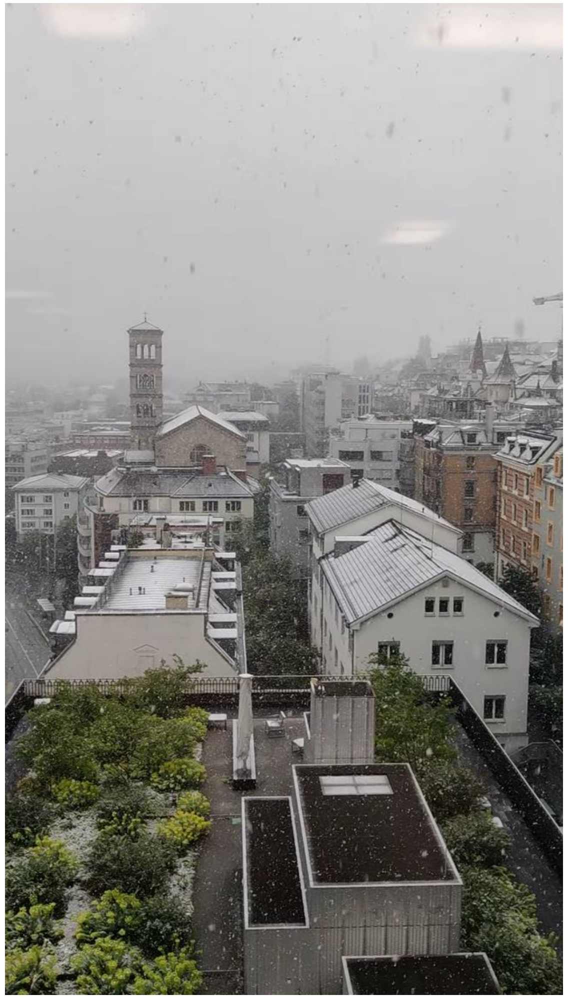
The complexities we are discussing in this workshop!

We need to reduce complexity without sacrificing (much) accuracy:

The most established technique is
time series aggregation (TSA)

Utilizing the fact that many hours in the year
show similar conditions

Why do we want to do time series aggregation?



April 18th, 2024, at 10 am

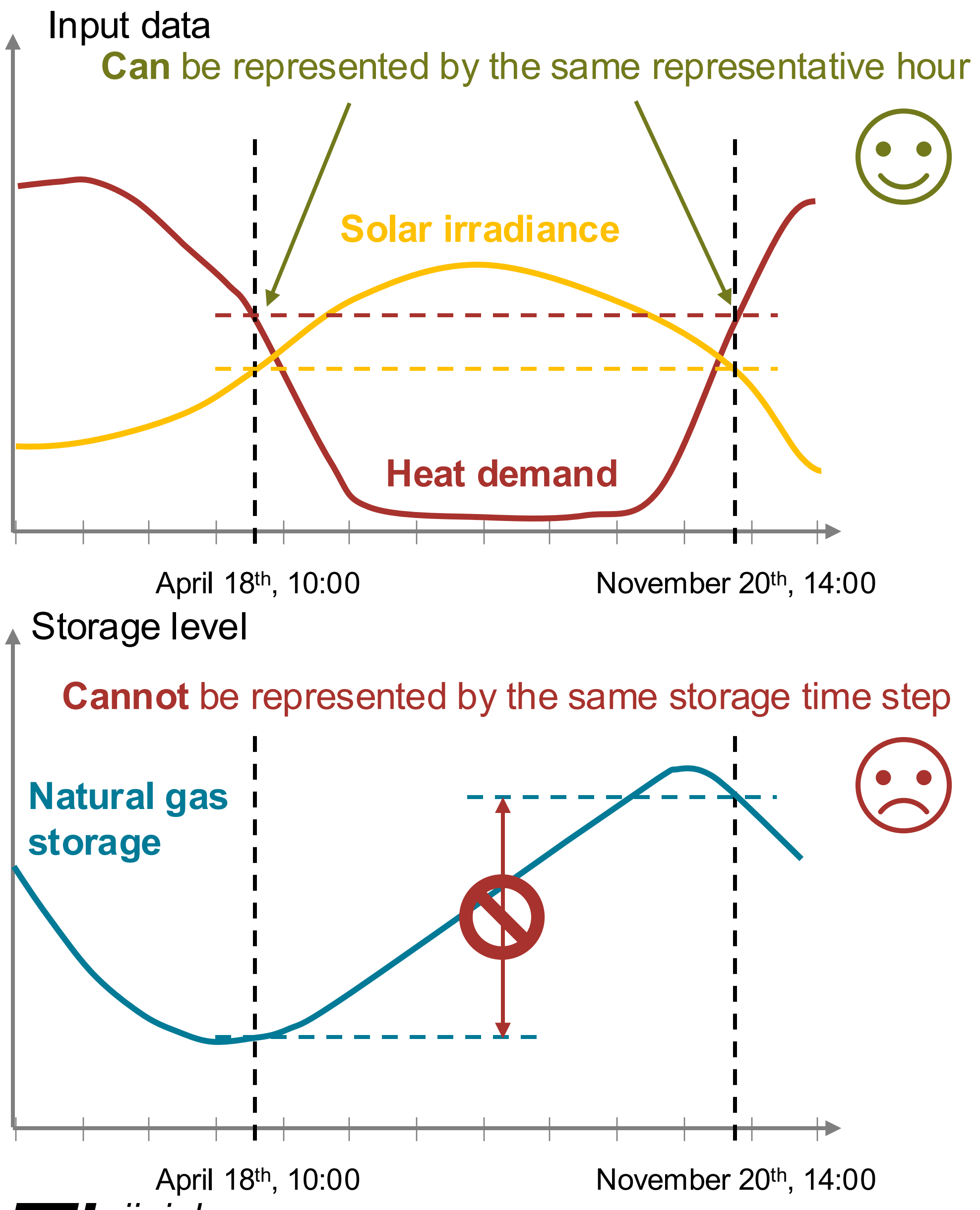
November 20th, 2025, at 2 pm

We need to reduce complexity without sacrificing (much) accuracy:

The most established technique is
time series aggregation (TSA)

Utilizing the fact that many hours in the year show similar conditions

What is the issue with time series aggregation?



Storage representation problem:
The storage level is time-coupled, i.e., the storage level depends on the previous time step

Let's formalize the problem:
The storage level is formulated as:

$$L_{s,t} = L_{s,t-1}(1-\varphi_s) + \underbrace{\underline{\eta}_s \underline{H}_{s,t} - \frac{\overline{H}_{s,t}}{\overline{\eta}_s}}$$

Storage level in **current** time step

Storage level in **previous** time step (including self-discharge)

Net charging flow (can contain more terms, e.g., spillage etc.) $\Delta H_{s,t}$

TSA loses chronology of time steps
→ "previous" time step does not exist anymore

Storage representation must rebuild the storage level chronology

What do we want to do today?

Compare clustering **representative days** (RD) and **representative hours** (RH)

Highlight the **limitations** of current RH storage representation methods

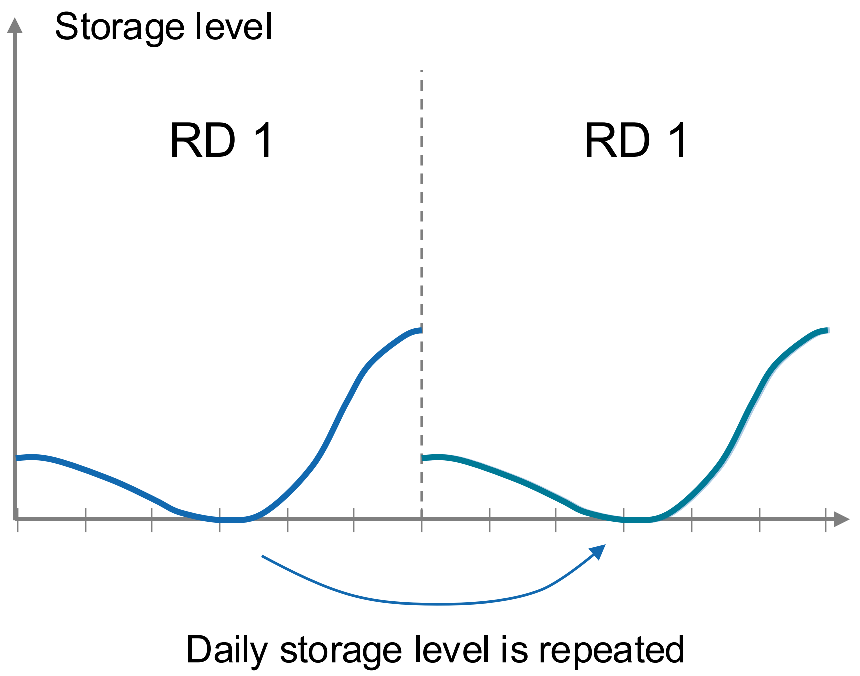
Present the novel RH storage representation method for **large-scale multi-storage** energy systems

Benchmark **existing methods** in terms of accuracy vs computational complexity

Representative days (RD) vs representative hours (RH)

The most common way to solve the storage representation problem [1]:

- Keep daily chronology of the clustered time steps = representative days (RD)
- For long(er)-term storage: Superpose daily storage levels with the storage levels of the RD



However, …
Clustering RD requires many representative time steps to approximate the input data well!

Representative hours (RH) require fewer time steps for the same approximation quality

Why are RH storage representation methods underrepresented?

Until now, the storage level cannot be reconstructed accurately and efficiently

[1] Kotzur, L., Markewitz, P., Robinius, M. & Stolten, D. Time series aggregation for energy system design: Modeling seasonal storage. *Applied Energy* (2018).

Limitations of current RH storage representation methods

Reconstructing the storage level chronology is especially challenging for non-chronological RH

Fully resolved storage level [1]

- Idea: Resolve the storage level with 8760 time steps per year → Fully resolved storage level
- Chronology: yes!
- Limitation: Introduces many additional variables and constraints → computationally challenging

State-space transitions [2]

- Idea: Track the storage level transition from one RH to another with central differences
- Chronology: yes!
- Limitation: Central differences approximate and smoothen storage level

Chronological RH [3]

- Idea: Cluster chronological RH → storage level time steps are the operational time steps
- Chronology: yes!
- Limitation: Strongly smoothens daily variability

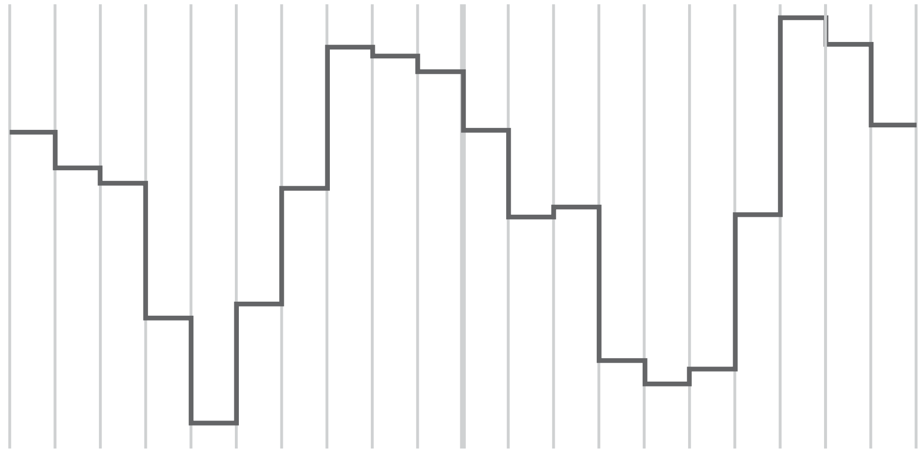
Proposed reduced RH storage representation

Idea: Resolve the storage level but only keep the necessary storage level time steps

[1] Gabrielli, P., Gazzani, M., Martelli, E. & Mazzotti, M. Optimal design of multi-energy systems with seasonal storage. *Applied Energy* (2018).
[2] Wogrin, S., Galbally, D. & Reneses, J. Optimizing Storage Operations in Medium- and Long-Term Power System Models. *IEEE Transactions on Power Systems* (2016).
[3] Pineda, S. & Morales, J. M. Chronological Time-Period Clustering for Optimal Capacity Expansion Planning With Storage. *IEEE Transactions on Power Systems* (2018).

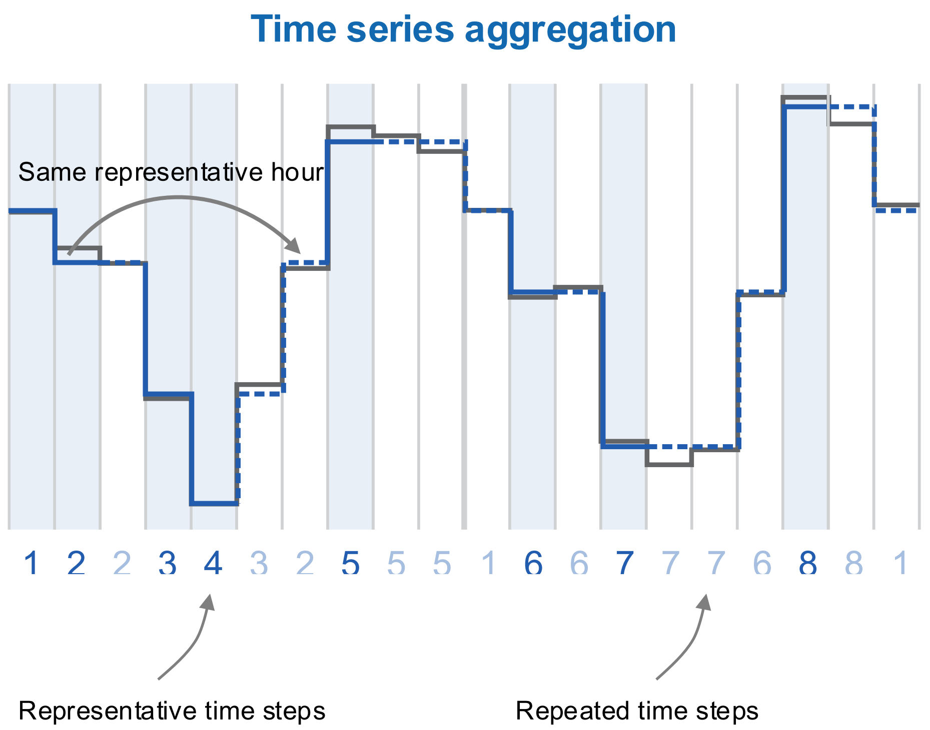
How to reduce the number of storage level variables and constraints without losing any information

Let’s take a generic input time series with 20 time steps



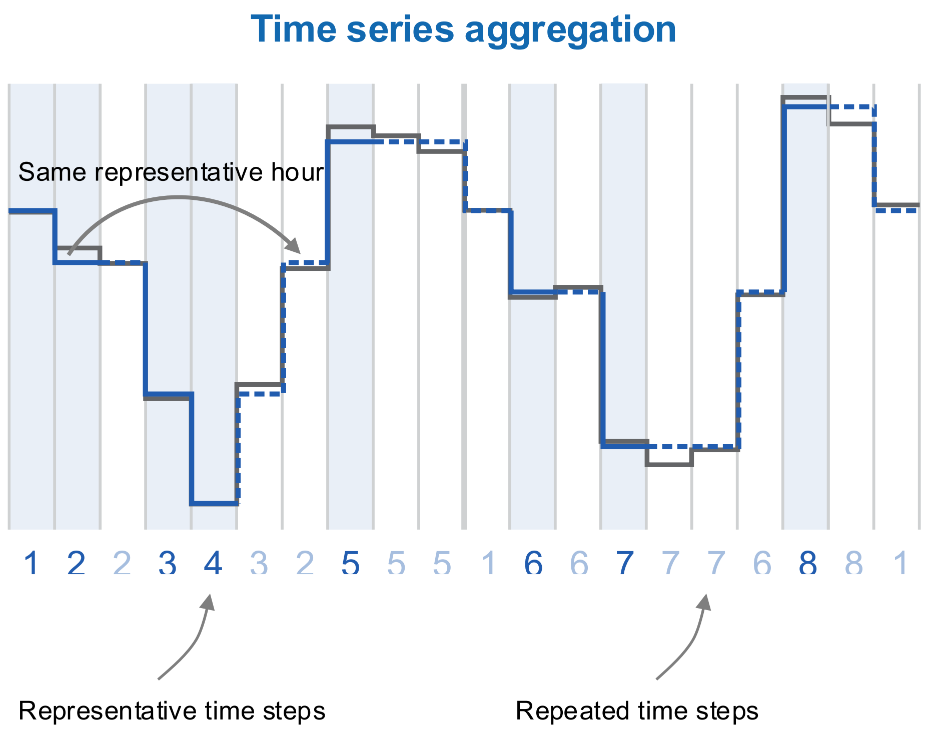
How to reduce the number of storage level variables and constraints without losing any information

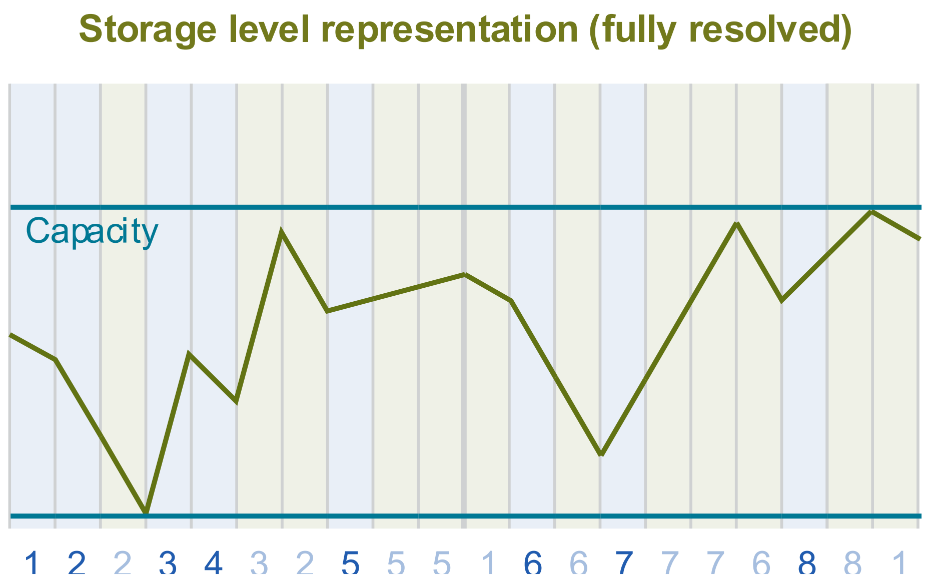
We cluster the input time series into 8 representative time steps



How to reduce the number of storage level variables and constraints without losing any information

We can reconstruct the storage level at every time step [1]
But can we reduce the number of storage level time steps? **Yes!**





[1] Gabrielli, P., Gazzani, M., Martelli, E. & Mazzotti, M. Optimal design of multi-energy systems with seasonal storage. *Applied Energy* (2018).

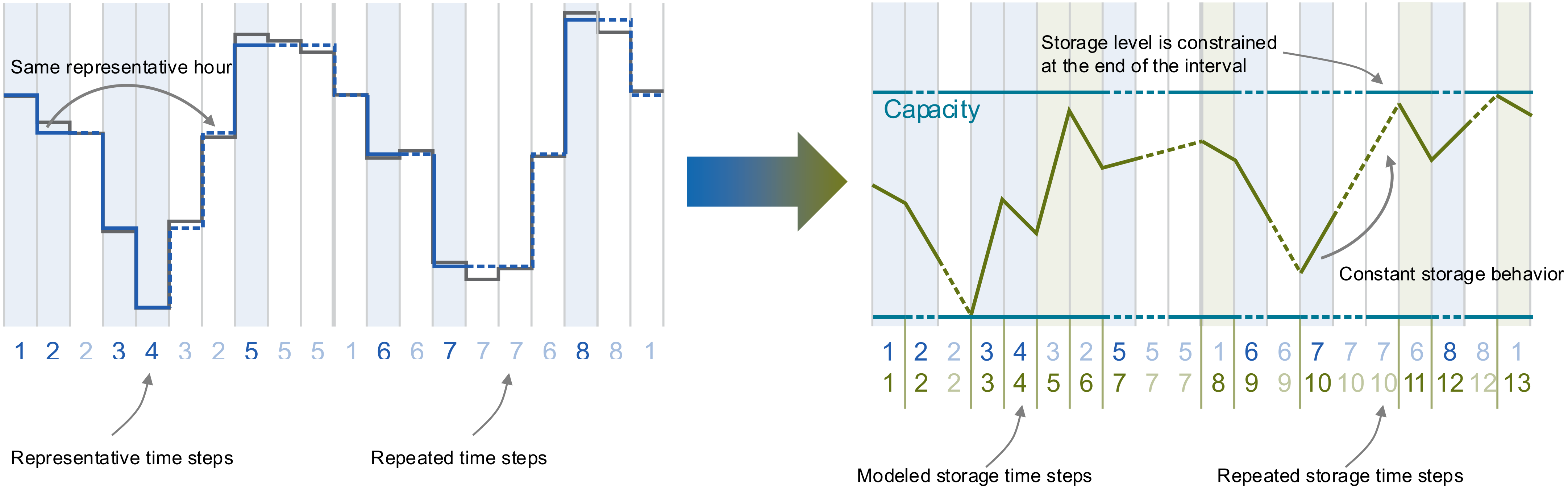
How to reduce the number of storage level variables and constraints without losing any information

Observation: The storage behavior does not change over the same representative hour
→ For every cluster of adjacent hours represented by the same RH: one storage time step
→ Every time we change the representative time step, we add one storage level time step
→ We can save many variables and constraints (from 20 to 13 storage level time steps)

We don't lose any information compared to full storage resolution!

Time series aggregation

Storage level representation (reduced)



Mathematical description of reduced RH storage representation with self-discharge

New indices:
Representative time steps: $i \in \mathrm{I}$
Storage level time steps: $j \in \mathrm{J}$

Constant net charging for RH $i = 5$:

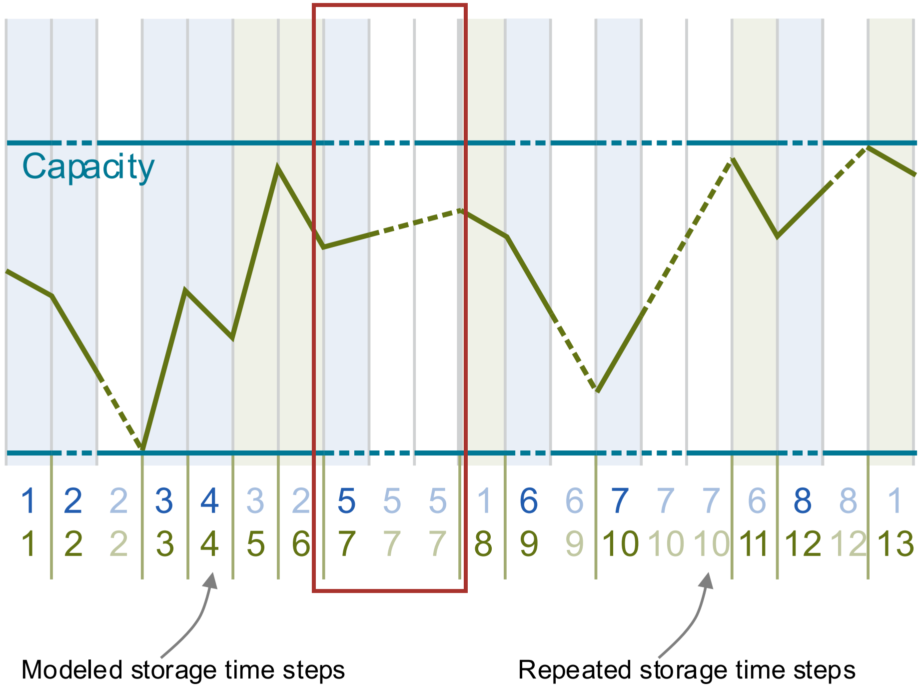
$$\Delta H_{s,t=8} = \Delta H_{s,t=9} = \Delta H_{s,t=10} = \Delta H_{s,i=5}$$

Storage level equations for $t = 8$, $t = 9$, and $t = 10$

$$\begin{aligned} L_{s,t=8} &= L_{s,t=7}(1-\varphi_s) + \Delta H_{s,i=5} \\ L_{s,t=9} &= L_{s,t=8}(1-\varphi_s) + \Delta H_{s,i=5} \\ L_{s,t=10} &= L_{s,t=9}(1-\varphi_s) + \Delta H_{s,i=5} \end{aligned}$$

We always add the same $\Delta H_{s,i=5}$!

Storage level representation (reduced)



$$L_{s,t=10} = L_{s,t=7}(1-\varphi_s)^3 + \Delta H_{s,i=5} \sum_{\tilde{j}=0}^{2} (1-\varphi_s)^{\tilde{j}}$$

$L_{s,j=7}$ $L_{s,j=6}$

Generic storage level constraint:

$$L_{s,j} = L_{s,j-1}(1-\varphi_s)^{d_j} + \Delta H_{s,\vartheta(j)} \sum_{\tilde{j}=0}^{d_j-1} (1-\varphi_s)^{\tilde{j}}$$

Duration of storage level time step

Unique mapping

Benchmarking on a greenfield optimization model of the European electricity and heating system

We benchmark the proposed method against four established methods:

1. *Superposition (RD)* [1]
2. *MinMax (RD)* [1]
3. *Full storage resolution (RH)* [2]
4. *Chrono (Chronological RH)* [3]

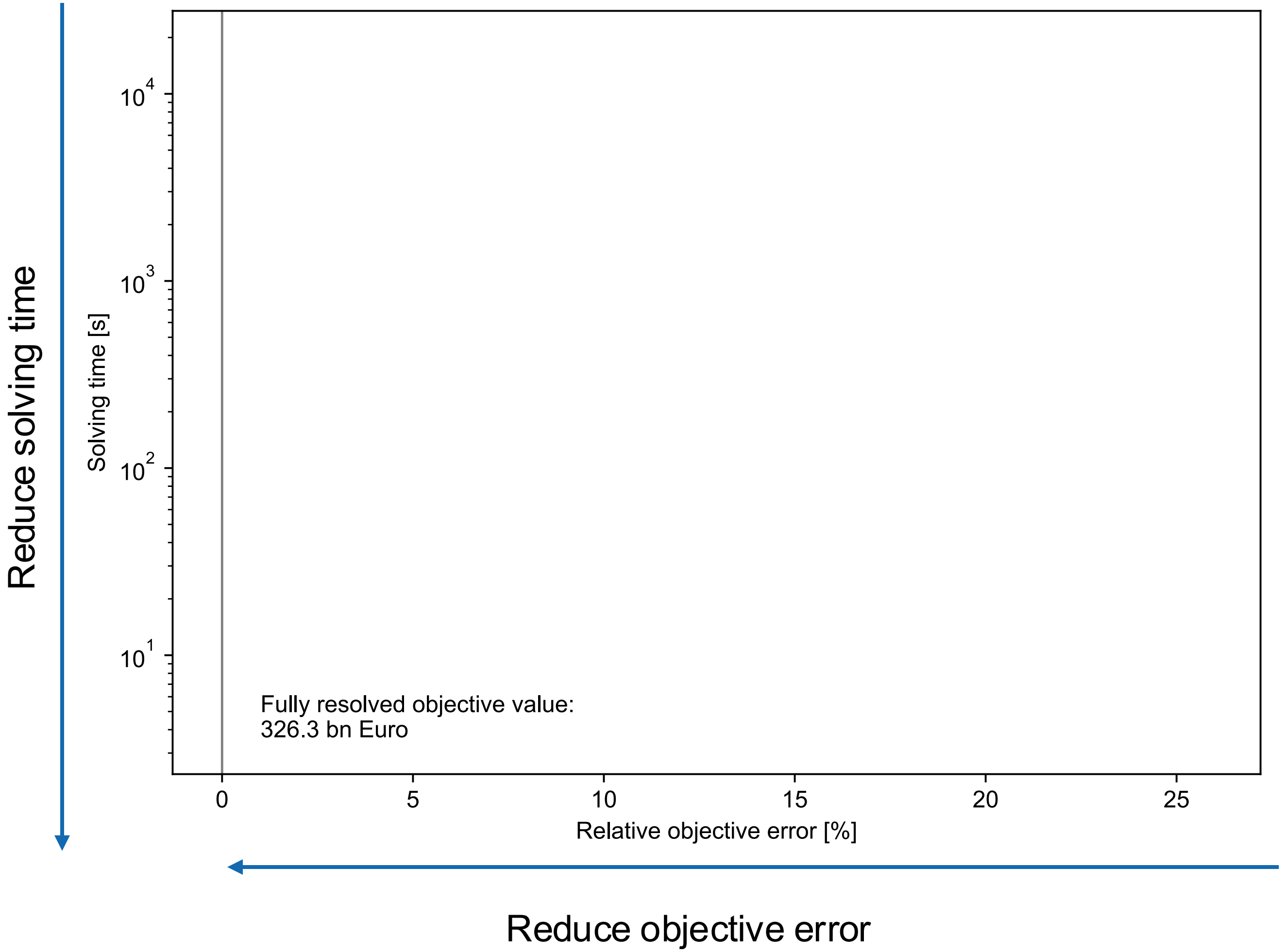
Only compare against methods that do not approximate the storage level

Optimize a **greenfield cost-minimization model** of the European (28 countries) electricity and heating system

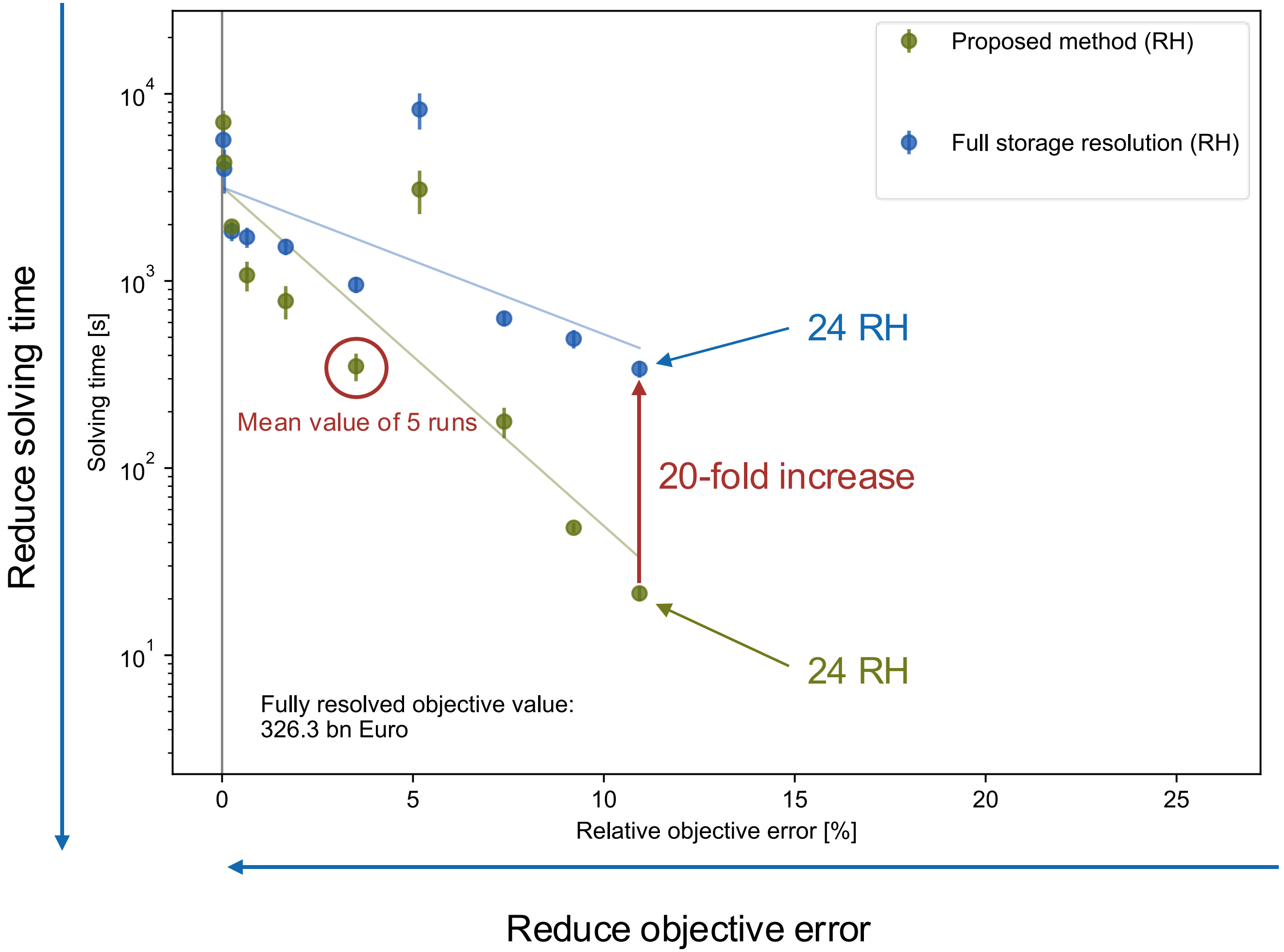
All storage representation methods are implemented in ZEN-garden [4] with hierarchical clustering (except for Chrono) and mean representation

[1] Kotzur, L., Markewitz, P., Robinius, M. & Stolten, D. Time series aggregation for energy system design: Modeling seasonal storage. *Applied Energy* (2018).
[2] Gabrielli, P., Gazzani, M., Martelli, E. & Mazzotti, M. Optimal design of multi-energy systems with seasonal storage. *Applied Energy* (2018).
[3] Pineda, S. & Morales, J. M. Chronological Time-Period Clustering for Optimal Capacity Expansion Planning With Storage. *IEEE Transactions on Power Systems* (2018).
[4] Mannhardt, J. *et al.* ZEN-garden: Optimizing energy transition pathways with user-oriented data handling. *SoftwareX* (2025).

RH methods better resolve the trade-off between accuracy and solving time



RH methods better resolve the trade-off between accuracy and solving time

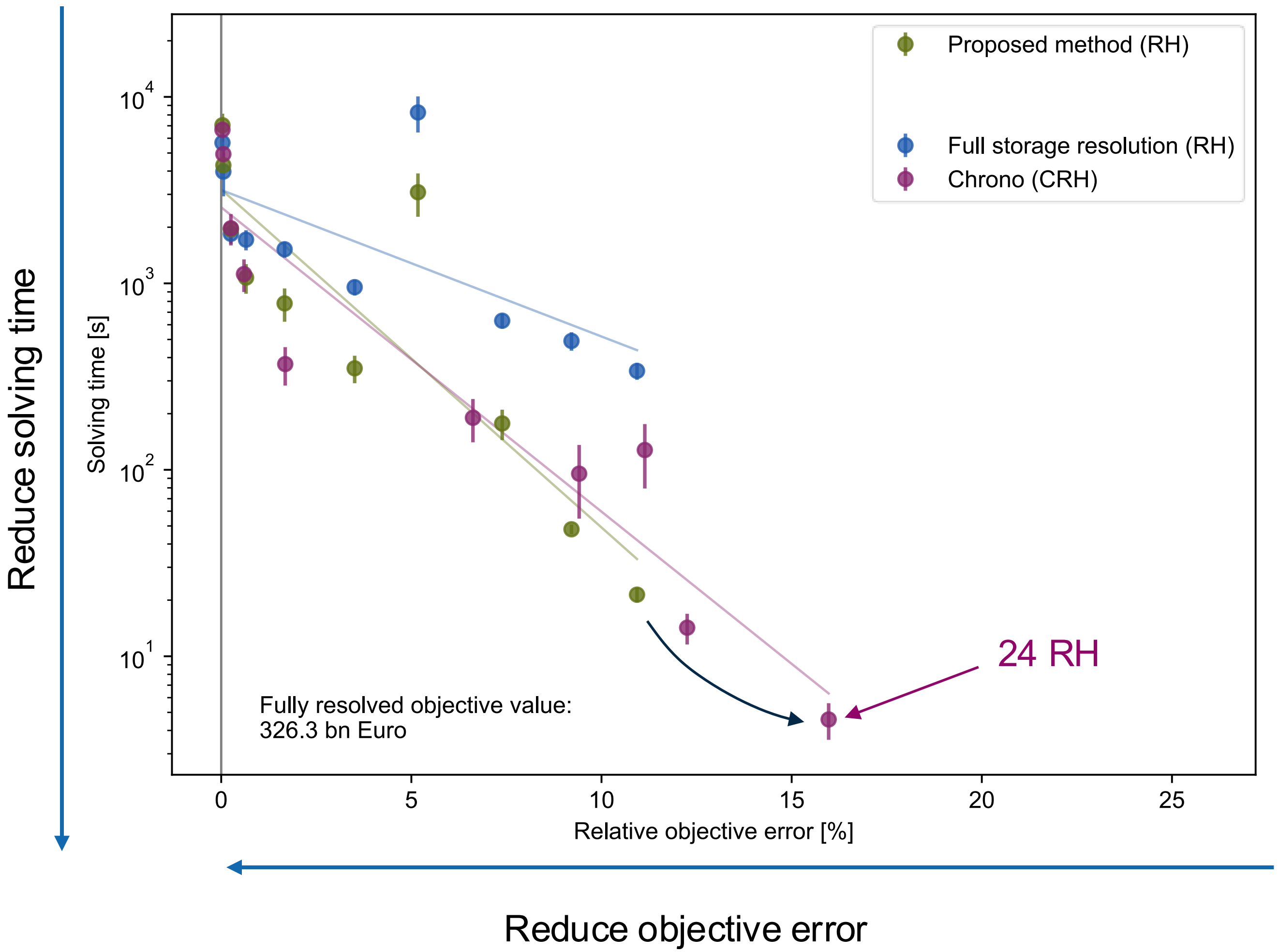


The proposed method shows lower solving times than *Full storage resolution* at the **same objective error**

Up to 20-fold solving time increase for *Full storage resolution.*

The strongest aggregation (24 RH) leads to an objective error of ~10%

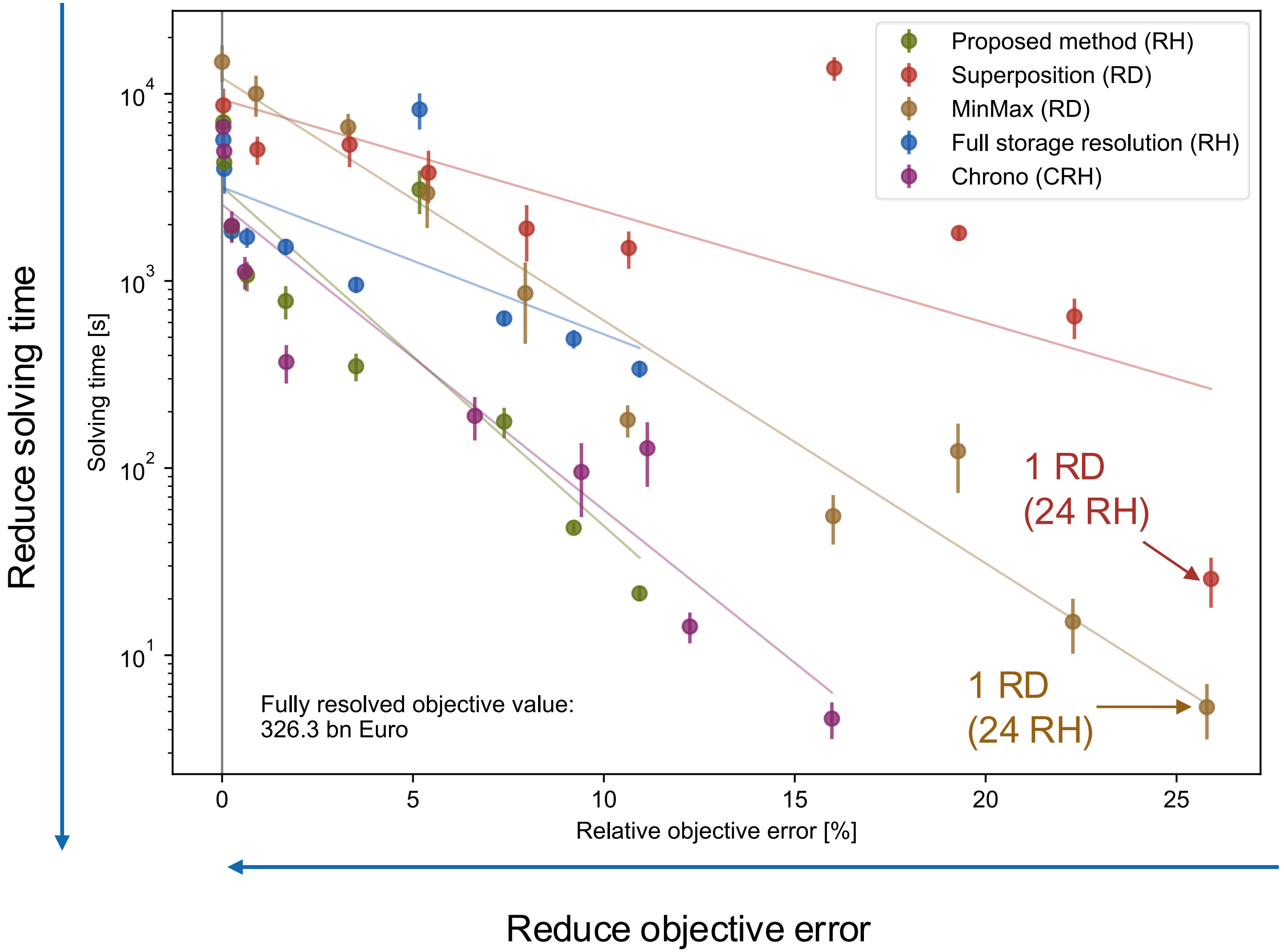
RH methods better resolve the trade-off between accuracy and solving time



Chrono shows similar accuracy and solving time as the proposed method

At the strongest aggregation, increased objective error (16%)

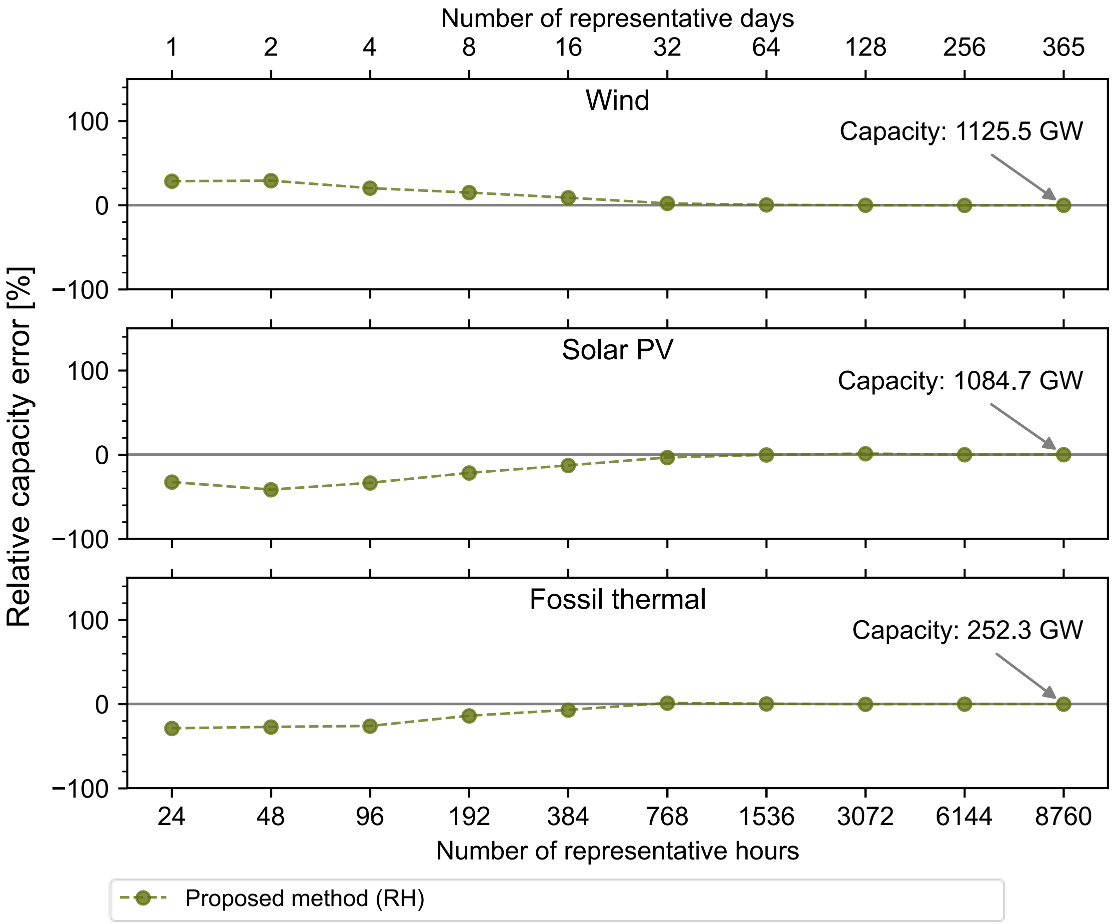
RH methods better resolve the trade-off between accuracy and solving time



The RD methods show the highest solving time and the highest objective error.

MinMax can reduce the solving time compared to *Superposition*

RH methods show a smooth reduction of the capacity error; Negligible error after 768 RH

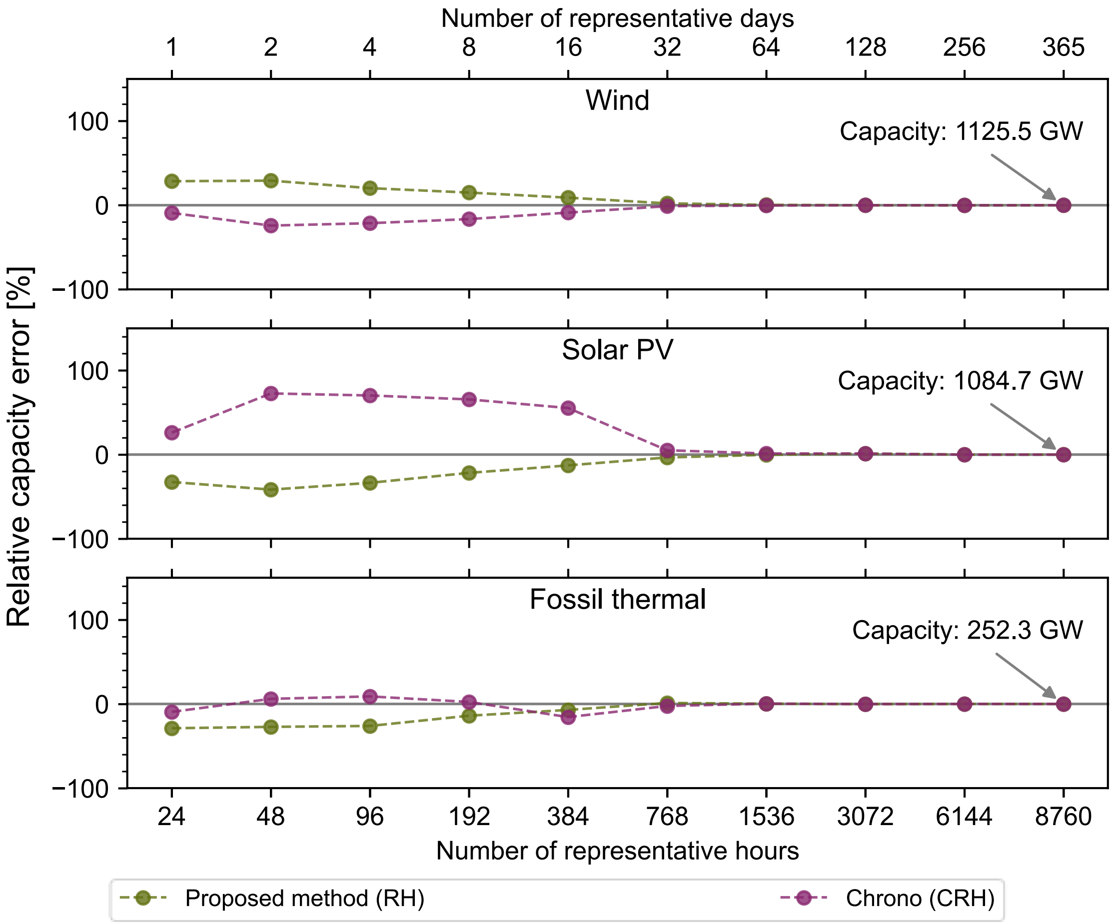


The proposed method **overestimates** onshore wind and **underestimates** solar PV

Maximum capacity error:
solar PV: - 40% (48 RH)

Lower fossil thermal capacity because time series are smoothed

RH methods show a smooth reduction of the capacity error; Negligible error after 768 RH

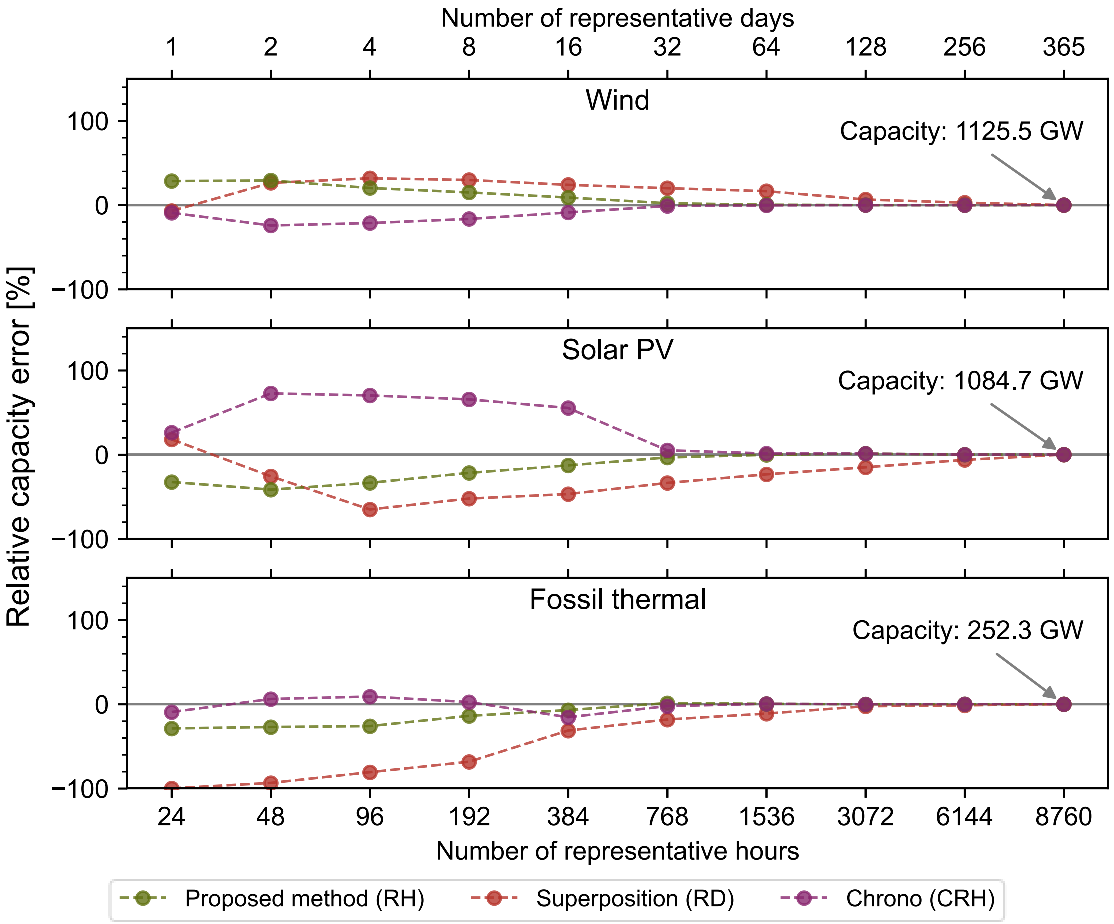


Chrono shows significantly higher solar PV capacity

Reason: daily profiles are removed entirely for strong aggregation
→ solar PV is more viable

At 768 RH: roughly 2 RH per day
→ daily behavior is captured
→ good approximation of solar PV

RH methods show a smooth reduction of the capacity error; Negligible error after 768 RH



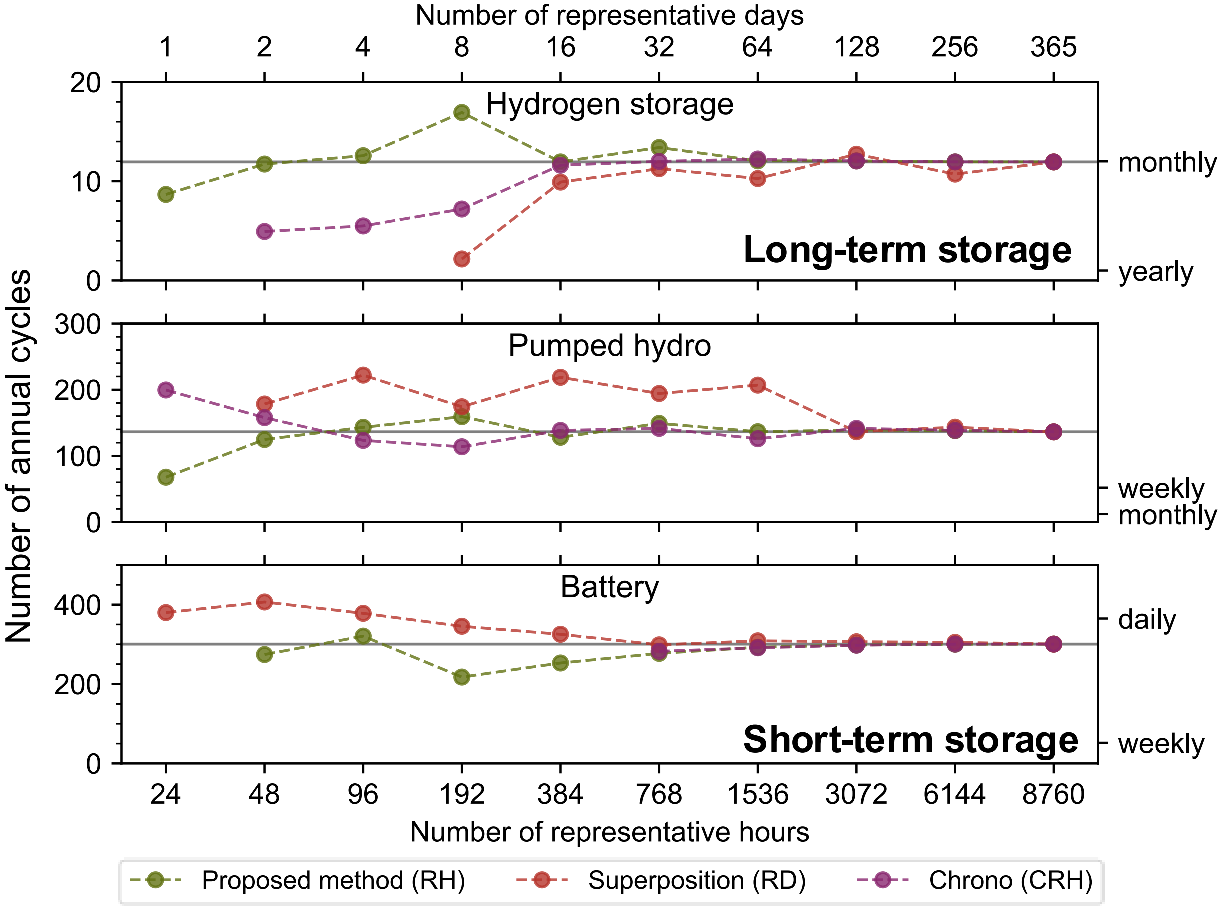
Superposition shows a persistent capacity error

Especially, in the required fossil thermal capacity

For solar PV, inverse effect of *Chrono*:
Only daily profile
→ lower viability of solar PV

ETH zürich

The proposed method approximates the storage behavior well



The proposed method shows a good approximation of the storage behavior, even at strong aggregation

Chrono shows lower hydrogen storage cycles and negligible battery capacity until 768 RH

Superposition **underestimates** the use of hydrogen storage and **overestimates** the use of batteries

Conclusions

Contribution: *RH-based storage level representation method with reduced storage level index*

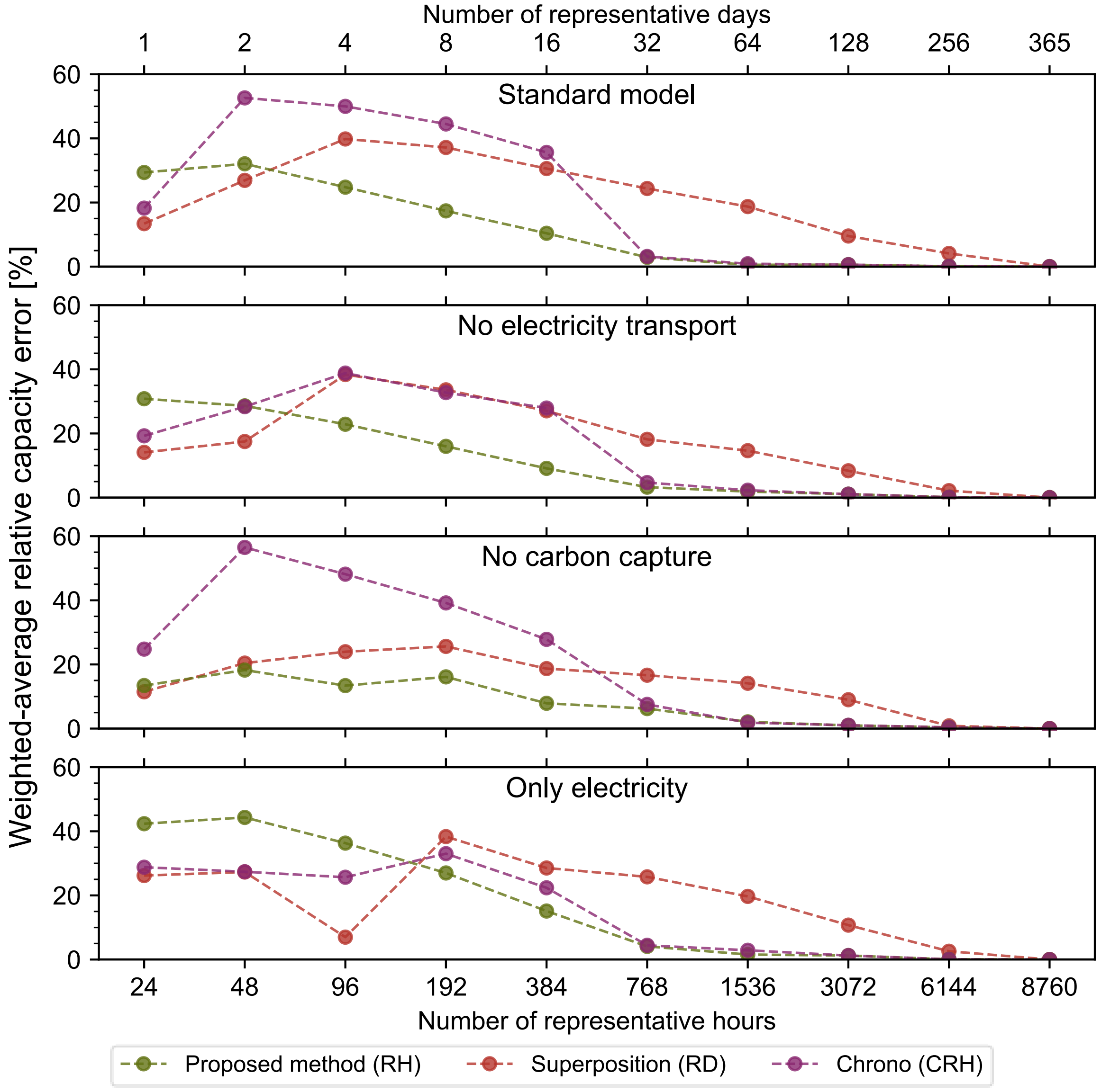
Clustering accuracy of using representative hours — Storage level representation to ensure computational tractability

Conclusions:

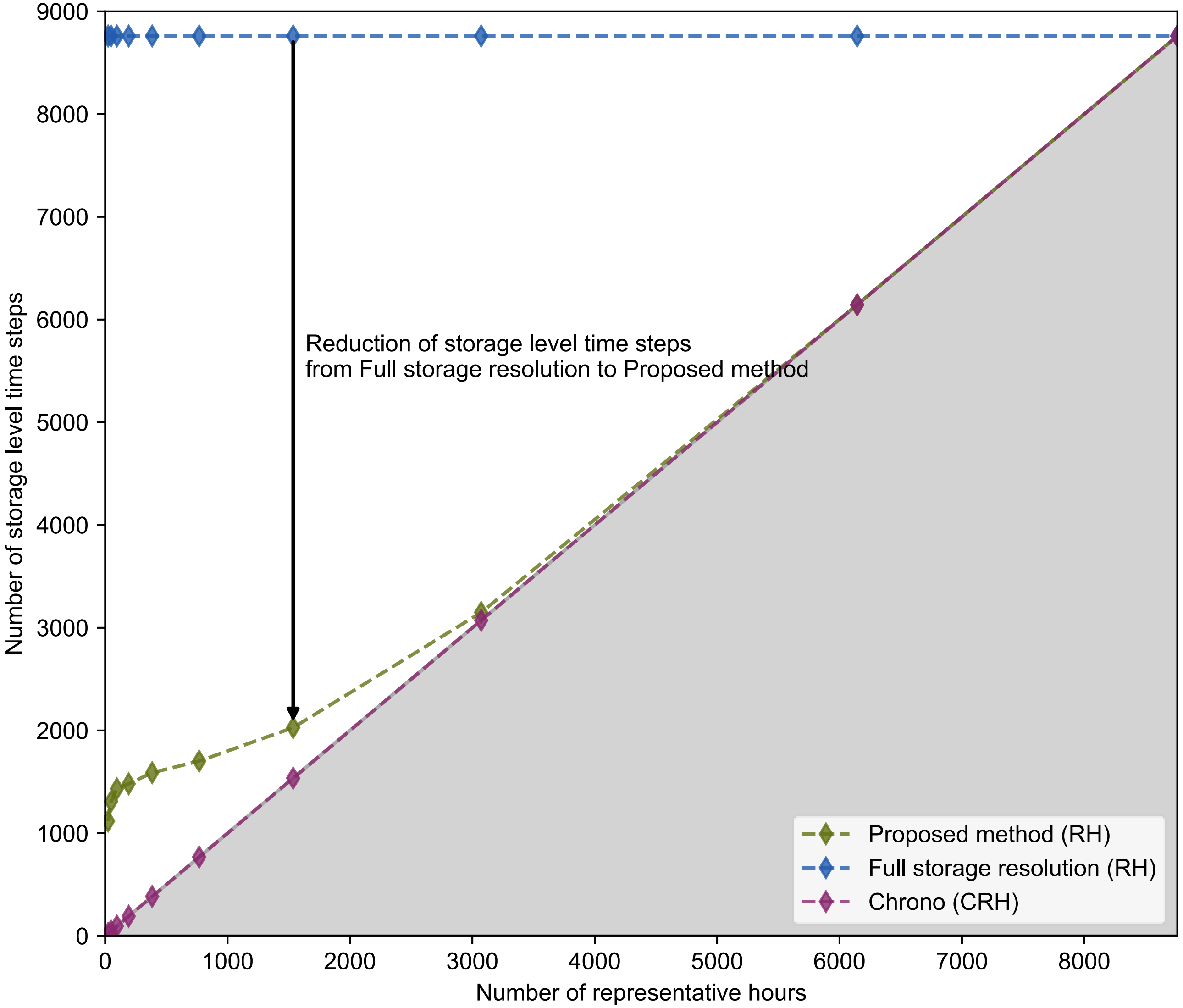
- RH represent the input data time series more accurately than RD with fewer representative time steps
- The proposed method shows a **strong reduction in solving time** while retaining a good accuracy
- The method can reproduce the operation of **short-term and long-term storage well**
- 100 to 500 representative hours are a good trade-off between accuracy and computational complexity for **large-scale multi-storage energy systems**

Appendix

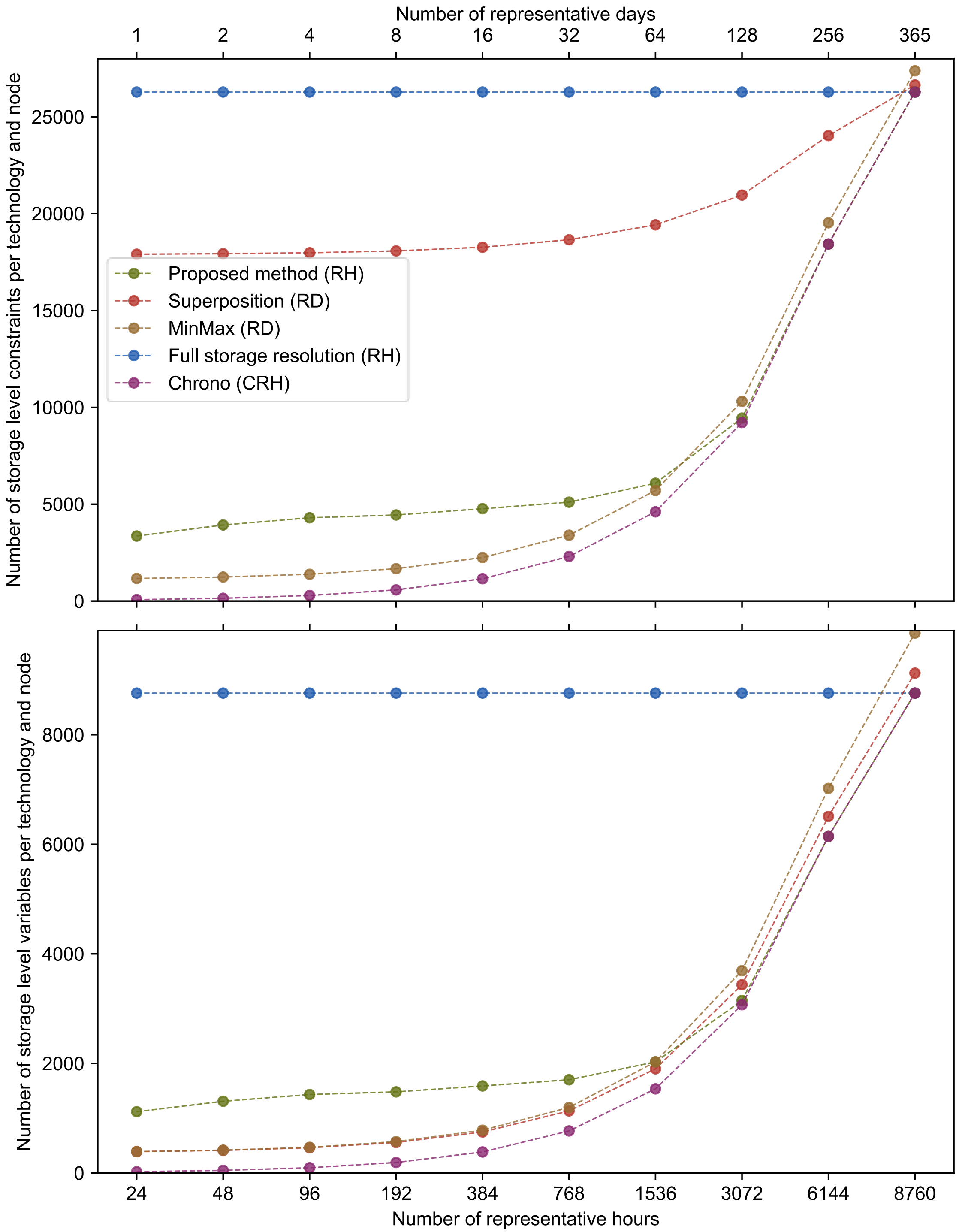
Robustness of methods for various large-scale energy system models



Number of time steps for RH methods

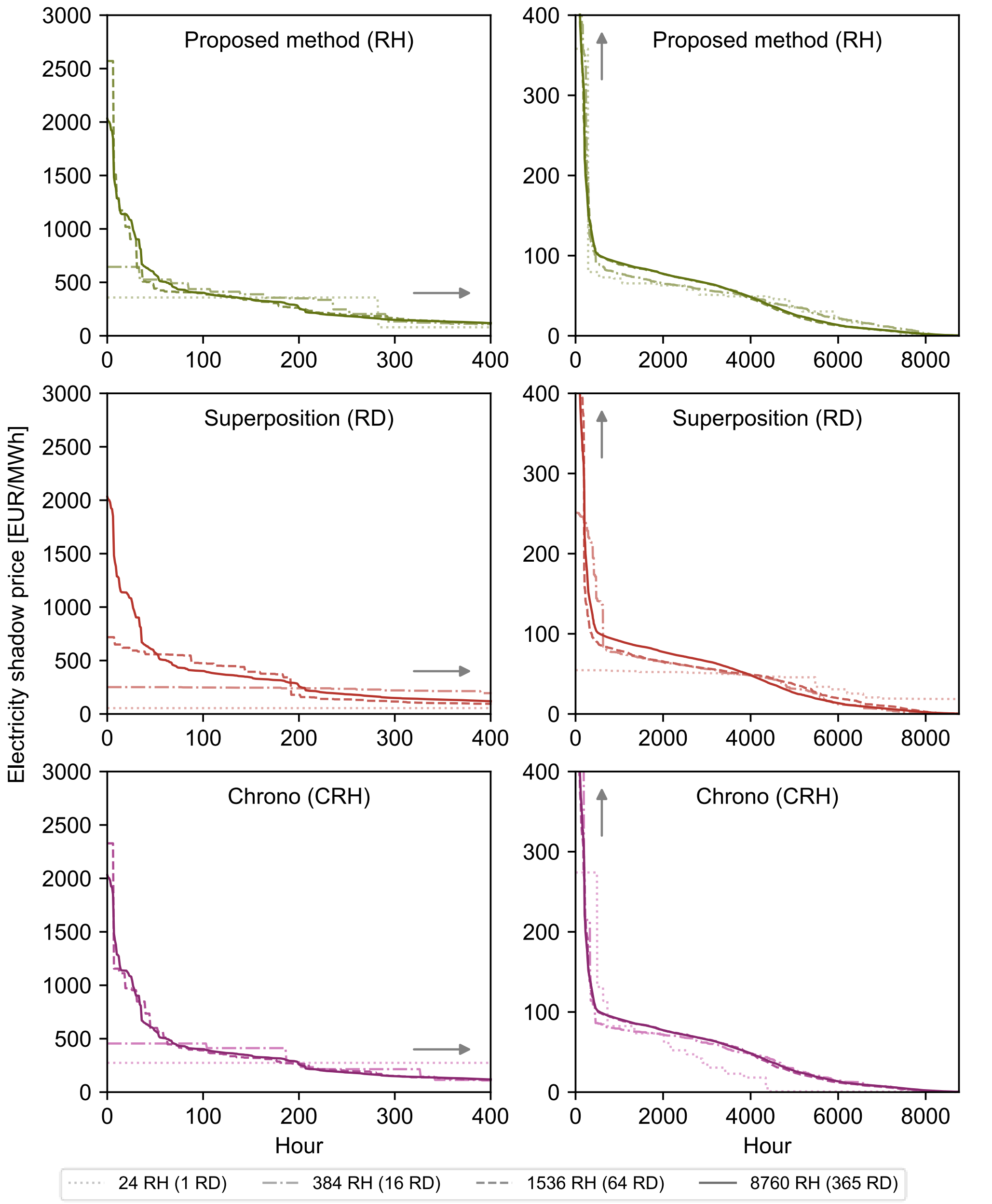


Number of constraints and variables

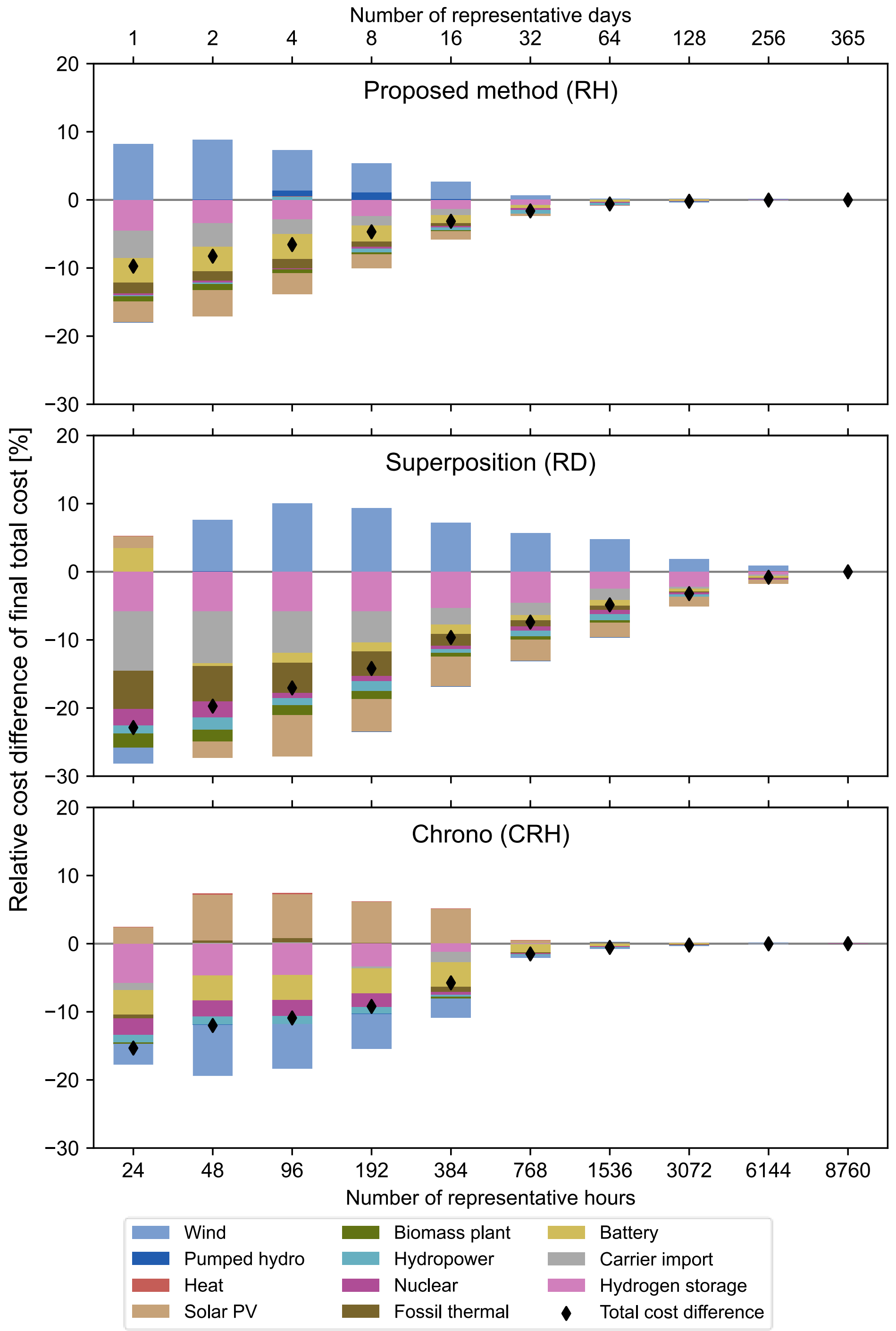


ETH zürich

Price duration curve of the shadow price of the electricity energy balance

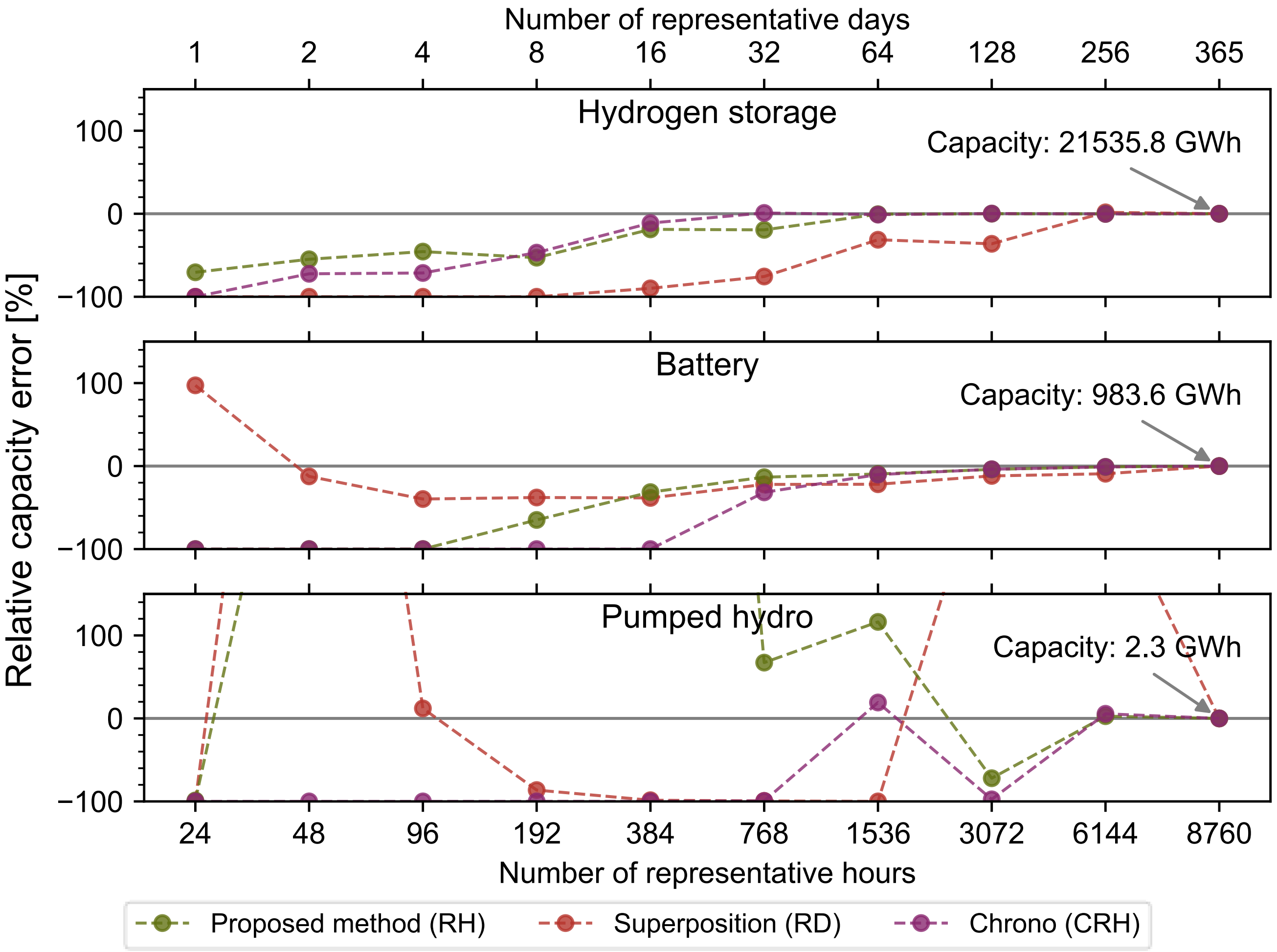


Cost error

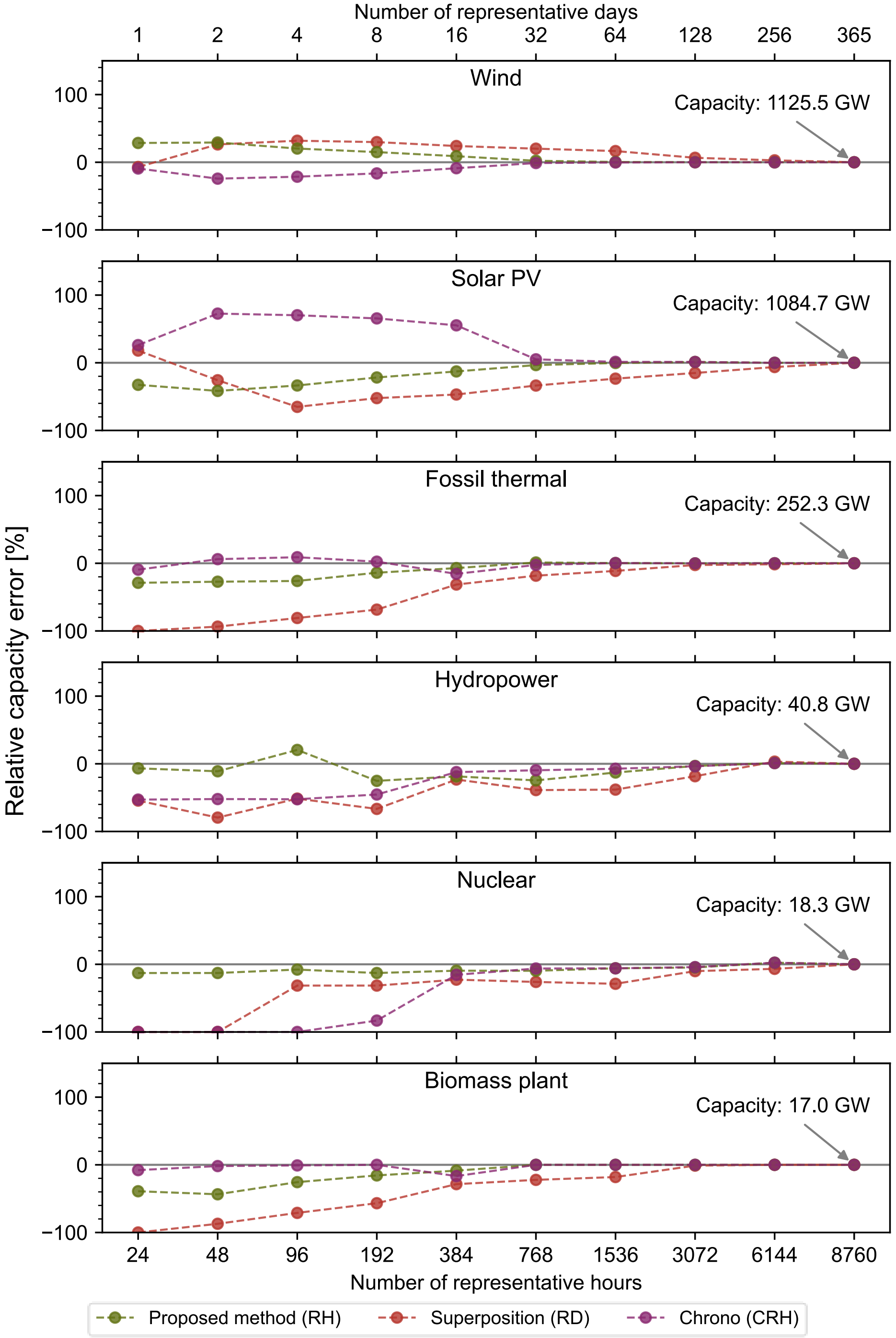


ETH zürich

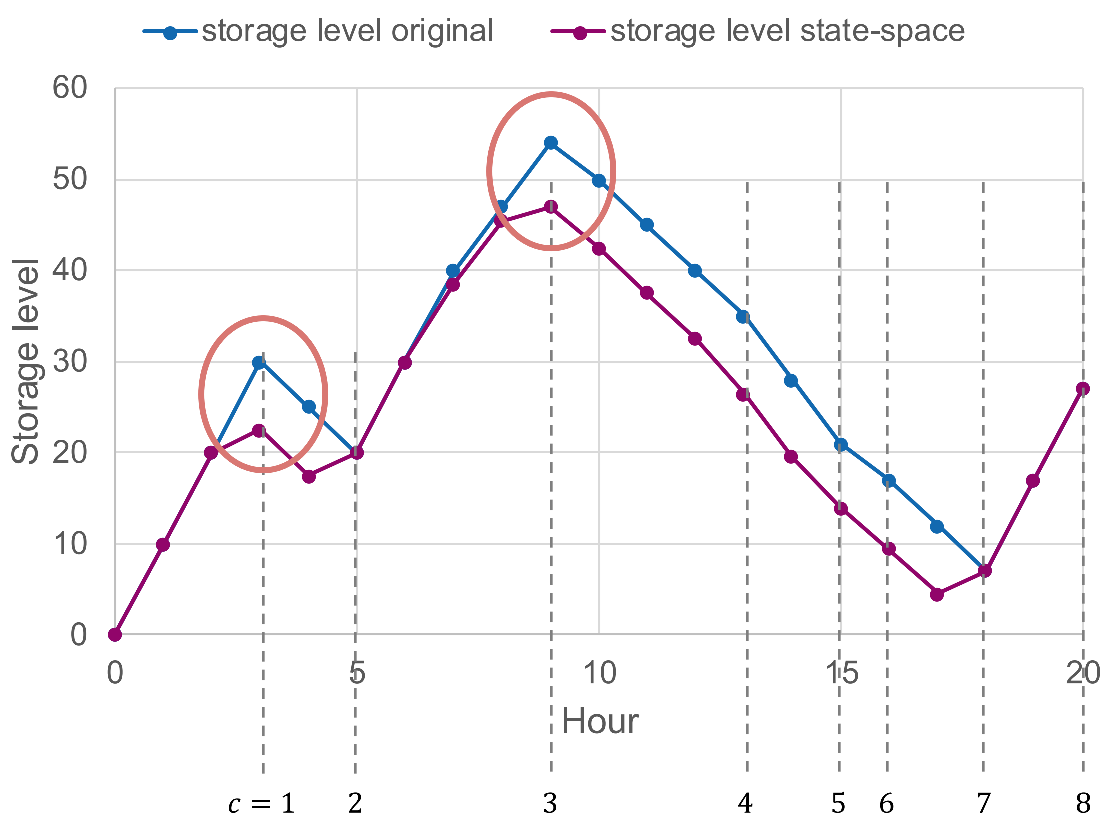
Capacity error storage



Capacity error conversion



State-space transition reconstruction



$$\Delta L_{s,k} = 0.5\left(\Delta H_{s,\kappa(k)} + \Delta H_{s,\kappa'(k)}\right), \text{ where } k = (i, i')$$

$$\kappa\colon \mathcal{K} \to \mathcal{I}, \kappa'\colon \mathcal{K} \to \mathcal{I}' \quad \text{(Transitions)}$$

$$L_s^0 + \sum_{k \in \mathcal{K}} \Delta L_{s,k} F_{k,c} \geq 0$$

$$L_s^0 + \sum_{k \in \mathcal{K}} \Delta L_{s,k} F_{k,c} \leq E_s$$

Storage level monotony

$$0 \leq L_{s,j} \leq E_s$$

Is $L_{s,j}$ monotonic over j? Yes!

$$L_{\hat{t}} = L_0(1-\varphi)^{\hat{t}} + \Delta H \sum_{\tilde{t}=0}^{\hat{t}-1} (1-\varphi)^{\tilde{t}}$$ for the intermediate time steps $\hat{t} \in [1, d_j]$

Two cases:

$\varphi = 0$

$$\frac{\mathrm{d}L_{\hat{t}}}{\mathrm{d}\hat{t}} = \Delta H$$

$0 < \varphi < 1$

$$\frac{\mathrm{d}L_{\hat{t}}}{\mathrm{d}\hat{t}} = \underbrace{\left(L_0 - \frac{\Delta H}{\varphi}\right) \ln(1-\varphi)}_{=\text{ constant } \forall \hat{t} \in [1, d_j]} (1-\varphi)^{\hat{t}}$$

Can't change its sign over $\hat{t} \in [1, d_j]$ → monotonic