Data aggregation for net-zero power systems (NetZero-Opt ERC Starting Grant)

Energy System Optimization Workshop @ TU Graz

Sonja Wogrin (PI), Luca Santosuosso, Thomas Klatzer, Benjamin Stöckl
Institute of Electricity Economics and Energy Innovation
Graz University of Technology

27.11.2025

Funded by the European Union
European Research Council
Established by the European Commission

Motivation. Dimensions of complexity of ESM

Meet the NetZero-Opt ERC Team

Uncertainty
E.g., scenario reduction.

Yannick

Check out Yannick's Poster

Technical
E.g., unit commitment, storage formulations.

Thomas

Spatial
E.g., node partitions and aggregations.

Benjamin

Temporal
Time series aggregation.

Sonja

Jakub

Luca

erc
European Research Council
Established by the European Commission

Agenda

I Time Series Aggregation. **Motivation** & **Starting Point** (Sonja)

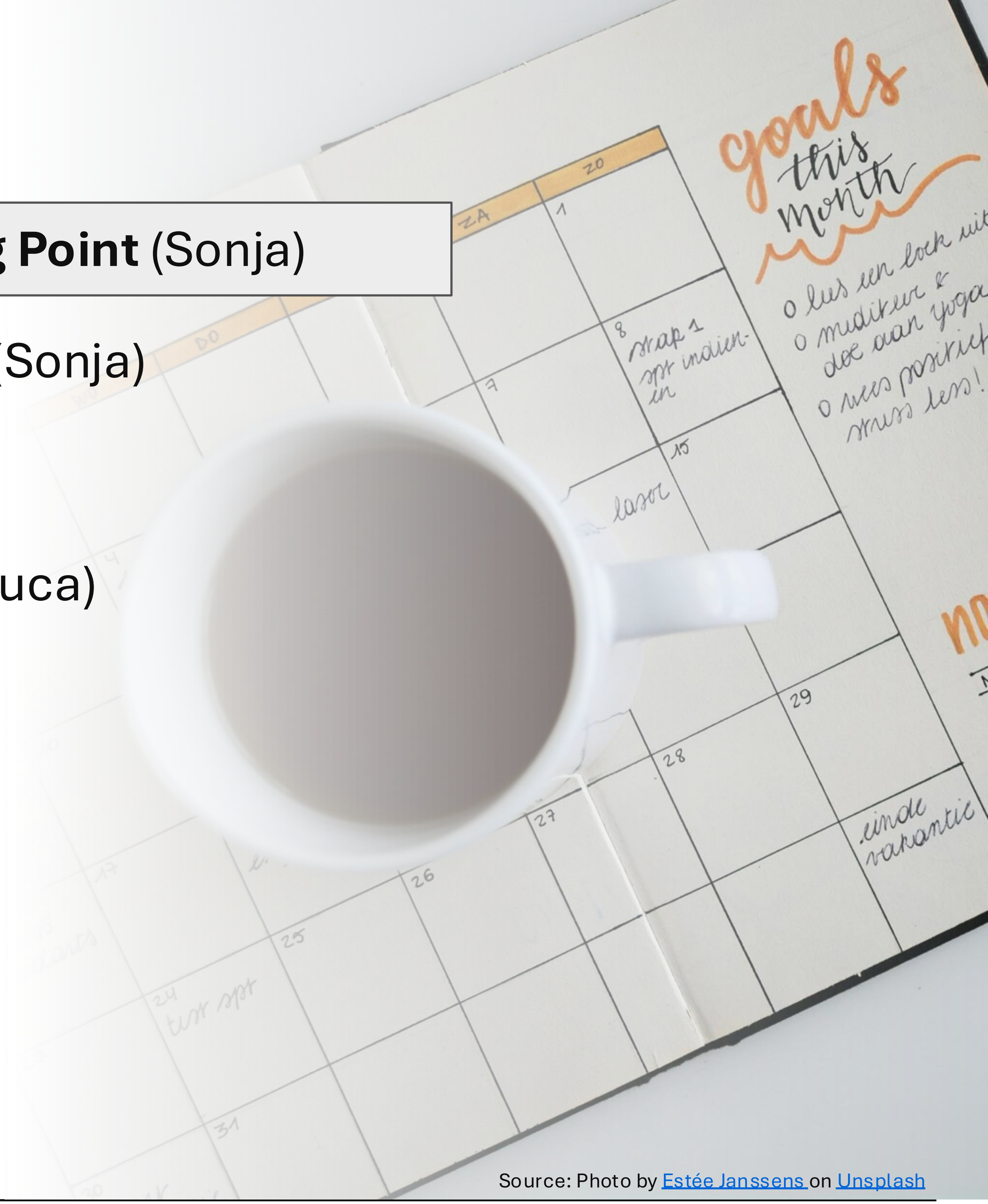
II Extension to **Network** & **Ramping** Constraints (Sonja)

III Extension to **Storage** Constraints (Thomas)

IV Time Series Aggregation with **Bounded Error** (Luca)

V Extension to **Grid Aggregation** (Benjamin)

VI Conclusions (Sonja)



Source: Photo by Estée Janssens on Unsplash

Computational Challenge

EU Power System Model

 # Variables: 630 M

NetZero-Opt

 # Variables: 0.5 M

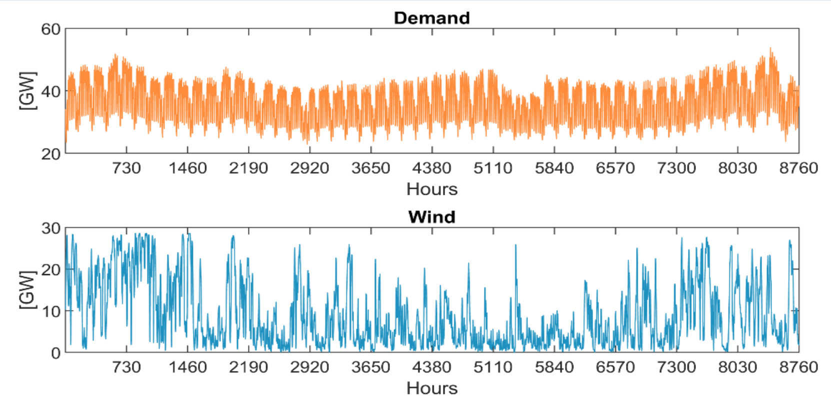
3 orders of magnitude

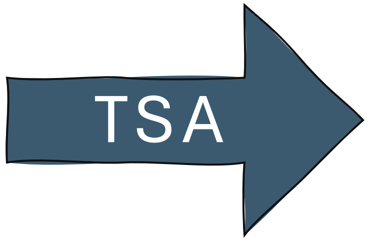
	Hourly Resolution	**NetZero-Opt**
World's Fastest Supercomputer* 600 M$	18 Days	**<1 ms**
Typical Workstation** 5000 €	Intractable	**12 Days**

Source*: Frontier (supercomputer) - Hewlett Packard Enterprise Frontier, or OLCF-5, 1102 exaFLOPS. Estimated cost: 600 M$.
Source**: Normal PC assumed to carry out 120 gigaFLOPS. Estimated cost: 5000€.

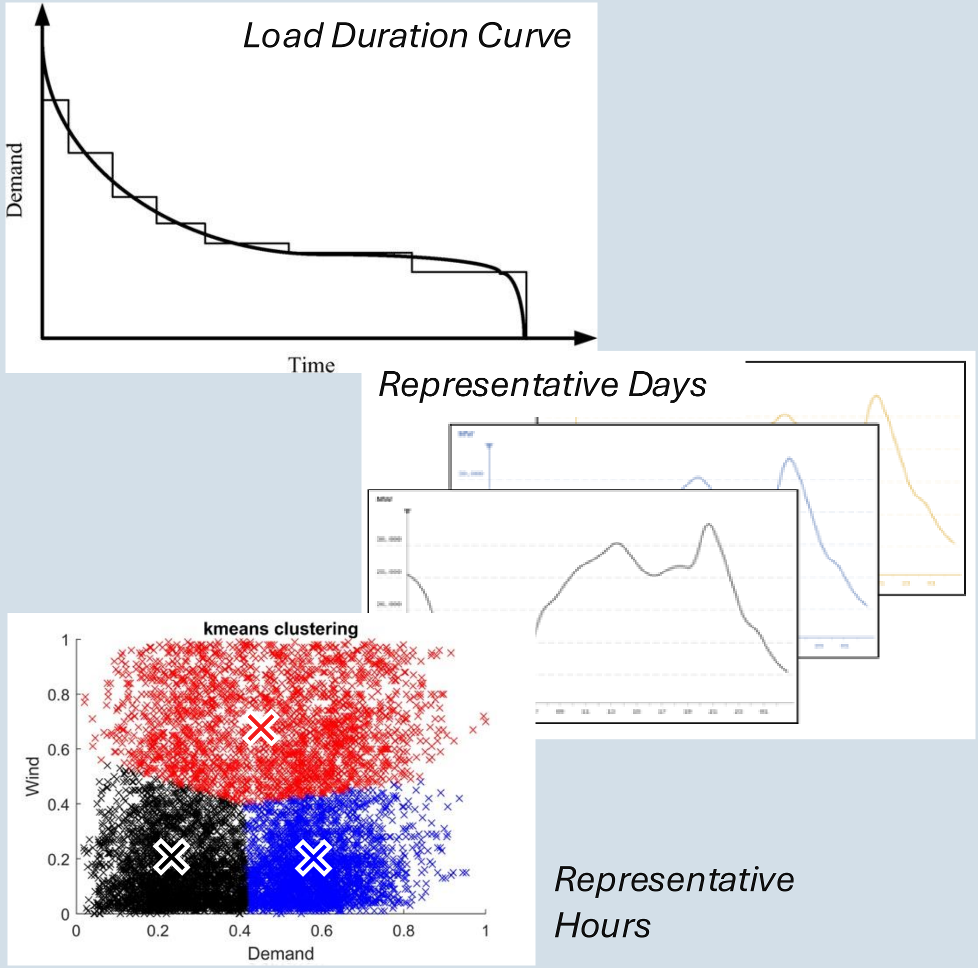
Time Series Aggregation (TSA)

Original Data Time Series





Times Series Aggregation



- Traditional TSA methods focus on the best **approximation of input data**.
- Usually, these methods do not provide **error bounds**!

Source:
Teichgraeber, H. and A.R. Brandt. "Time-series aggregation for the optimization of energy systems: Goals, challenges, approaches, and opportunities." Renewable and Sustainable Energy Reviews (2022)
Li, C. et al. "On representative day selection for capacity expansion planning of power systems under extreme operating conditions." International Journal of Electrical Power & Energy Systems (2022)
Hoffmann, M. et al. "A review on time series aggregation methods for energy system models." Energies (2020)
Hilbers, A.P. et al. "Importance subsampling: improving power system planning under climate-based uncertainty." Applied Energy (2019).

Full versus aggregated Economic Dispatch (ED)

Full Model

For generators *g* and time periods *k*:

- Minimize operating cost
- s.t.: Demand balance
- Lower and upper bounds

$$min \sum_{g,k} C_g p_{g,k} + \sum_k C^{nsp} nsp_k$$

$$\text{s.t.} \quad \sum_g p_{g,k} + nsp_k = D_k \quad \forall k$$

$$\underline{P}_g \leq p_{g,k} \leq \overline{P}_{g,k} \quad \forall g, k$$

Example: For 2 generators and 8760 time periods, this yields **26.280 variables**.

Aggregated Model

For generators *g* and **representative periods** *r (r<<k)* :

- Minimize **aggregated** operating cost, s.t.:
- s.t.: Demand balance
- Lower and upper bounds

$$min (\sum_{g,r} C_g p_{g,r} + \sum_r C^{nsp} nsp_r) W_r$$

$$\text{s.t.} \quad \sum_g p_{g,r} + nsp_r = D_r \quad \forall r$$

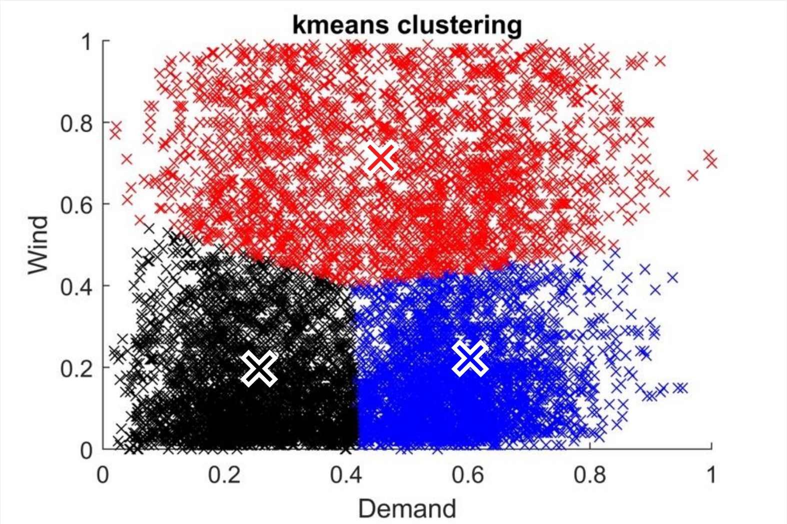
$$\underline{P}_g \leq p_{g,r} \leq \overline{P}_{g,r} \quad \forall g, r$$

Example: For 2 generators and 3 representative time periods, this yields **9 variables**.

Starting Point. Economic Dispatch (single node)

Traditional Framework

- Approximate 8760 hours of original time series using only **3 representative hours** with kmeans clustering:

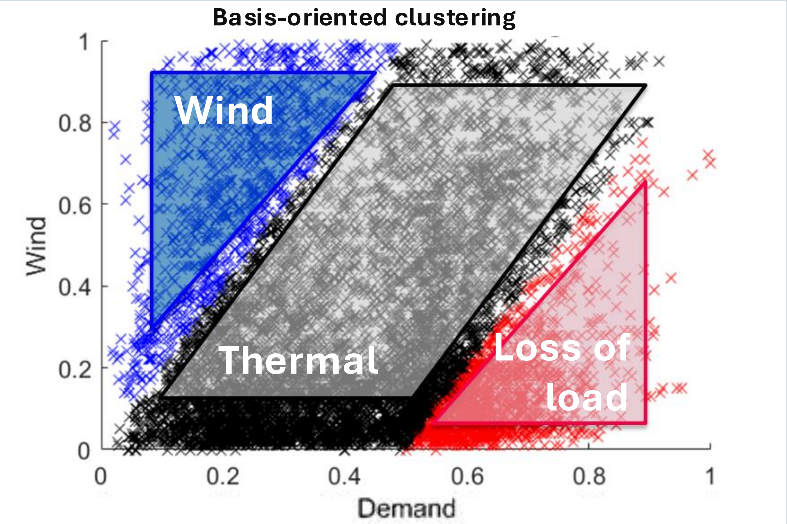


Results of aggregated economic dispatch problem:

- Relative **total system cost error 91%.**
- Relative **error** in estimated **loss of load 100%.**

NetZero-Opt

- Also uses 3 representative hours, but chosen within the same **simplex basis (i.e. active constraint sets)***:

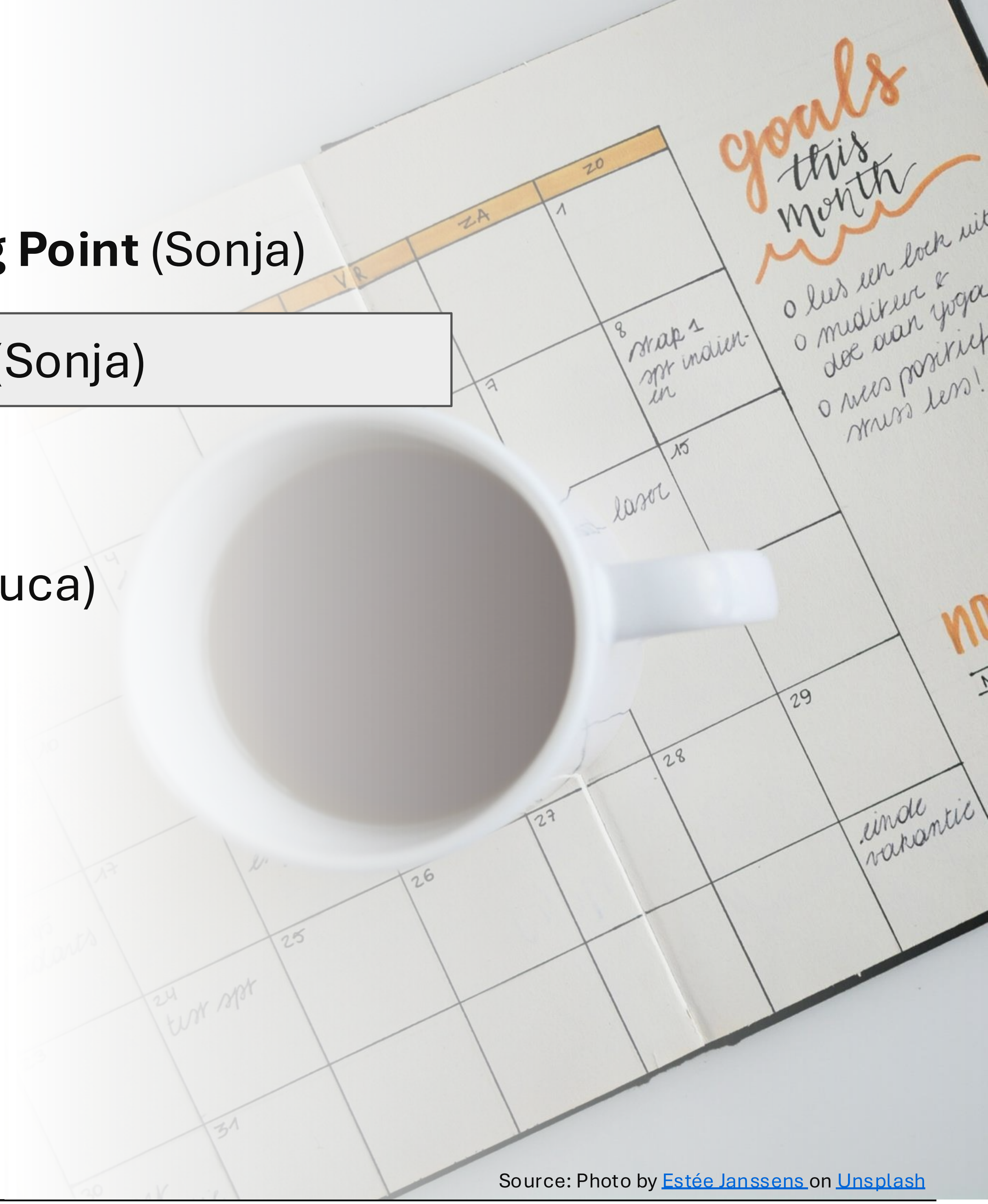


- Results yields **proven** relative **error of 0%!**
- Aggregation potential of **3 orders of magnitude** (measured in number of model variables).

Source: **Wogrin, S.** "Time series aggregation for optimization: One-size-fits-all?" IEEE Transactions on Smart Grid (2023).

Agenda

I	Time Series Aggregation. **Motivation** & **Starting Point** (Sonja)
II	Extension to **Network** & **Ramping** Constraints (Sonja)
III	Extension to **Storage** Constraints (Thomas)
IV	Time Series Aggregation with **Bounded Error** (Luca)
V	Extension to **Grid Aggregation** (Benjamin)
VI	Conclusions (Sonja)



Source: Photo by Estée Janssens on Unsplash

The methodology extends to transport problems

Formulation

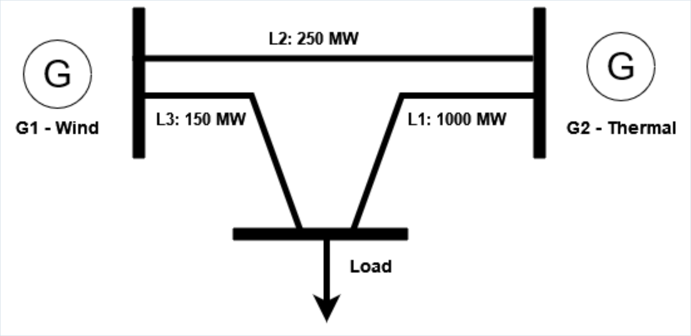
$$\min \sum_{k,g} C^g p_{g,k} + \sum_{k,i} C^{nsp} nsp_{i,k} + \sum_{k,i,j} C^N f_{k,j,i}$$

$$\text{s.t.} \quad \underline{P_w} \leq p_{w,k} \leq CF_k \overline{P_w} \quad \forall k, w$$

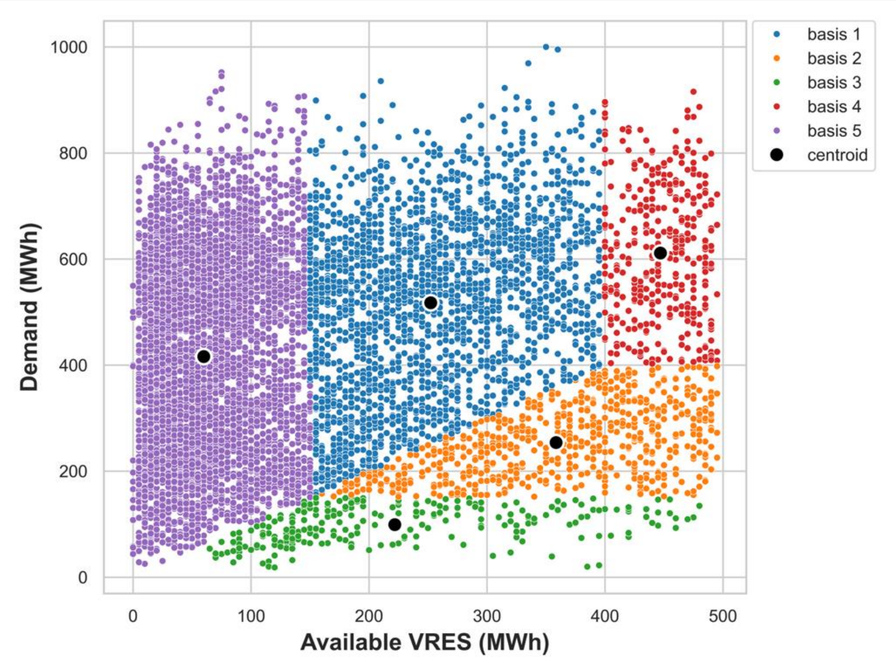
$$\underline{P_t} \leq p_{t,k} \leq \overline{P_t} \quad \forall k, t$$

$$\boxed{\begin{aligned} &\sum_j f_{k,j,i} - \sum_j f_{k,i,j} + nsp_{i,k} + \sum_{g \in i} p_{g,k} = D_{k,i} \forall k, i \\ &f_{k,i,j} \leq \overline{F_l} \quad \forall k, i, j \\ &f_{k,j,i} \leq \overline{F_l} \quad \forall k, i, j \end{aligned}}$$

Add network constraints



Aggregation Results



- More active constraint sets because of **line constraints**
- Only **five representative periods** required for exact aggregated model
- **Achieves 99.94% reduction** of number of variables

Source: **D. Cardona-Vasquez** et al. "Enhancing time series aggregation for power system optimization models: Incorporating network and ramping constraints." Electric Power Systems Research 230 (2024): 110267.

The methodology extends to time-linking constraints

Ramping Constraints

Formulation

$$
\min \sum_{k,g} C^g p_{g,k} + \sum_{k,i} C^{nsp} nsp_{i,k} + \sum_{k,i,j} C^N f_{k,j,i}
$$

$$
\text{s.t.} \quad \underline{P_w} \leq p_{w,k,i} \leq CF_k \overline{P_w} \quad \forall k, w, i
$$

$$
\underline{P_t} \leq p_{t,k,i} \leq \overline{P_t} \quad \forall k, t, i
$$

$$
\sum_j f_{k,j,i} - \sum_j f_{k,i,j} + nsp_{i,k} + \sum_{g \in i} p_{g,k} = D_{k,i} \forall k, i
$$

$$
f_{k,i,j} \leq \overline{F_l} \quad \forall k, i, j
$$

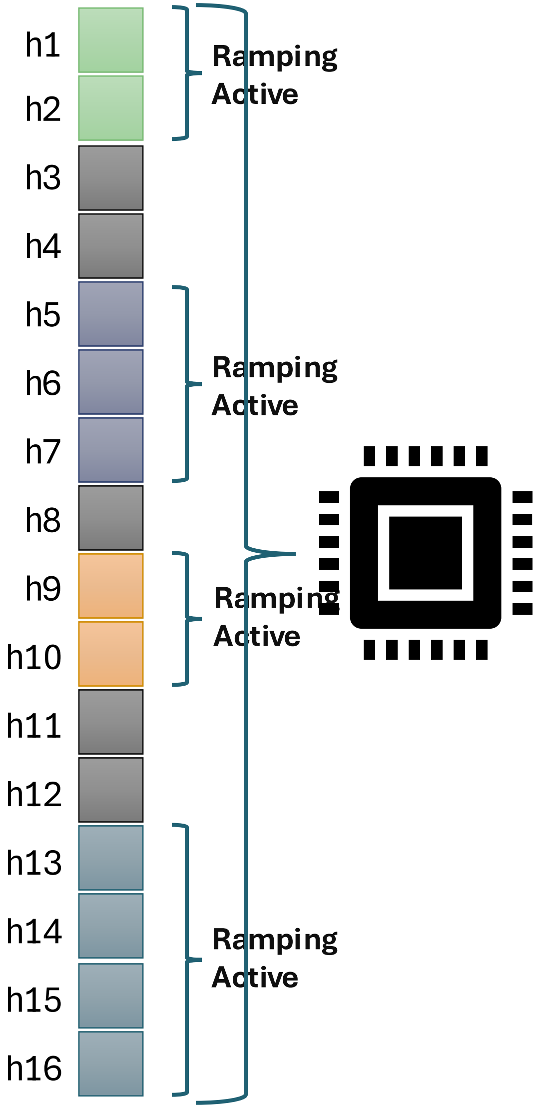
$$
f_{k,j,i} \leq \overline{F_l} \quad \forall k, i, j
$$

Model Structure

$$
\begin{array}{c}
\begin{array}{cccccc}
\multicolumn{2}{c}{\overbrace{\hphantom{1 \quad 1}}^{k=1}} & \multicolumn{2}{c}{\overbrace{\hphantom{0 \quad 0}}^{k=2}} & \multicolumn{2}{c}{\overbrace{\hphantom{0 \quad 0}}^{k=3}}
\end{array} \\
\left(\begin{array}{cc:cc:cc}
1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 \\
\hdashline
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
\hdashline
-1 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 1 & 0
\end{array}\right)
\end{array}
\begin{pmatrix} p_{T,1} \\ p_{W,1} \\ p_{T,2} \\ p_{W,2} \\ p_{T,3} \\ p_{W,3} \end{pmatrix}
\leq
\begin{pmatrix} D_1 \\ D_2 \\ D_3 \\ \overline{P_T} \\ \overline{P_{W,1}} \\ \overline{P_T} \\ \overline{P_{W,2}} \\ \overline{P_T} \\ \overline{P_{W,3}} \\ RU_T \\ RU_T \end{pmatrix}
$$

Source: **D. Cardona-Vasquez** et al. "Disaggregation of energy system optimization models using machine learning for identification of active constraints." Sustainable Energy, Grids and Networks 43 (2025): 101772.

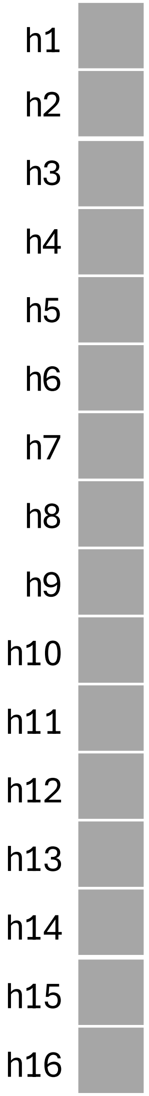
Identifying "active constraint sets" which are time-linked



1. Identify **active ramping constraints**

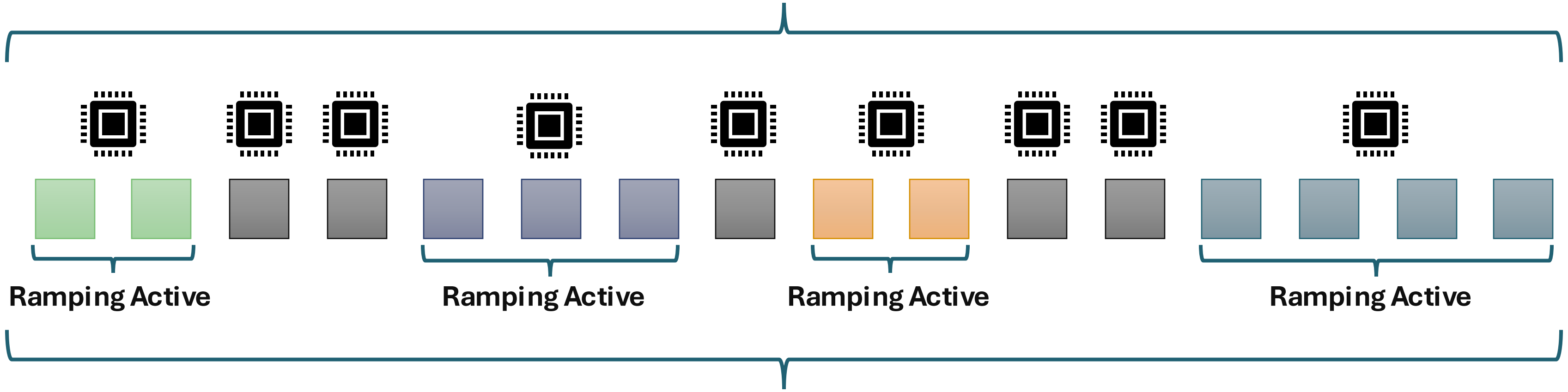
Source: **D. Cardona-Vasquez** et al. "Enhancing time series aggregation for power system optimization models: Incorporating network and ramping constraints." Electric Power Systems Research 230 (2024): 110267.

Identifying "Bases" with time linking



1. **Identify** active ramping constraints
2. **Disaggregate** into independent submodels
3. **Solve** submodels **in parallel**
4. **Aggregate** the submodel results

Disaggregate full model into submodels and solve in parallel

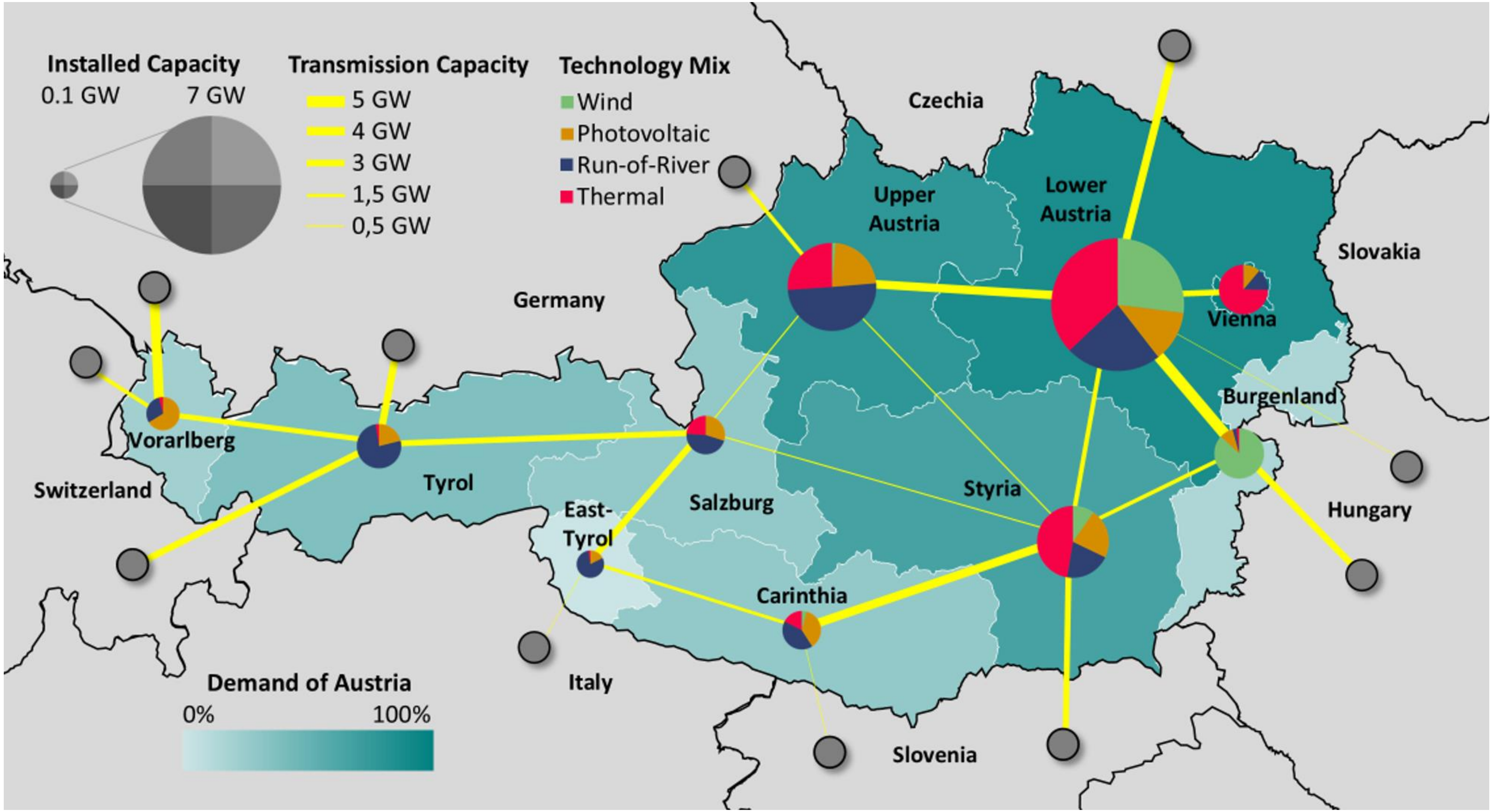


Aggregate results and compare with full hourly run

Source: **D. Cardona-Vasquez** et al. "Disaggregation of energy system optimization models using machine learning for identification of active constraints." Sustainable Energy, Grids and Networks 43 (2025): 101772.

www.tugraz.at

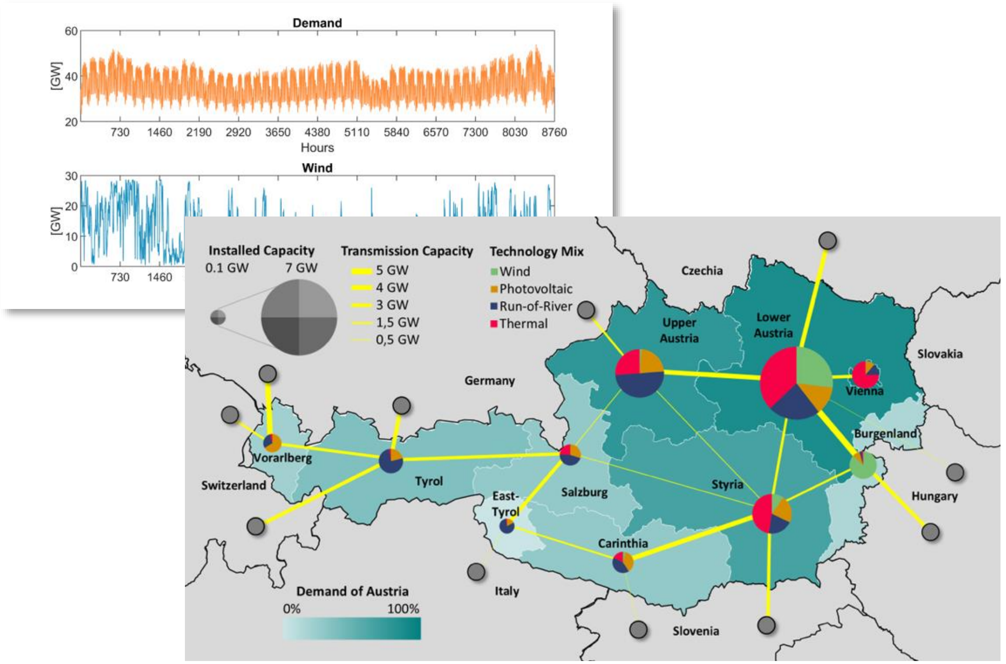
Case Study: Austrian Power System

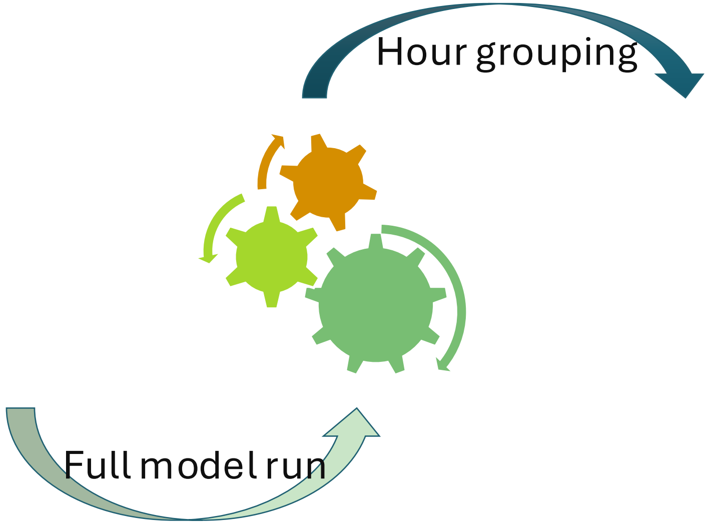


Source: **D. Cardona-Vasquez** et al. "Disaggregation of energy system optimization models using machine learning for identification of active constraints." Sustainable Energy, Grids and Networks 43 (2025): 101772.

Forecasting active ramping constraints

Training a classifier to predict time-linked periods





H1	
H2	
H3	
H4	
H5	
H6	
H7	
H8	
H9	
H10	
H11	
H12	
H13	
H14	
H15	
H16	
H17	
H18	
H19	
H20	
H21	
H22	
H23	
H24	
H25	
H26	
⋮	⋮
H8733	
H8734	
H8735	
H8736	

At least one ramping constraint is active

Add previous and next hour

- **Goal**: A classifier to identify the colored column from input data only
- Each group (i.e., “basis”) can be run independently of the others
- Non-colored (**gray**) hours are **temporally unlinked**
- Result: **accuracy of classifier 92%**

Source: **D. Cardona-Vasquez** et al. "Disaggregation of energy system optimization models using machine learning for identification of active constraints." Sustainable Energy, Grids and Networks 43 (2025): 101772.

Results: Optimization Model Outputs

Takeaways Model Results

- Relative error in **total system cost** is 0.004%.
- Relative error in **total production** per technology <0.01%.
- One outlier of 17.78% (small in absolute terms).
- **Robust solution quality** across 6 other test years.

Relative error of energy produced per technology and federal state

Federal State	Run-of-River	Solar-PV	Thermal	Wind	Imports
Burgenland	0.19%	0.00%	0.00%	0.00%	0.00%
Carinthia	0.00%	0.00%	0.65%	0.00%	0.00%
Lower Austria	0.00%	0.01%	0.56%	0.00%	0.00%
Upper Austria	0.00%	0.00%	0.63%	0.00%	0.00%
East Tyrol	0.00%	0.00%	0.00%	0.00%	0.00%
Styria	0.01%	0.02%	0.14%	0.00%	0.00%
Salzburg	0.00%	0.00%	0.00%	0.00%	0.00%
Tyrol	0.00%	0.00%	0.00%	0.00%	17.78%
Vorarlberg	0.00%	0.00%	0.00%	0.00%	0.00%
Vienna	0.00%	0.00%	0.24%	0.00%	0.00%

Colour scale: 0.0%, 0.1%, 0.2%, 0.3%, 0.4%, 0.5%, 0.6%

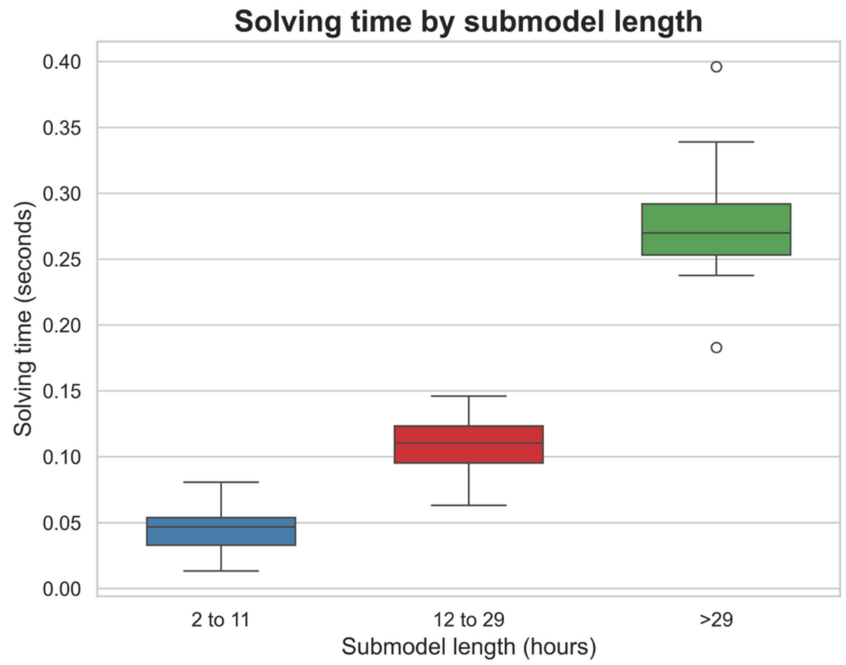
Source: **D. Cardona-Vasquez** et al. "Disaggregation of energy system optimization models using machine learning for identification of active constraints." Sustainable Energy, Grids and Networks 43 (2025): 101772.

Results: Computational Improvement

Takeaways Computational Results

- Full solution solving time: **115 s**
- Longest solving time (for a single submodel): **0.4 s** – (lower bound for a parallelized run)
- **300 times** faster
- Disaggregation based on active constraint sets seems promising

Length (hours)	Submodels (#)
2 to 11	266
12 to 29	123
>29	11
Total	**400**



Source: **D. Cardona-Vasquez** et al. "Disaggregation of energy system optimization models using machine learning for identification of active constraints." Sustainable Energy, Grids and Networks 43 (2025): 101772.

Key messages

Network & Ramping Constraints

Exact TSA with network & ramping

- **Exact TSA** based on active constraint sets extends to network constraints.
- Exact TSA is possible with **time-linking** constraints (i.e., ramping) when model structure is accounted for.
- In practice, the combination of **disaggregation** of the full-scale model, **parallelization** & **aggregation** of submodels leads to high model accuracy and considerable computational improvements.

🚀 Check out our papers

Disaggregation of energy system optimization models using machine learning for identification of active constraints

David Cardona-Vasquez [a, b,*], Alexander Konrad [a], Yannick Werner [a], Sonja Wogrin [a, b]

[a] Institute of Electricity Economics and Energy Innovation (IEE), Graz University of Technology, Inffeldgasse 18, Graz, Austria
[b] Research Center ENERGETIC, Graz University of Technology, Rechbauerstraße 12, Graz, Austria

HIGHLIGHTS

- Machine learning assisted constraint identification achieves a speedup of up to 300 times in solving time.
- The approach allows for the parallelization of time-linked energy system models.
- Disaggregation error is less than 1 % with respect to the complete model.
- For training, active constraint identification only requires a previous run of the complete model.

Enhancing time series aggregation for power system optimization models: Incorporating network and ramping constraints

David Cardona-Vasquez [a,*], Thomas Klatzer [a,b], Bettina Klinz [c,b], Sonja Wogrin [a,b]

[a] Institute of Electricity Economics and Energy Innovation (IEE), Graz University of Technology, Inffeldgasse 18, Graz, Austria
[b] Research Center ENERGETIC, Graz University of Technology, Rechbauerstraße 12, Graz, Austria
[c] Institute for Discrete Mathematics, Graz University of Technology, Steyrergasse 30, Graz, Austria

ARTICLE INFO

Keywords:
Power systems optimization
Mathematical modeling
Dimensionality reduction
Renewable energy sources

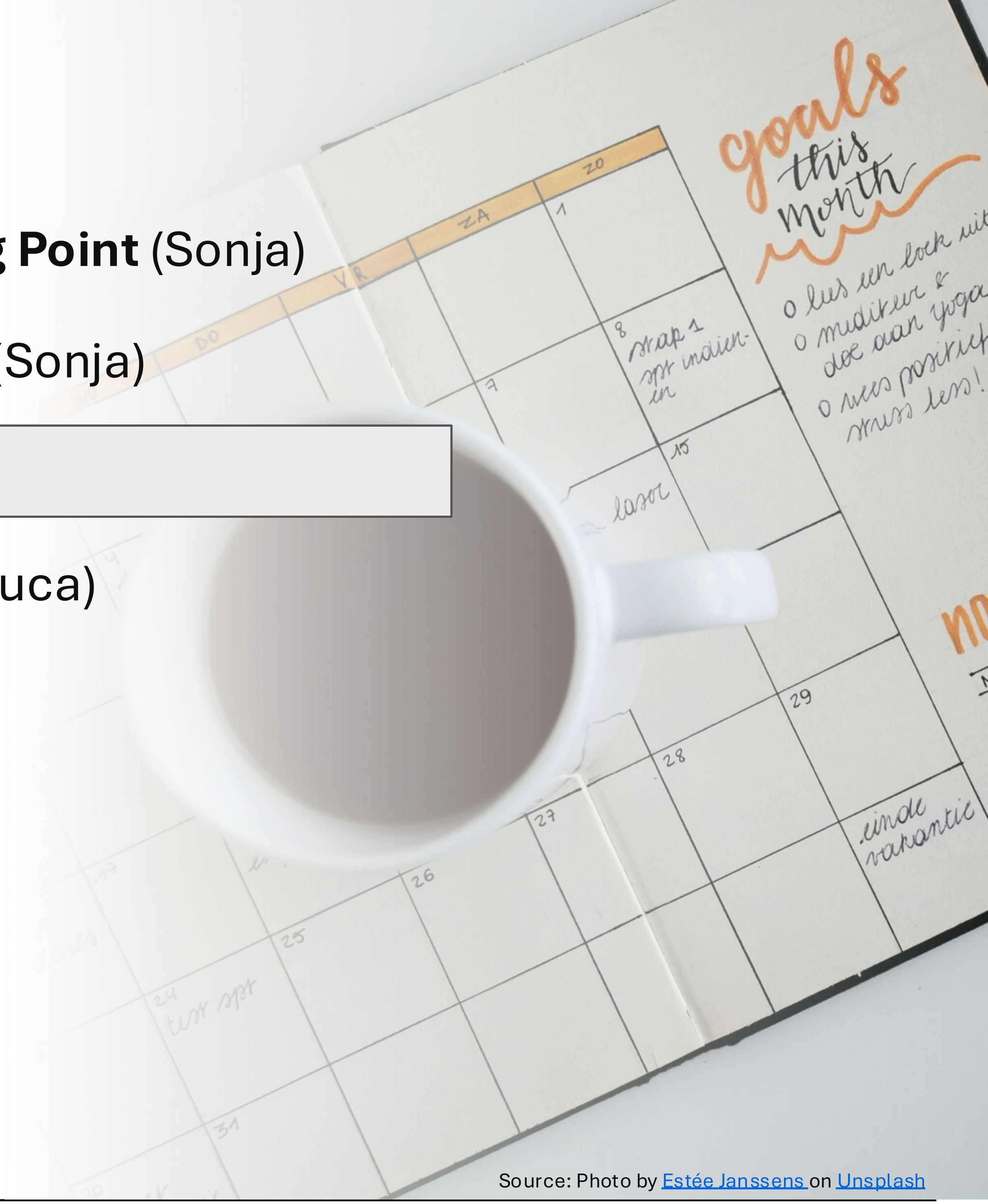
ABSTRACT

In this paper, we extend a recently developed Basis-Oriented time series aggregation approach for aggregating input-data in power system optimization models which has proven to be exact in simple economic dispatch problems. We extend this methodology to include network and ramping constraints, for the latter, to handle temporal linking, we developed a heuristic that, in its current version, relies on the dual solution to find a

Sources: **D. Cardona-Vasquez** et al. "Enhancing time series aggregation for power system optimization models: Incorporating network and ramping constraints." Electric Power Systems Research 230 (2024): 110267.
D. Cardona-Vasquez et al. "Disaggregation of energy system optimization models using machine learning for identification of active constraints." Sustainable Energy, Grids and Networks 43 (2025): 101772.

Agenda

I	Time Series Aggregation. **Motivation** & **Starting Point** (Sonja)
II	Extension to **Network** & **Ramping** Constraints (Sonja)
III	Extension to **Storage** Constraints (Thomas)
IV	Time Series Aggregation with **Bounded Error** (Luca)
V	Extension to **Grid Aggregation** (Benjamin)
VI	Conclusions (Sonja)



Source: Photo by Estée Janssens on Unsplash

Ramping versus Storage Time-Coupling Constraints

Ramping

- **Inequality time-coupling** constraints
- **Active** or **inactive**

Storage

- **Equality time-coupling** constraint
- **Always active**, i.e., during all periods

Questions addressed in this talk

- Exact TSA with storage: criteria for disaggregation and aggregation?
- Practical implementation?

Source: Photos by Alex Simpson on Unsplash and Vardan Papikyan on Unsplash

Source: **T. Klatzer** et al. "Towards Exact Temporal Aggregation of Time-Coupled Energy Storage Models via Active Constraint Set Identification and Machine Learning." arXiv preprint arXiv:2510.14451 (2025).

The methodology extends to storage constraints

Disaggregation into independent submodels

Formulation

Storage constraints

$$z_i = \min \sum_{r\in\mathcal{R}^i} W_r\Big(\sum_{g\in\mathcal{G}} C_g p_{r,g} + \sum_{s\in\mathcal{S}} C^{\mathrm{d}}_s p^{\mathrm{d}}_{r,s} + C^{\mathrm{ns}} p^{\mathrm{ns}}_r\Big) \quad (1\mathrm{a})$$

$$\sum_{g\in\mathcal{G}} p_{r,g} + \sum_{s\in\mathcal{S}} \big(p^{\mathrm{d}}_{r,s} - p^{\mathrm{c}}_{r,s}\big) + p^{\mathrm{ns}}_r = \widetilde{D}_r \quad (\mu^{\mathrm{bal}}_r) \quad \forall r\in\mathcal{R}^i \quad (1\mathrm{b})$$

$$0 \le p_{r,t} \le \overline{P}_t \quad (\underline{\lambda}^{\mathrm{t}}_{r,t}, \overline{\lambda}^{\mathrm{t}}_{r,t}) \quad \forall r\in\mathcal{R}^i, t\in\mathcal{G}^{\mathcal{T}} \quad (1\mathrm{c})$$

$$0 \le p_{r,v} \le \overline{P}_v \widetilde{F}_{r,v} \quad (\underline{\lambda}^{\mathrm{v}}_{r,v}, \overline{\lambda}^{\mathrm{v}}_{r,v}) \quad \forall r\in\mathcal{R}^i, v\in\mathcal{G}^{\mathcal{V}} \quad (1\mathrm{d})$$

$$0 \le p^{\mathrm{ns}}_r \le \widetilde{D}_r \quad (\underline{\lambda}^{\mathrm{ns}}_r, \overline{\lambda}^{\mathrm{ns}}_r) \quad \forall r\in\mathcal{R}^i \quad (1\mathrm{e})$$

$$e_{r,s} = \underline{E}_s + W_r\big(\eta^{\mathrm{c}}_s p^{\mathrm{c}}_{r,s} - p^{\mathrm{d}}_{r,s}/\eta^{\mathrm{d}}_s\big) \quad (\mu^{\mathrm{ini}}_{r,s}) \quad r=1,\ \forall s\in\mathcal{S} \quad (1\mathrm{f})$$

$$e_{r,s} = e_{r-1,s} + W_r\big(\eta^{\mathrm{c}}_s p^{\mathrm{c}}_{r,s} - p^{\mathrm{d}}_{r,s}/\eta^{\mathrm{d}}_s\big) \quad (\mu^{\mathrm{intra}}_{r,s}) \quad \forall r=2,\dots,R^i,\ \forall s\in\mathcal{S} \quad (1\mathrm{g})$$

$$e_{r,s} = \underline{E}_s \quad (\mu^{\mathrm{fin}}_{r,s}) \quad r=R^i,\ \forall s\in\mathcal{S} \quad (1\mathrm{h})$$

$$\underline{E}_s \le e_{r,s} \le \overline{E}_s \quad (\underline{\lambda}^{\mathrm{soc}}_{r,s}, \overline{\lambda}^{\mathrm{soc}}_{r,s}) \quad \forall r\in\mathcal{R}^i, s\in\mathcal{S} \quad (1\mathrm{i})$$

$$0 \le p^{\mathrm{c}}_{r,s} \le \overline{P}^{\mathrm{c}}_s \quad (\underline{\lambda}^{\mathrm{c}}_{r,s}, \overline{\lambda}^{\mathrm{c}}_{r,s}) \quad \forall r\in\mathcal{R}^i, s\in\mathcal{S} \quad (1\mathrm{j})$$

$$0 \le p^{\mathrm{d}}_{r,s} \le \overline{P}^{\mathrm{d}}_s \quad (\underline{\lambda}^{\mathrm{d}}_{r,s}, \overline{\lambda}^{\mathrm{d}}_{r,s}) \quad \forall r\in\mathcal{R}^i, s\in\mathcal{S} \quad (1\mathrm{k})$$

Model Structure

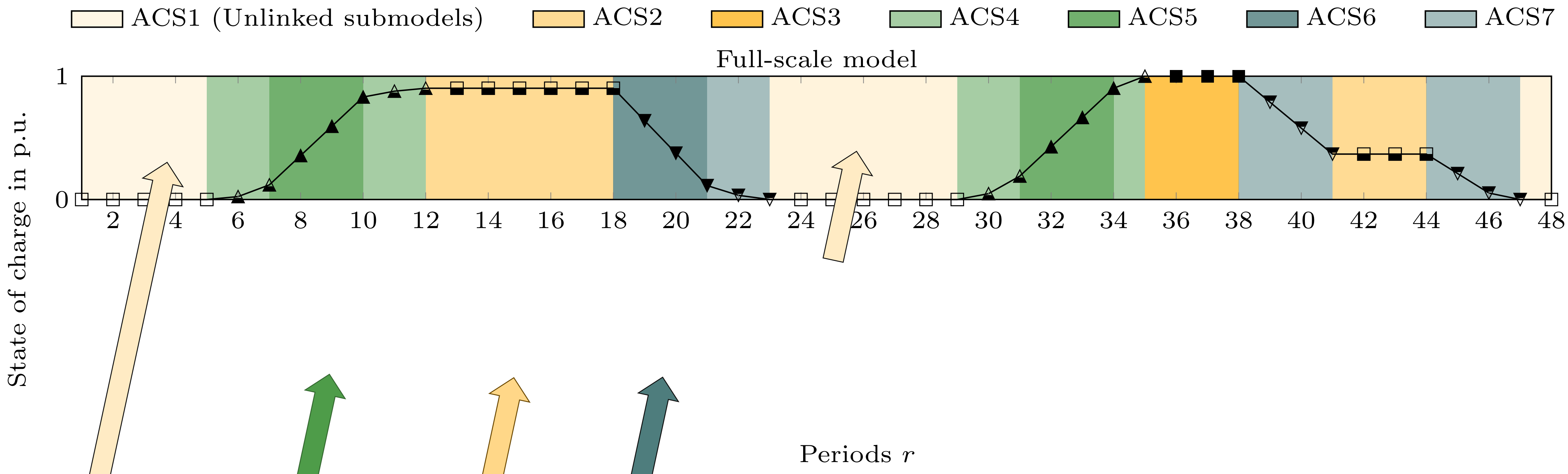
$$\left(\begin{array}{ccc:ccc} \multicolumn{3}{c}{\overbrace{\hspace{4cm}}^{r-1}} & \multicolumn{3}{c}{\overbrace{\hspace{4cm}}^{r}} \\ 1 & -W_{r-1}\eta^{\mathrm{c}}_s & W_{r-1}/\eta^{\mathrm{d}}_s & -1 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 1 & -W_r\eta^{\mathrm{c}}_s & W_r/\eta^{\mathrm{d}}_s \end{array}\right) \begin{pmatrix} e_{r-1,s} \\ p^{\mathrm{c}}_{r-1,s} \\ p^{\mathrm{d}}_{r-1,s} \\ e_{r,s} \\ p^{\mathrm{c}}_{r,s} \\ p^{\mathrm{d}}_{r,s} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

- **Exact disaggregation** into **independent submodels** can happen when the storage state of charge is zero (in two consecutive time steps).
- Disaggregation allows for **parallelization** of submodels.

Source: **T. Klatzer** et al. "Towards Exact Temporal Aggregation of Time-Coupled Energy Storage Models via Active Constraint Set Identification and Machine Learning." arXiv preprint arXiv:2510.14451 (2025).

The methodology extends to storage constraints

TSA within a submodel according to active constraint sets (ACS)



1. **Disaggregation** into unlinked submodels happens when storage is empty (ACS1)

2. **Aggregation** within a submodel: consecutive time periods can be **aggregated perfectly** within the **same active constraint sets** (e.g., **ACS2 storage idle**; **ACS4 storage charging**; **ACS6 storage discharging at full capacity**)

Source: **T. Klatzer** et al. "Towards Exact Temporal Aggregation of Time-Coupled Energy Storage Models via Active Constraint Set Identification and Machine Learning." arXiv preprint arXiv:2510.14451 (2025).

Conceptual 4-step argument for exact TSA with storage

In the case of perfect information

Step 1: Compute full-scale model and identify **ACSs and corresponding dual information** for each period.

- **Full-scale model**: $|\mathcal{I}| = 1, \ |\mathcal{R}^1| = 8736, \ W_r = 1$
- **Computational burden**: 187×10^{-3} work units

Step 2: **Disaggregate** full-scale model into independent **unlinked** submodels $\mathcal{U}$ and **linked** submodels $\mathcal{L}$ such that $\mathcal{I} = \mathcal{U} \cup \mathcal{L}$.

- **Unlinked** submodels represent 1377 periods using: $i \in \mathcal{U} = \{1, 2, \ldots, 1377\}, \ |\mathcal{R}^i| \in \mathcal{R}^{\mathcal{U}} = \{1, 1, \ldots, 1\}, \ W_r = 1$
- **Linked** submodels represent 7359 periods using: $i \in \mathcal{L} = \{1, 2 \ldots, 176\}, \ |\mathcal{R}^i| \in \mathcal{R}^{\mathcal{L}} = \{4, 5, \ldots, 231, 253, 395\}, \ W_r = 1$
- **Lower bound on computational burden**: 77×10^{-4} work units $\rightarrow$ **24-fold** reduction versus full-scale model

Step 3: **Aggregate unlinked** submodels, and aggregate adjacent periods within **linked** submodels sharing the same ACS and duals.

- **Unlinked** submodels aggregate into $|\mathcal{R}^i| \in \mathcal{R}^{\mathcal{U}} = \{1, 1, \ldots, 1\}$ with $W_r \geq 1 \ \rightarrow \ \sum_{|\mathcal{R}^i| \in \mathcal{R}^{\mathcal{U}}} = 8$ periods
- **Linked** submodels aggregate into $|\mathcal{R}^i| \in \mathcal{R}^{\mathcal{L}} = \{2, 2, \ldots, 31, 27, 5\}$ with $W_r \geq 1 \ \rightarrow \ \sum_{|\mathcal{R}^i| \in \mathcal{R}^{\mathcal{L}}} = 925$ periods

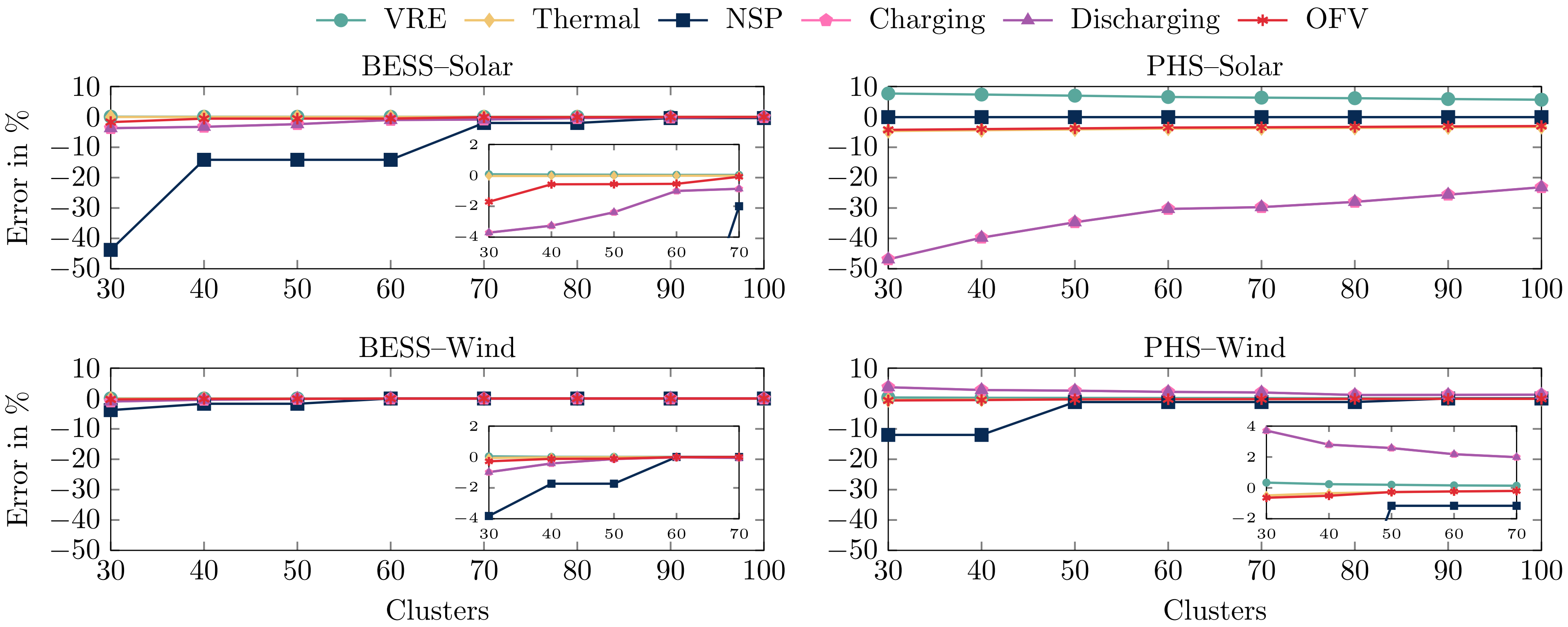
Step 4: **Solve disaggregated and subsequently aggregated submodels** in parallel and compute total OFV $Z = \sum_{i \in \mathcal{I}} z_i$.

- **TSA** with 933 periods achieves **zero error** in the OFV and aggregated decision variables
- **Lower bound on computational burden**: 51×10^{-5} work units $\rightarrow$ **369-fold** reduction versus full-scale model

Source: **T. Klatzer** et al. "Towards Exact Temporal Aggregation of Time-Coupled Energy Storage Models via Active Constraint Set Identification and Machine Learning." arXiv preprint arXiv:2510.14451 (2025).

Exact TSA with storage in practice

ML-based disaggregation (random forest classifier) and aggregation (hierarchical clustering)



Results

- **Relative output error** depends on the number of clusters in aggregation (**-5** to **-1%** for the **objective function value** using only **30 clusters**)
- 20 to 120-fold **computational speed-up** versus full-scale model

Source: **T. Klatzer** et al. "Towards Exact Temporal Aggregation of Time-Coupled Energy Storage Models via Active Constraint Set Identification and Machine Learning." arXiv preprint arXiv:2510.14451 (2025).

Key messages

Storage Constraints

Exact TSA with storage

- **NetZero-Opt demonstrates: exact TSA** with **storage** is **possible!**
- **Disaggregation** allows for **parallelization – aggregation** further reduces **problem size**

Future research

- Extend to **investment, multi-storage** & **long-duration energy storage** problems
- Explore other **ML-driven approaches** to identify suitable periods for **disaggregation** & **ML-driven aggregation**

🚀 Check out our paper

Towards Exact Temporal Aggregation of Time-Coupled Energy Storage Models via Active Constraint Set Identification and Machine Learning

Thomas Klatzer, David Cardona-Vasquez, Luca Santosuosso and Sonja Wogrin
Institute of Electricity Economics and Energy Innovation
Graz University of Technology
Graz, Austria
{thomas.klatzer, david.cardonavasquez, luca.santosuosso, wogrin}@tugraz.at

***Abstract*—Time series aggregation (TSA) methods aim to construct temporally aggregated optimization models that accurately represent the output space of their full-scale counterparts while using a significantly reduced dimensionality in the input space. This paper presents the first approach that achieves an exact TSA of a full-scale power system model – even in the presence of energy storage time-coupling constraints – by leveraging active constraint sets and dual information. This advances the state of the art beyond existing TSA approaches, which typically cannot guarantee solution accuracy or rely on iterative procedures to determine the required number of representative periods. To bridge the gap between our theoretical analysis and their practical application, we employ machine learning approaches, i.e., classification and clustering, to inform TSA in models that co-schedule variable renewable energy sources and energy storage. Numerical results demonstrate substantially improved computational performance relative to the full-scale model, while maintaining high solution accuracy.**

$\widetilde{F}_{r,v}$	Average capacity factor of v in r (p.u.).
$\overline{P}_g$	Max. power generation of g (MW).
$\overline{P}^{c}_{s}, \overline{P}^{d}_{s}$	Max. charging/discharging power of s (MW).
$\eta^{c}_{s}, \eta^{d}_{s}$	Charging/discharging efficiency of s (–).
$\underline{E}_s, \overline{E}_s$	Min./max. state of charge of s (MWh).

C. Primal Variables

$p_{r,g}$	Power generation of g (MW).
p^{ns}_{r}	Non-supplied power (MW).
$p^{c}_{r,s}, p^{d}_{r,s}$	Charging/discharging power of s (MW).
$e_{r,s}$	State of charge of s (MWh).

D. Dual Variables

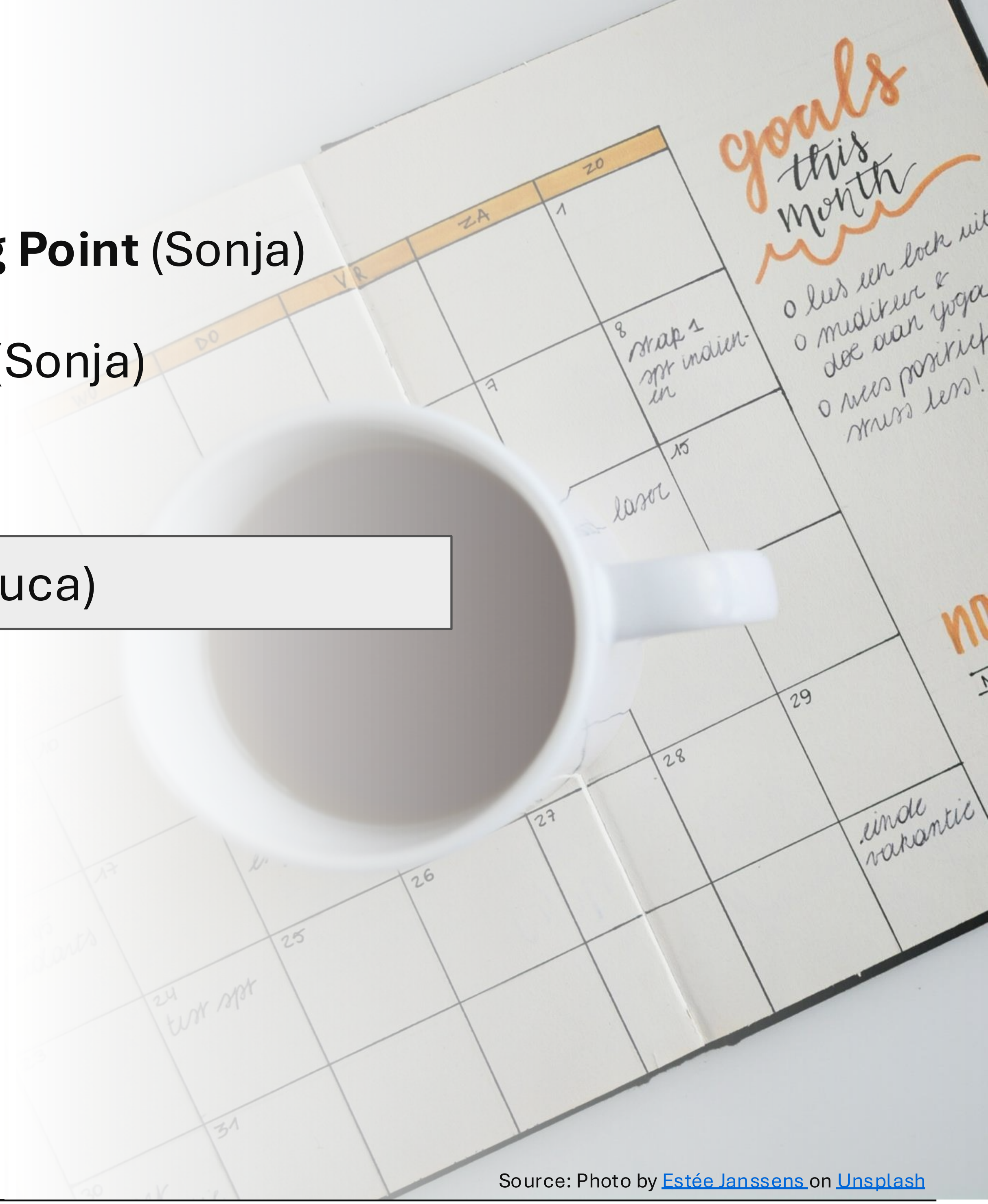
μ^{bal}_{r}	Dual of the power balance constraint.
$\underline{\lambda}^{v}_{r,v}, \overline{\lambda}^{v}_{r,v}$	Duals of lower/upper bounds of $p_{r,v}$.

Source: Photos by Vardan Papikyan on Unsplash

Source: **T. Klatzer** et al. "Towards Exact Temporal Aggregation of Time-Coupled Energy Storage Models via Active Constraint Set Identification and Machine Learning." arXiv preprint arXiv:2510.14451 (2025).

Agenda

- **I** Time Series Aggregation. **Motivation** & **Starting Point** (Sonja)
- **II** Extension to **Network** & **Ramping** Constraints (Sonja)
- **III** Extension to **Storage** Constraints (Thomas)
- **IV** Time Series Aggregation with **Bounded Error** (Luca)
  - Bounded error in the objective function
  - What are we clustering for?
  - From static optimization to optimal control
- **V** Extension to **Grid Aggregation** (Benjamin)
- **VI** Conclusions (Sonja)



Source: Photo by Estée Janssens on Unsplash

Time series aggregation with bounded error

Motivation: Traditional TSA does not guarantee bounded error! (**heuristic**)

Bounded error in the objective function

We demonstrate that an **appropriately* constructed aggregated model** yields a lower bound on the optimal objective function value of the full-scale model.

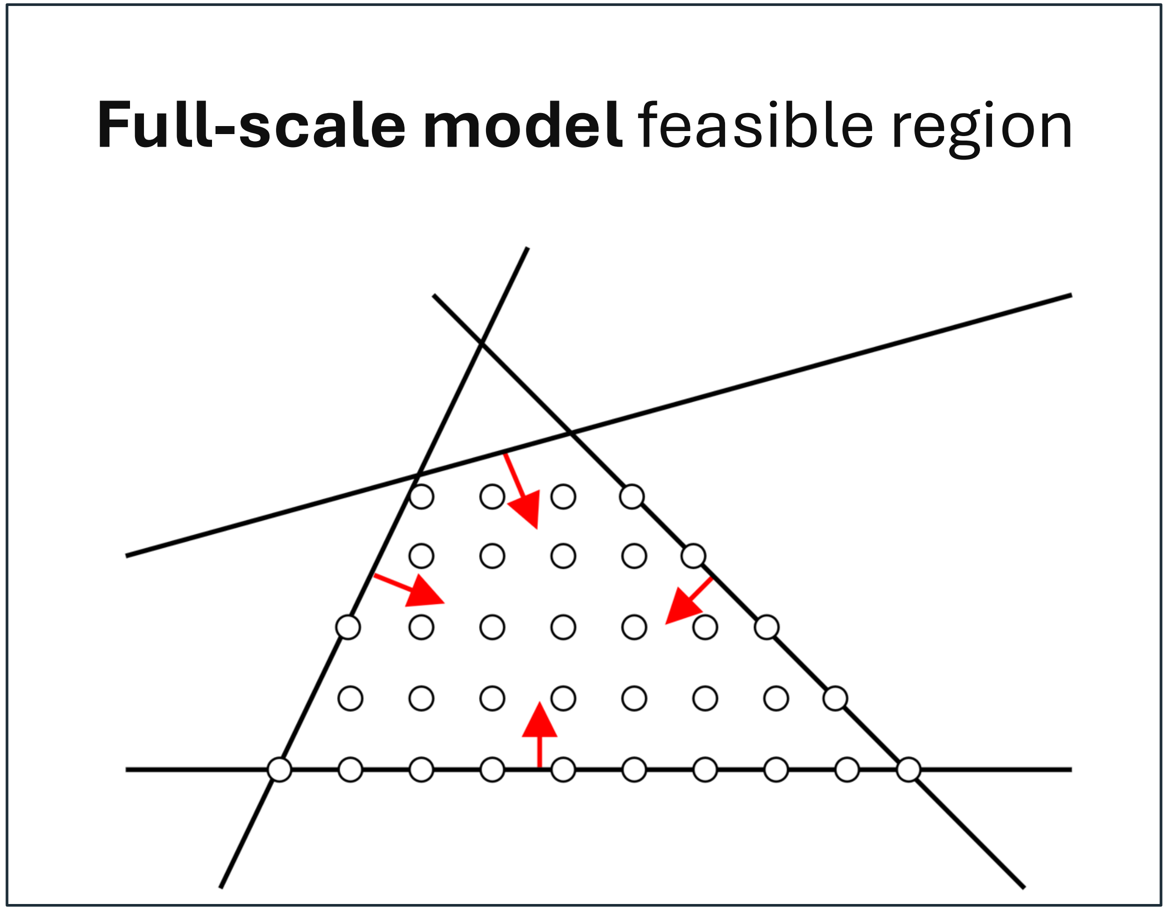
D. Main Theoretical Result

This subsection presents our main theoretical result.

Proposition 1. Let $\boldsymbol{z}$ be a feasible solution to the full-scale model (2). Let $\hat{\boldsymbol{z}}$ be derived from $\boldsymbol{z}$ accordingly to (3)–(6). Then, $\hat{\boldsymbol{z}}$ is a feasible solution to the aggregated model (8) and it holds that

$$J\left(\boldsymbol{z}\right) = \hat{J}\left(\hat{\boldsymbol{z}}\right).$$

Proof. Using (5) and (6), the power balance constraints (8b)



***Source: L. Santosuosso** and S. Wogrin, "Optimal virtual power plant investment planning via time series aggregation with bounded error," *arXiv preprint arXiv:2504.19699* (2025).

Bounded error in the objective function

We demonstrate that an **appropriately* constructed aggregated model** yields a lower bound on the optimal objective function value of the full-scale model.

D. Main Theoretical Result

This subsection presents our r

Proposition 1. Let z be a feas

model (2). Let $\hat{z}$ be derived fr

Then, $\hat{z}$ is a feasible solution to the aggregated model (8) and it holds that

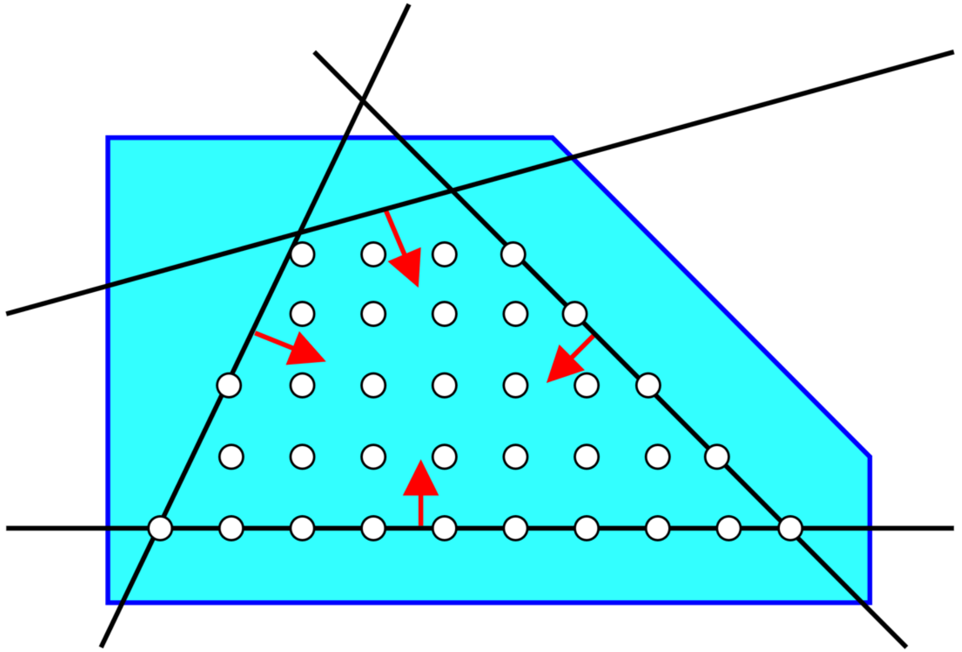
$$J(z) = \hat{J}(\hat{z}) \,.$$

Proof. Using (5) and (6), the power balance constraints (8b)

Full-scale obj.
=
Aggregated obj.

***Source: L. Santosuosso** and S. Wogrin, "Optimal virtual power plant investment planning via time series aggregation with bounded error," *arXiv preprint arXiv:2504.19699* (2025).

The **aggregated model** is a **relaxation** of the full-scale model



Bounded error in the objective function

Algorithm 1 Time Series Aggregation with Bounded Error in the Objective Function

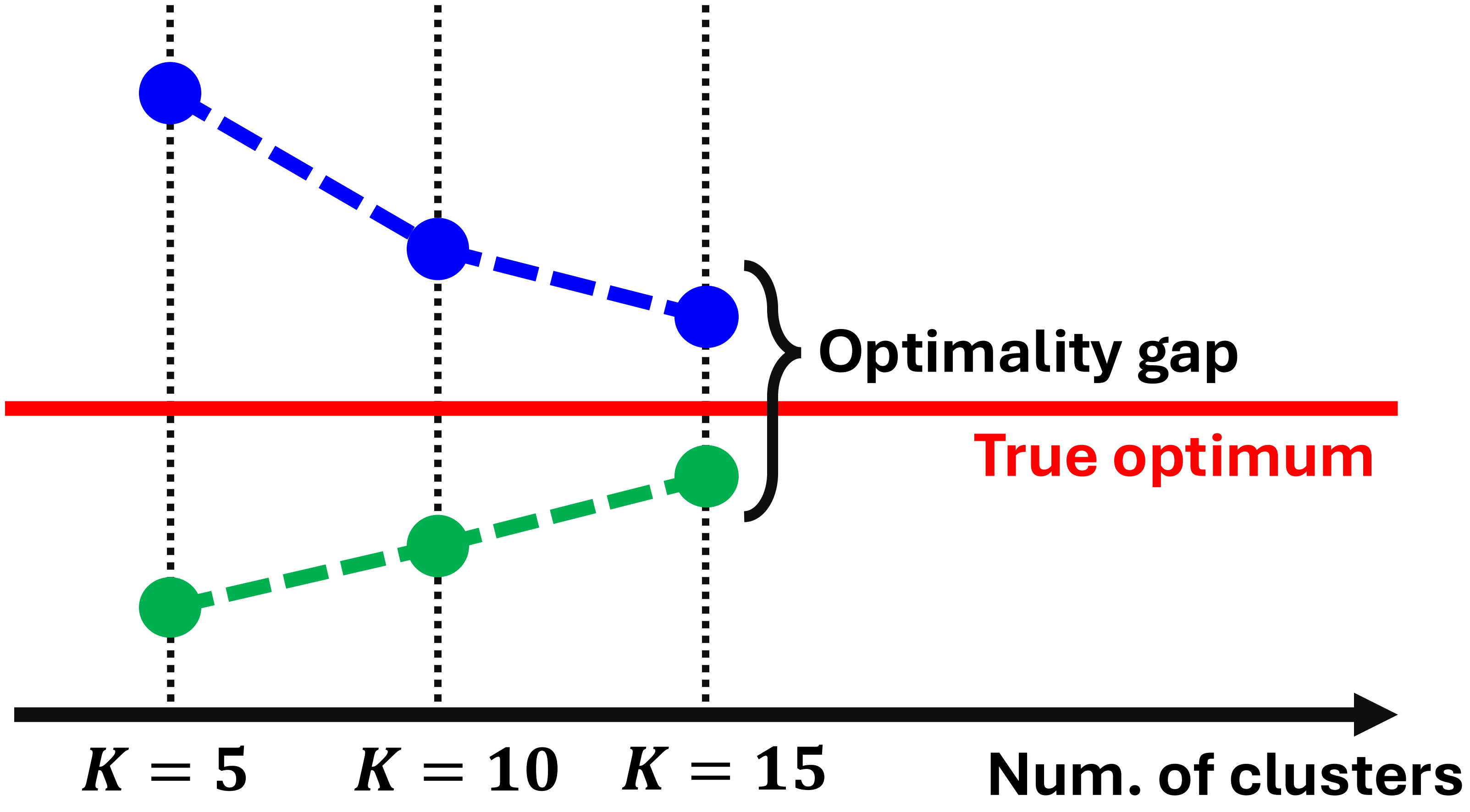
Input: Parameters $\{F_{g,t}, D_t, \overline{X}_g, \underline{X}_g \mid g \in \boldsymbol{G}, t \in \boldsymbol{T}\}$, initial number of clusters K^0, step size α, optimality threshold

(1) TSA with any clustering technique + solve aggregated model

(2) Fix binaries from aggregated model and solve operational prob.

10: $J^{\mathrm{LB}^{i+1}} \leftarrow \max\left(J^{\mathrm{LB}^i}, \tilde{J}^{\mathrm{LB}}\right)$;
11: $J^{\mathrm{UB}^{i+1}} \leftarrow \min\left(J^{\mathrm{UB}^i}, \tilde{J}^{\mathrm{UB}}\right)$;
12: **end if**
13: $\epsilon^{i+1} \leftarrow$ Evaluate (11) for $J^{\mathrm{LB}^{i+1}}$ and $J^{\mathrm{UB}^{i+1}}$;
14: $K^{i+1} \leftarrow K^i + \alpha \lfloor \epsilon^{i+1} \rfloor$;
15: $i \leftarrow i + 1$;
16: **end while**
17: $J^{\mathrm{UB}\star} \leftarrow J^{\mathrm{UB}^i}$ and $J^{\mathrm{LB}\star} \leftarrow J^{\mathrm{LB}^i}$;

Source: **L. Santosuosso** and S. Wogrin, "Optimal virtual power plant investment planning via time series aggregation with bounded error," *arXiv preprint arXiv:2504.19699* (2025).



Bounded error in the objective function

Algorithm 1 Time Series Aggregation with Bounded Error in the Objective Function

Input: Parameters $\{F_{g,t}, D_t, \overline{X}_g, \underline{X}_g \,|\, g \in \boldsymbol{G}, t \in \boldsymbol{T}\}$, initial number of clusters K^0, step size α, optimality threshold ϵ^{thr}, and maximum number of iterations $\overline{I}$

(1) TSA with any clustering technique + solve aggregated model

(2) Fix binaries from aggregated model and solve operational prob.

$J^{\text{LB}^{i+1}} \leftarrow \max\left(J^{\text{LB}^i}, \tilde{J}^{\text{LB}}\right)$;
$J^{\text{UB}^{i+1}} \leftarrow \min\left(J^{\text{UB}^i}, \tilde{J}^{\text{UB}}\right)$;
end if
$\epsilon^{i+1} \leftarrow$ Evaluate (11) for $J^{\text{LB}^{i+1}}$ and $J^{\text{UB}^{i+1}}$;
$K^{i+1} \leftarrow K^i + \alpha \lfloor \epsilon^{i+1} \rceil$;
$i \leftarrow i + 1$;
end while
$J^{\text{UB}\star} \leftarrow J^{\text{UB}^i}$ and $J^{\text{LB}\star} \leftarrow J^{\text{LB}^i}$;

Source: **L. Santosuosso** and S. Wogrin, "Optimal virtual power plant investment planning via time series aggregation with bounded error," *arXiv preprint arXiv:2504.19699* (2025).

Lower bound:

- From a **reduced MILP** model.
- Valid for **any** clustering technique.

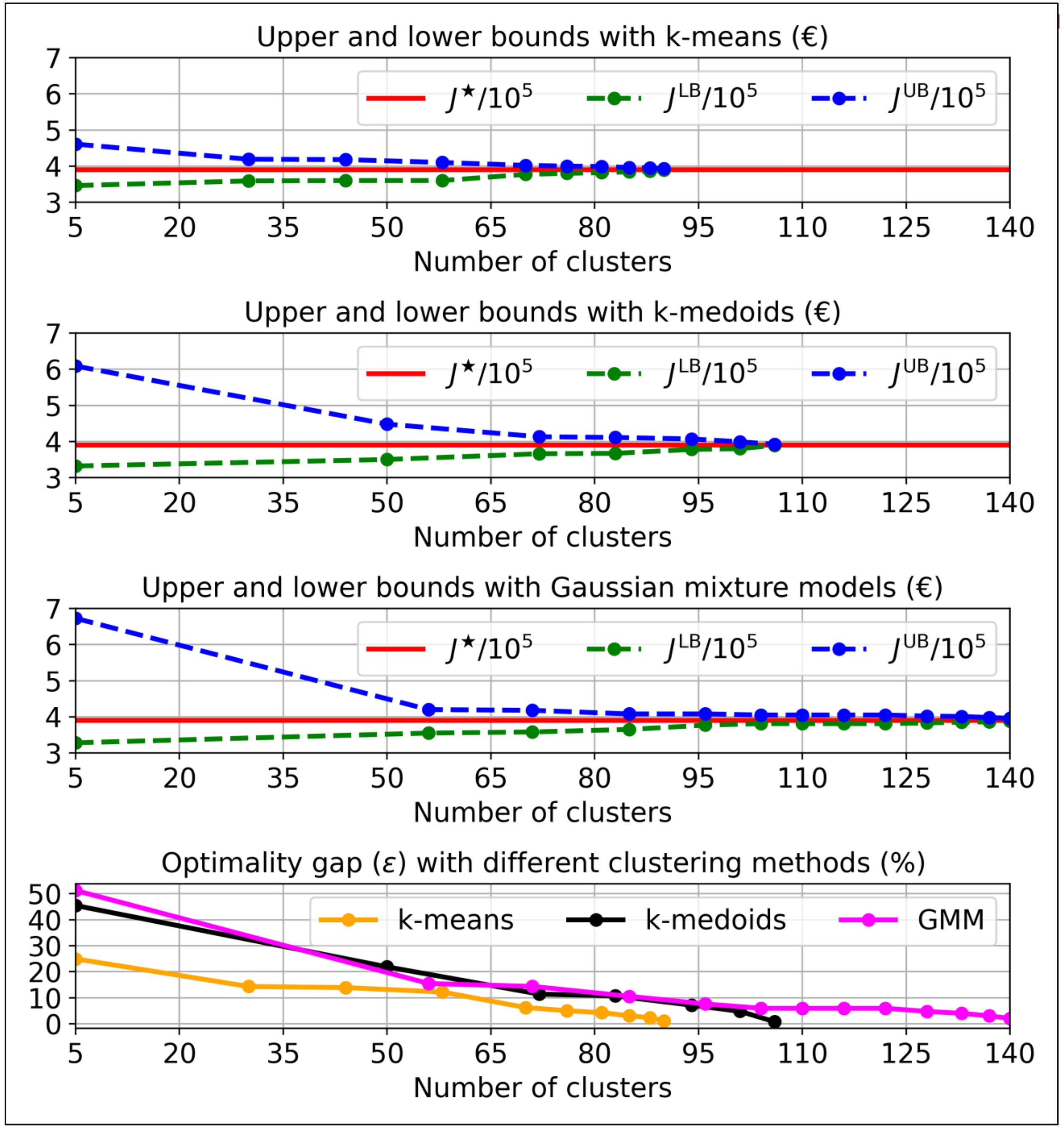
Upper bound:

- From an **LP** model.
- Provides a **feasible solution** at each iteration.

Results

- The **bounds validity** is independent of the clustering technique.
- Over **50%** computational savings!

Settings	Computational time (s)			
	F-S	TSA with bounded error		
		k-means	k-medoids	GMM
$\lvert T\rvert = 8760$, $\lvert G\rvert = 100$	273	431 (**+58%**)	516 (+89%)	689 (+152%)
$\lvert T\rvert = 17520$, $\lvert G\rvert = 100$	793	956 (**+21%**)	1360 (+72%)	1892 (+139%)
$\lvert T\rvert = 8760$, $\lvert G\rvert = 1000$	4032	3346 (**−17%**)	5113 (+27%)	7629 (+89%)
$\lvert T\rvert = 17520$, $\lvert G\rvert = 1000$	16911	7060 (**−58%**)	11862 (−30%)	15540 (−8%)



Check out our paper

Content: (1) Demonstration of the lower-bound property of the aggregated model; (2) TSA-based solution algorithm with bounded error; (3) Numerical results.

Optimal Virtual Power Plant Investment Planning via Time Series Aggregation with Bounded Error

Luca Santosuosso
Institute of Electricity Economics and Energy Innovation
Graz University of Technology
Graz, Austria
luca.santosuosso@tugraz.at

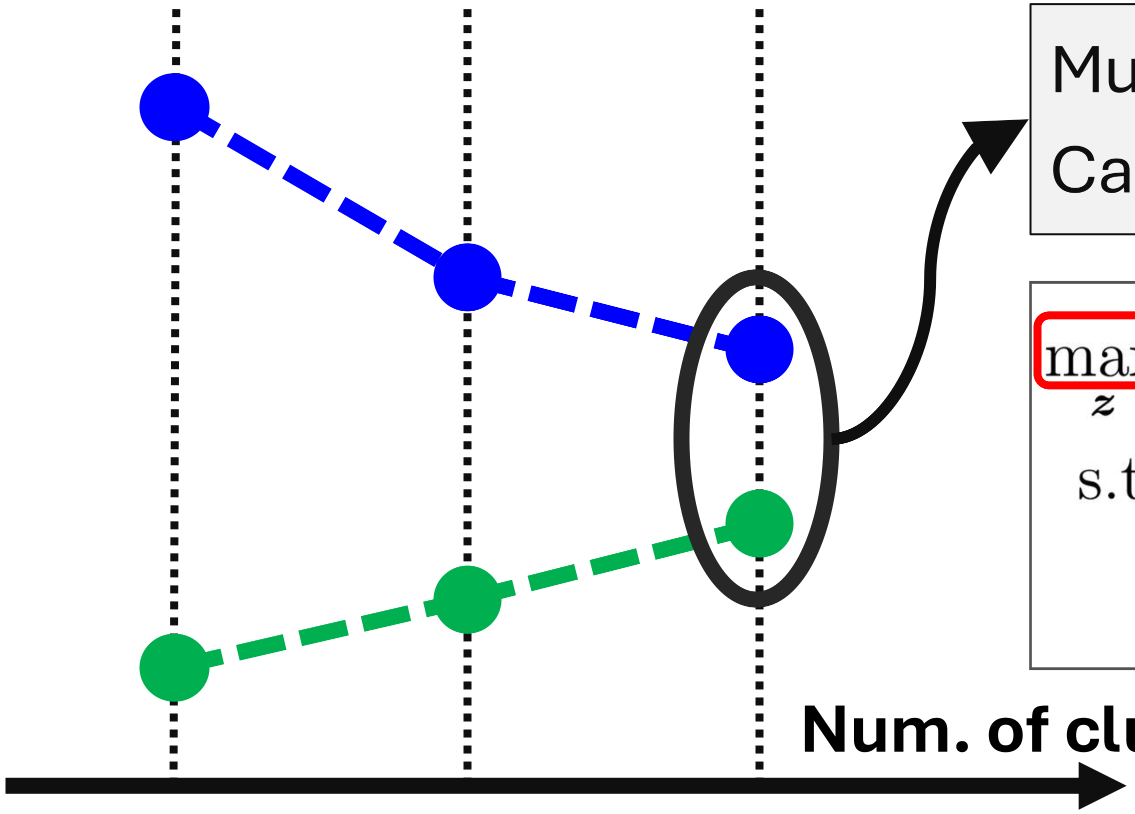
Sonja Wogrin
Institute of Electricity Economics and Energy Innovation
Graz University of Technology
Graz, Austria
wogrin@tugraz.at

***Abstract*—This study addresses the investment planning problem of a virtual power plant (VPP), formulated as a mixed-integer linear programming (MILP) model. As the number of binary variables increases and the investment time horizon extends, the**

tistical characteristics of the input data. Common approaches include k-means [6], k-medoids [7], and hierarchical clustering [8], among others. These methods have been extensively

What are we clustering for?

Our generation expansion planning model is used by **different stakeholders**, each with their own **metric of interest** $M(\boldsymbol{z})$.



Multiple solutions exist within the objective bounds.

Can we find those that minimize / maximize $M(\boldsymbol{z})$?

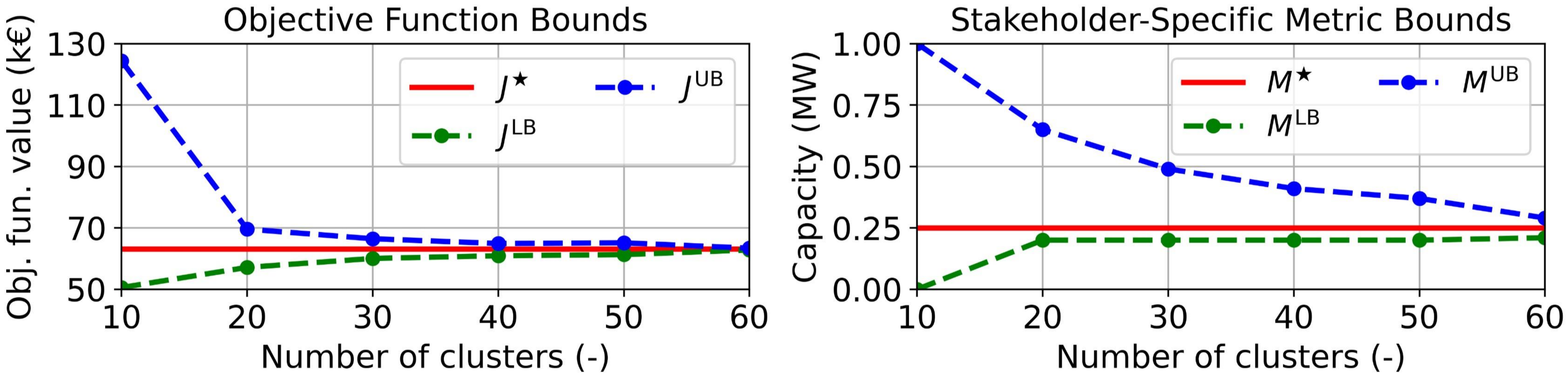
$$\begin{aligned} \max_{\boldsymbol{z}} \quad & M(\boldsymbol{z}) \\ \text{s.t.} \quad & \text{Full-scale dynamics} \\ & J^{\text{LB}} \leq J(\boldsymbol{z}) \leq J^{\text{UB}} \end{aligned}$$

$$\begin{aligned} \min_{\boldsymbol{z}} \quad & M(\boldsymbol{z}) \\ \text{s.t.} \quad & \text{Full-scale dynamics} \\ & J^{\text{LB}} \leq J(\boldsymbol{z}) \leq J^{\text{UB}} \end{aligned}$$

Source: **L. Santosuosso**, B. Klinz, and S. Wogrin, "What are we clustering for? Establishing performance guarantees for time series aggregation in generation expansion planning," *arXiv preprint arXiv:2510.09357* (2025).

What are we clustering for?

Objective function bounds and corresponding bounds for the **stakeholder-specific metric** of interest, namely the capacity investment in a specific storage unit.



Source: **L. Santosuosso**, B. Klinz, and S. Wogrin, "What are we clustering for? Establishing performance guarantees for time series aggregation in generation expansion planning," *arXiv preprint arXiv:2510.09357* (2025).

Check out our paper

What Are We Clustering For? Establishing Performance Guarantees for Time Series Aggregation in Generation Expansion Planning

Luca Santosuosso[a], Bettina Klinz[b], Sonja Wogrin[a]

[a]Institute of Electricity Economics and Energy Innovation, Graz University of Technology, Inffeldgasse 18, Graz, 8010, Austria
[b]Institute for Discrete Mathematics, Graz University of Technology, Steyrergasse 30, Graz, 8010, Austria

Abstract

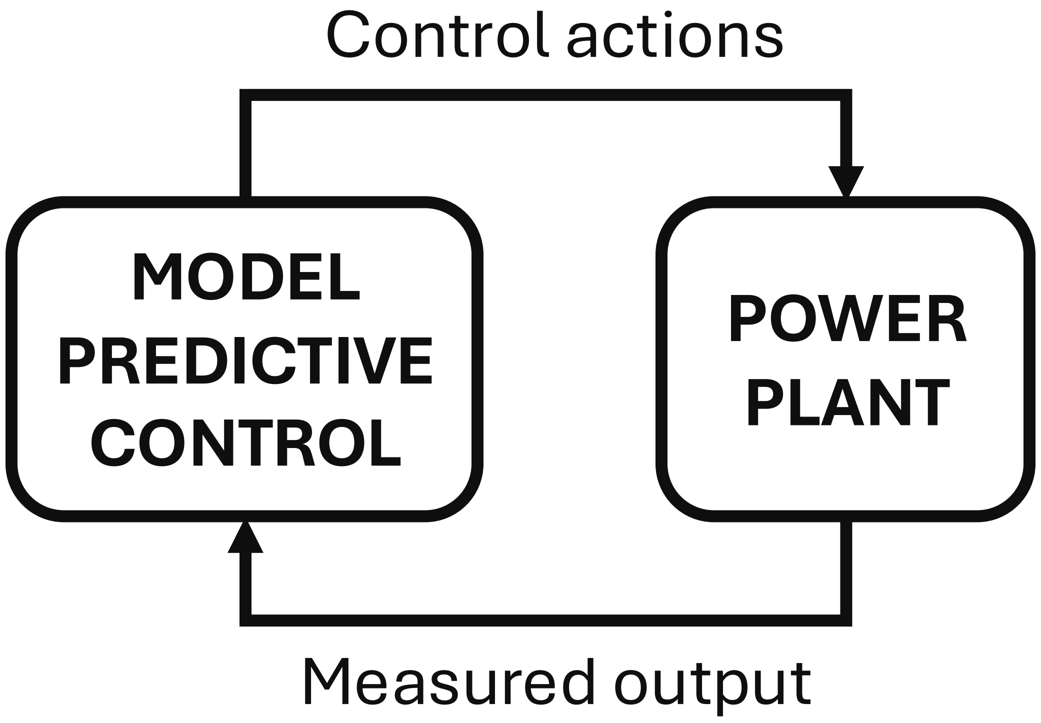
Generation expansion planning (GEP) is a prominent example of capacity expansion problems in operations research. Being generally NP-hard, GEP optimization models can become intractable when nonconvex dynamics, time-coupling constraints, and complex asset interactions are involved. Time series aggregation (TSA) tackles this by reducing temporal complexity via input data clustering. However, existing TSA methods either focus solely on preserving the statistical features of the input data, yielding heuristics without

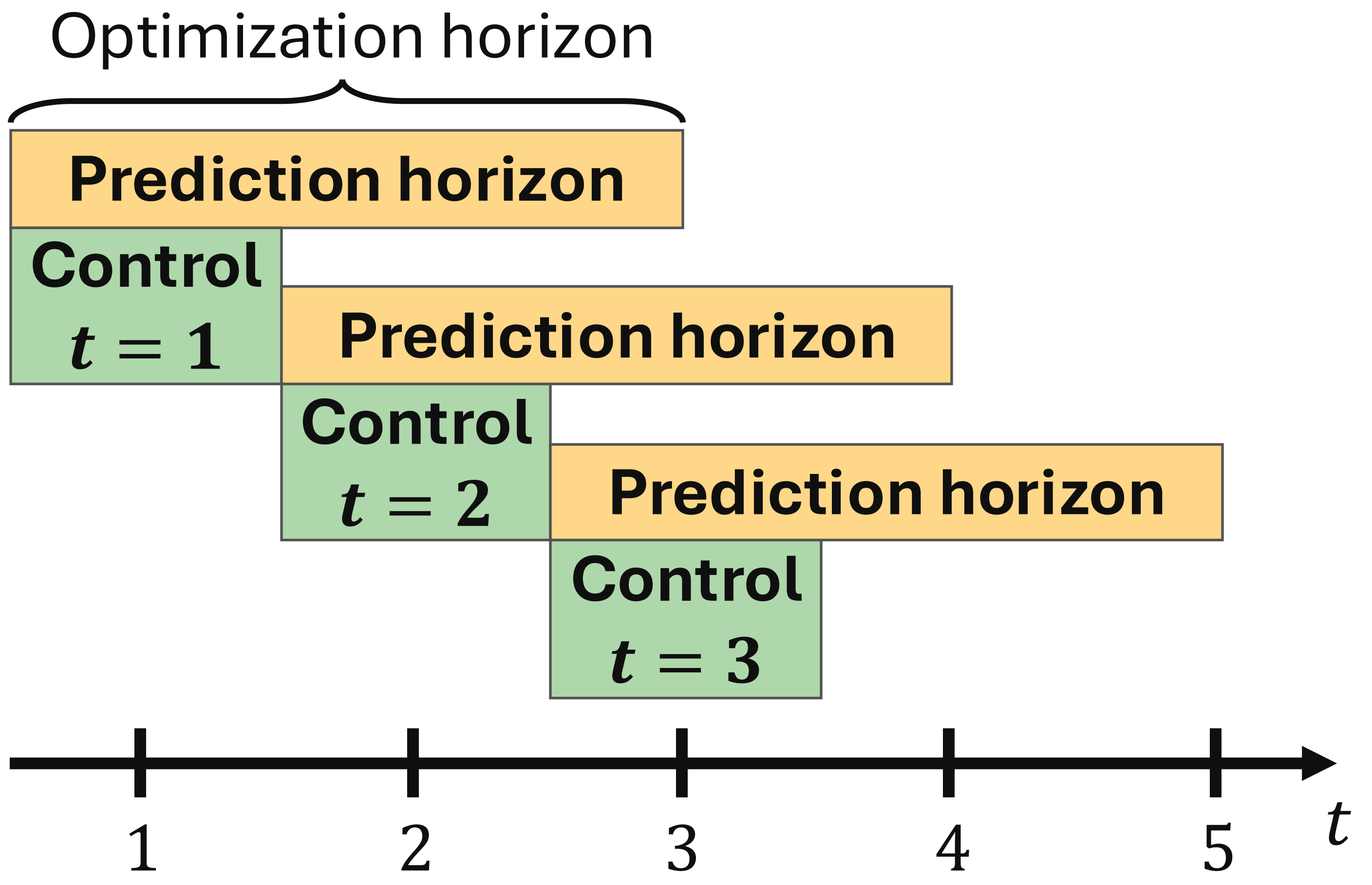
Content:

- Extension of lower-bound aggregated model property to **MIQP** formulations and **storage constraints**.
- TSA-based solution algorithm with **bounded error**.
- A comparison with **Benders**.
- Deriving bounds for stakeholder-specific metrics (***what are we clustering for?***).
- Numerical results.

From static optimization to optimal control

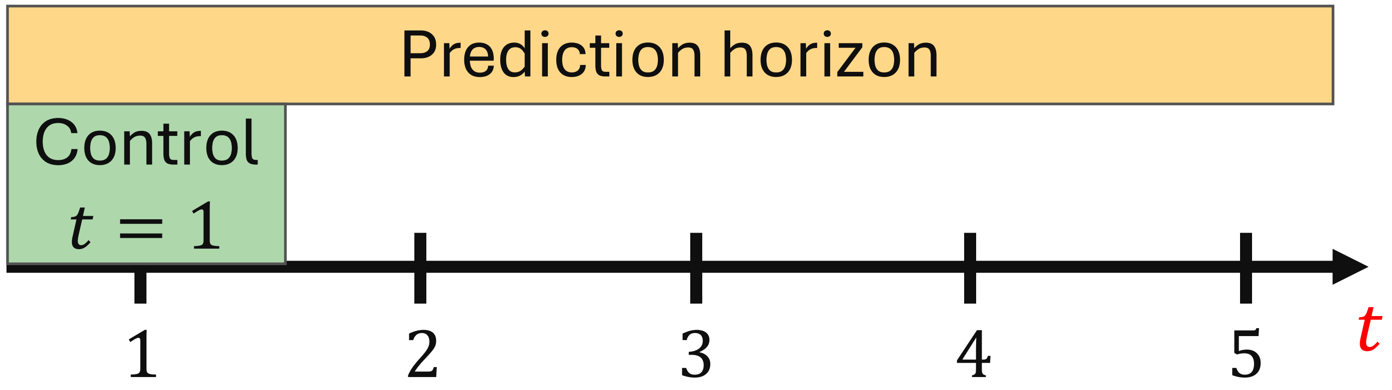
Real-time optimal control problems require **fast, reactive control decisions computed online at high resolution.**



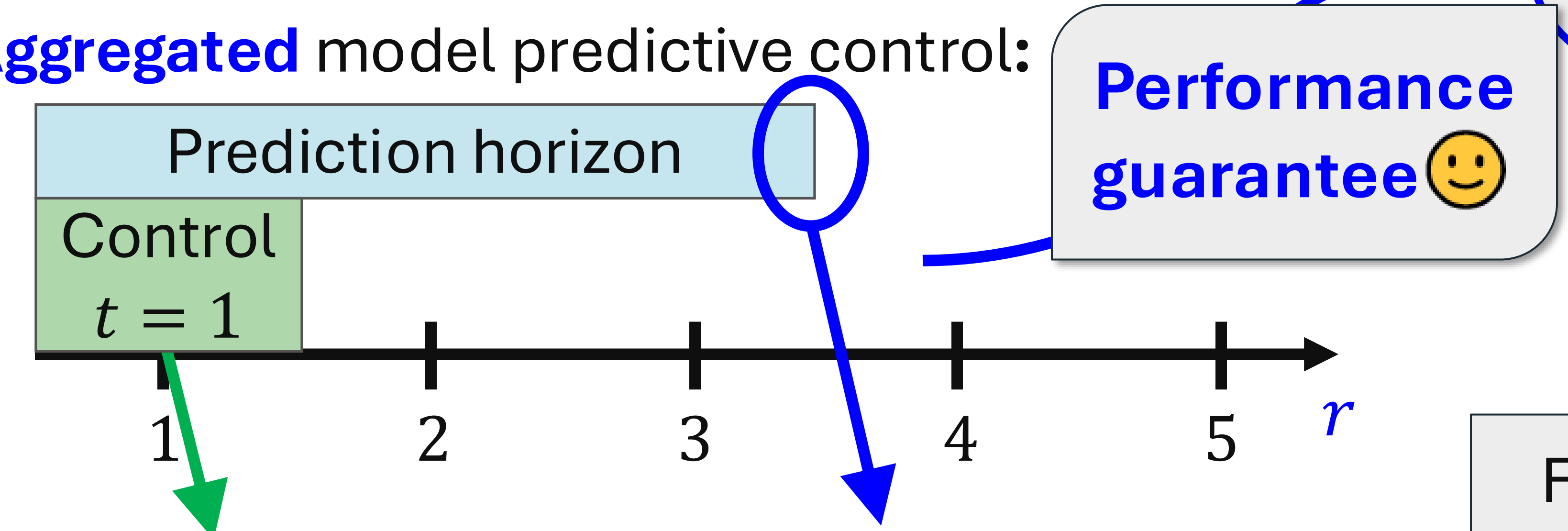


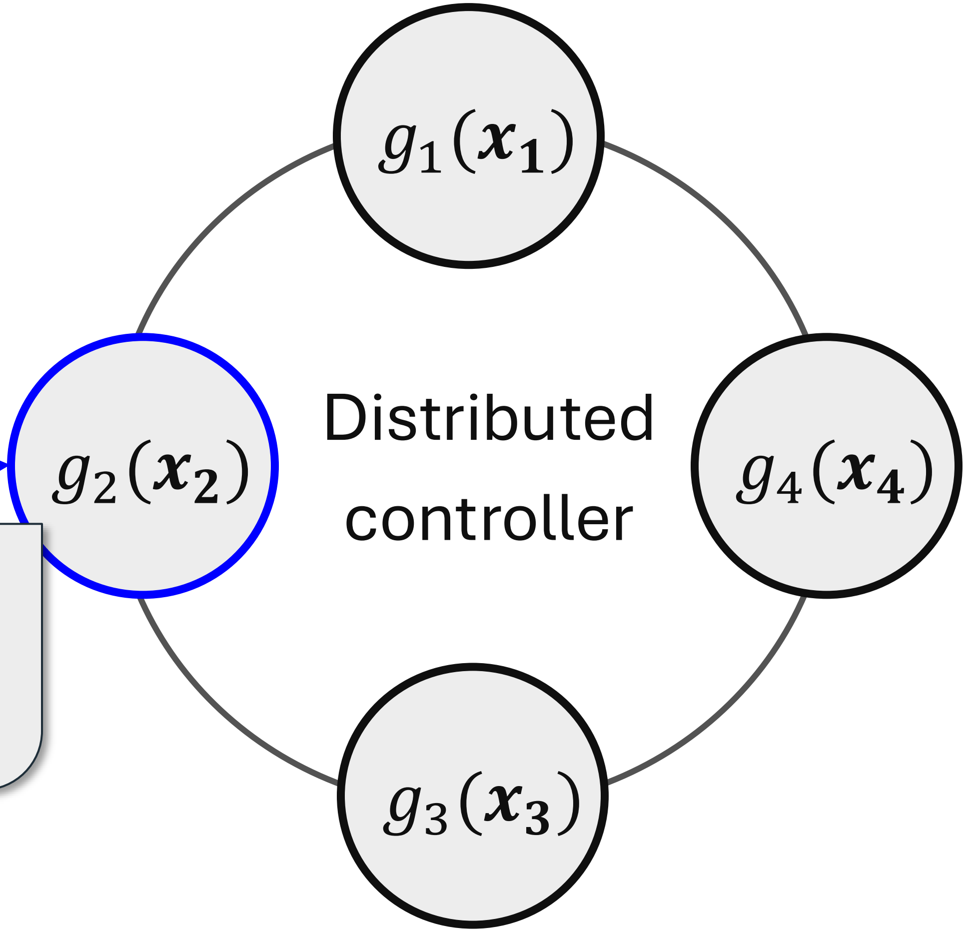
From static optimization to optimal control

Full-scale model predictive control:



Aggregated model predictive control:





Full-scale control granularity

Aggregated prediction horizon tail

Fuse with **distributed control** to crush the **other complexity layers** (e.g., assets, scenarios).

Check out our paper

Content: (1) Performance-guaranteed TSA applied to stochastic MPC; (2) TSA + distributed optimization to address multiple complexity dimensions.

Distributed Stochastic Model Predictive Control with Temporal Aggregation for the Joint Dispatch of Cascaded Hydropower and Renewables

Luca Santosuosso, Sonja Wogrin
Institute of Electricity Economics and Energy Innovation
Graz University of Technology
Graz, Austria
{luca.santosuosso, wogrin}@tugraz.at

Abstract—**This paper addresses the real-time energy dispatch of a hybrid system comprising cascaded hydropower plants, wind, and solar units, jointly participating in the day-ahead energy market under inflow, renewable generation, and price uncer-**

time control [5]. This paper focuses specifically on the latter stage of this sequential decision-making process.

The real-time control of hydropower plants typically oper-

Key messages

Bounded error

Time series aggregation with bounded error

- TSA restores tractability in complex energy models.
- ...but traditional TSA is still **heuristic-driven**. ⚠

NetZero-Opt:

✓ TSA with formal **performance guarantees** (bounded error in objective + solutions). 🙂

✓ Unified framework for **optimization** and **optimal control**. 🚀

🚀 **Check out our papers**

Optimal Virtual Power Plant Investment Planning via Time Series Aggregation with Bounded Error

Luca Santosuosso
Institute of Electricity Economics and Energy Innovation
Graz University of Technology
Graz, Austria
luca.santosuosso@tugraz.at

Sonja Wogrin
Institute of Electricity Economics and Energy Innovation
Graz University of Technology
Graz, Austria
wogrin@tugraz.at

What Are We Clustering For? Establishing Performance Guarantees for Time Series Aggregation in Generation Expansion Planning

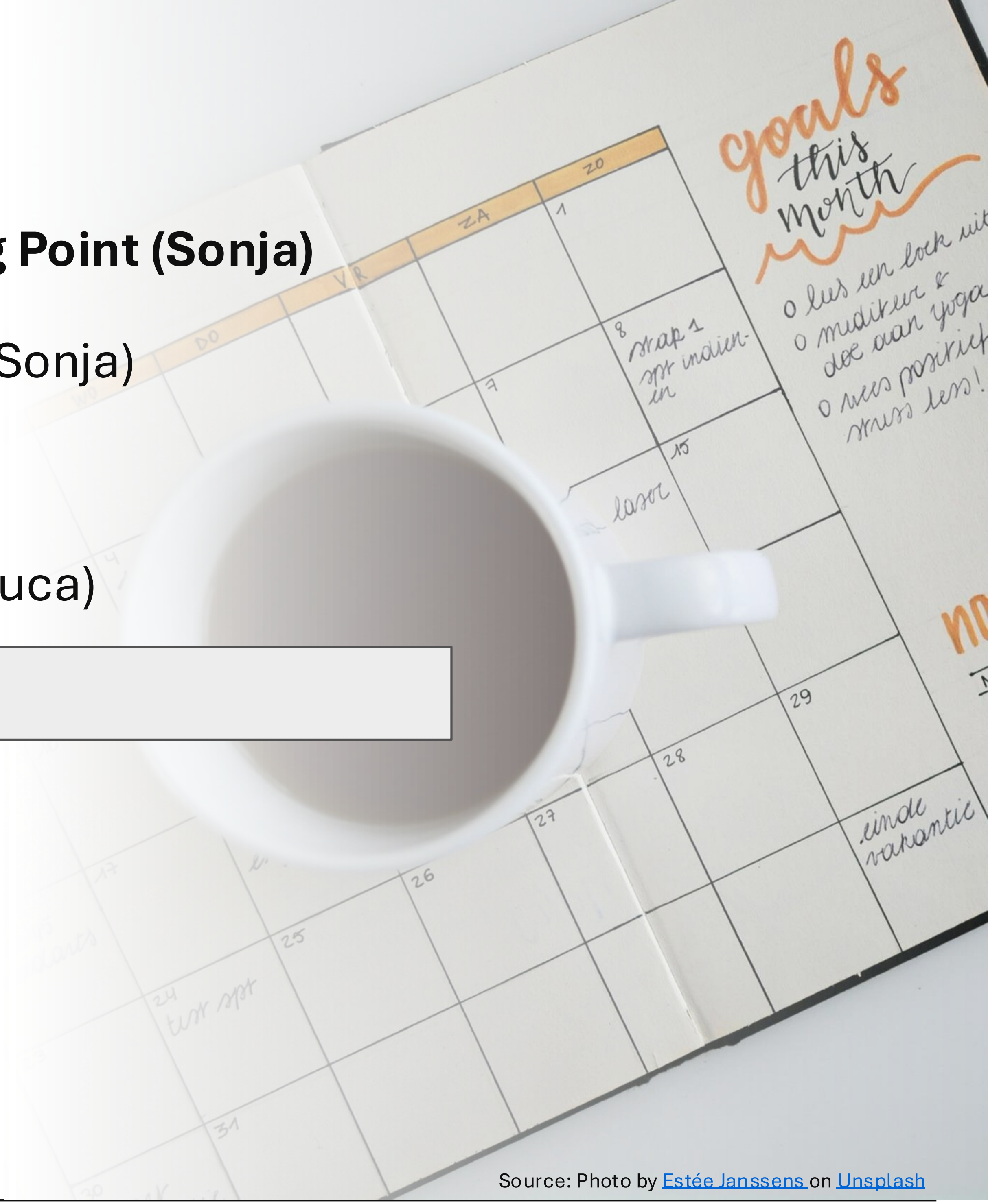
Luca Santosuosso[a], Bettina Klinz[b], Sonja Wogrin[a]

Distributed Stochastic Model Predictive Control with Temporal Aggregation for the Joint Dispatch of Cascaded Hydropower and Renewables

Luca Santosuosso, Sonja Wogrin

Agenda

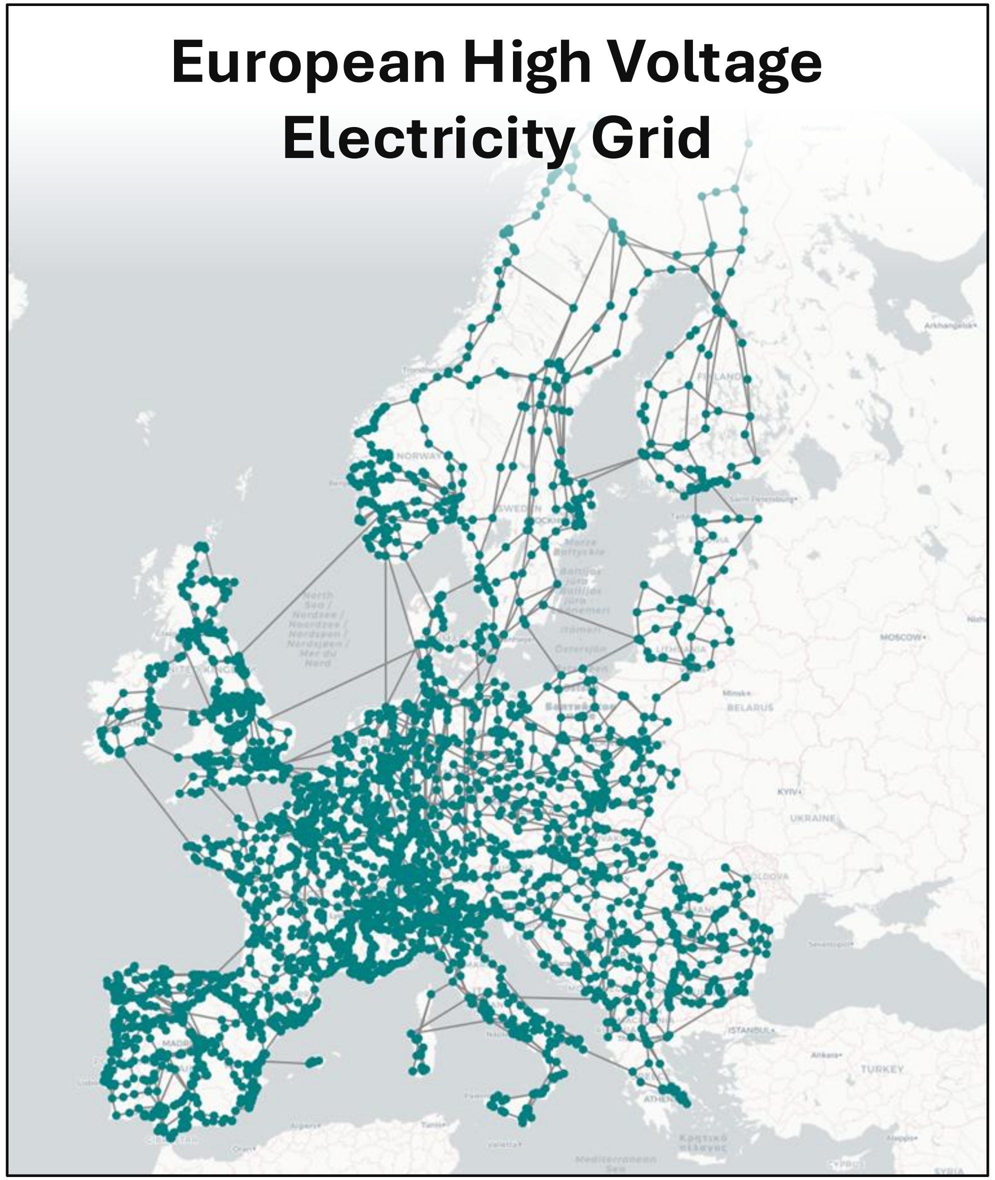
- **I** Time Series Aggregation. **Motivation** & **Starting Point (Sonja)**
- **II** Extension to Network & **Ramping** Constraints (Sonja)
- **III** Extension to **Storage** Constraints (Thomas)
- **IV** Time Series Aggregation with **Bounded Error** (Luca)
- **V** Extension to **Grid Aggregation** (Benjamin)
- **VI** Conclusions (Sonja)

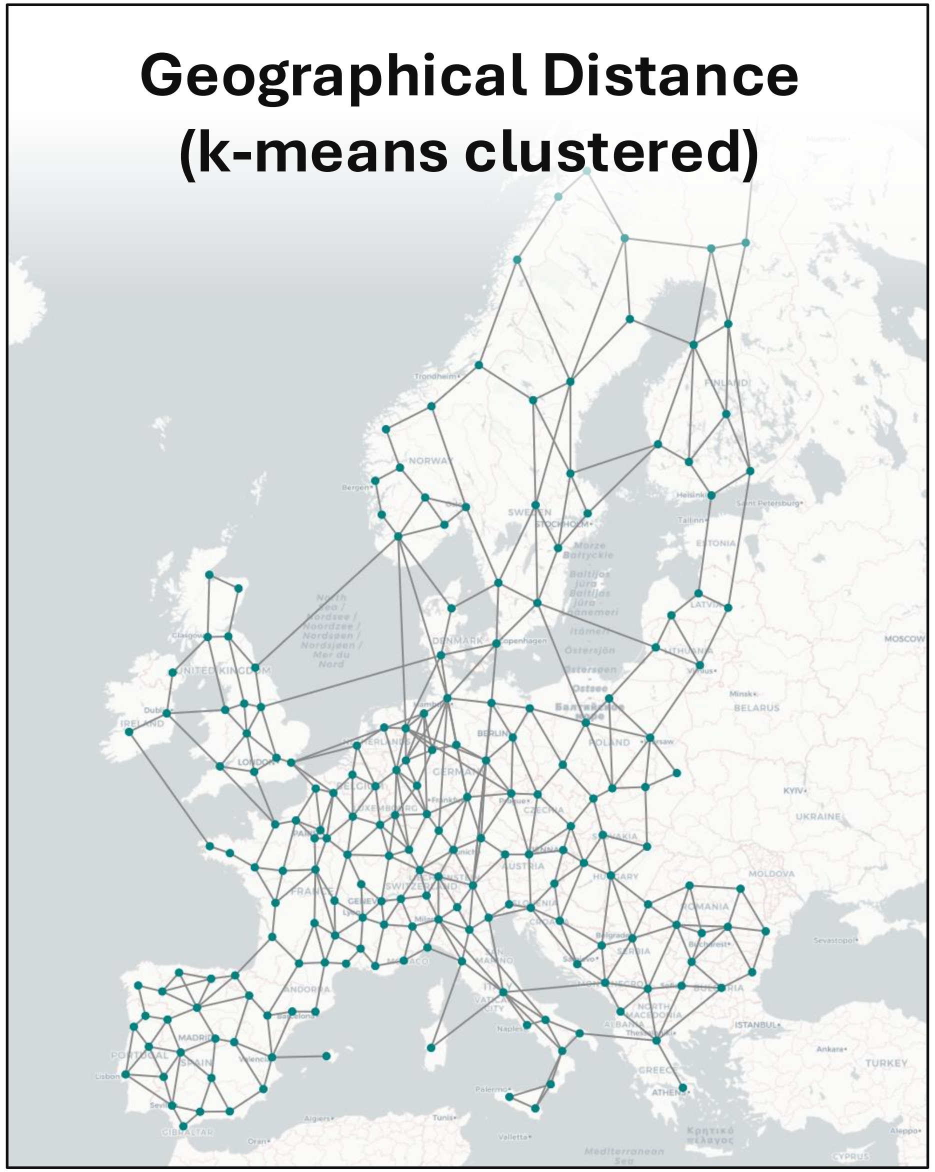


Source: Photo by Estée Janssens on Unsplash

Motivation

Grid Aggregation





Definition Grid Partitioning & Aggregation

Partitioning

- **General Attributes**
  Geographical Distance
- **Power Network**
  Electrical Distance
- **Duals Based**
  LMPs, NCP

Aggregation

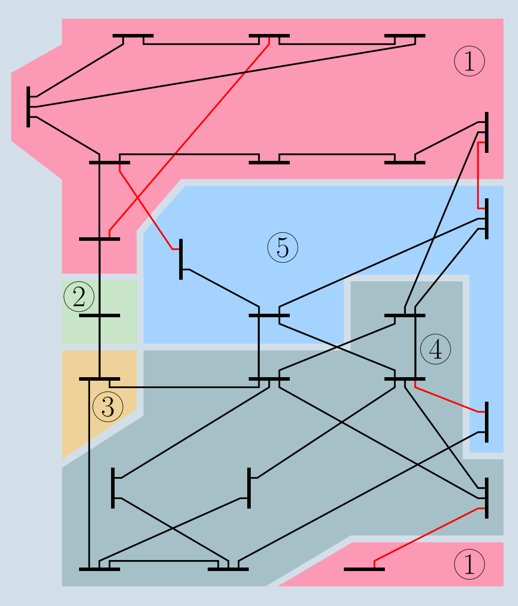
- **Transport Problem**
- **Parallel Reactances**
- **Reduced PTDF**
- **Kron-Reduction**

How to Partition Nodes?

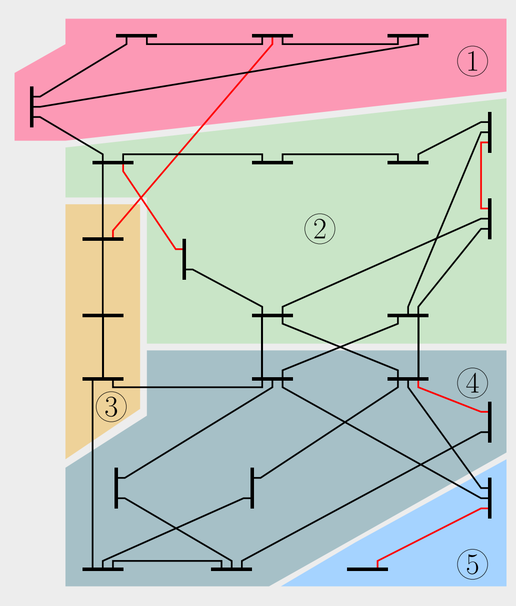
Duals based

Locational Marginal Price (LMP)

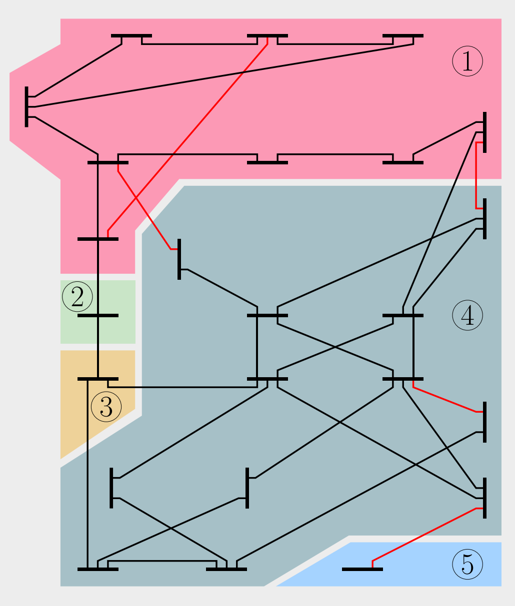
KMeans



Spectral Clustering

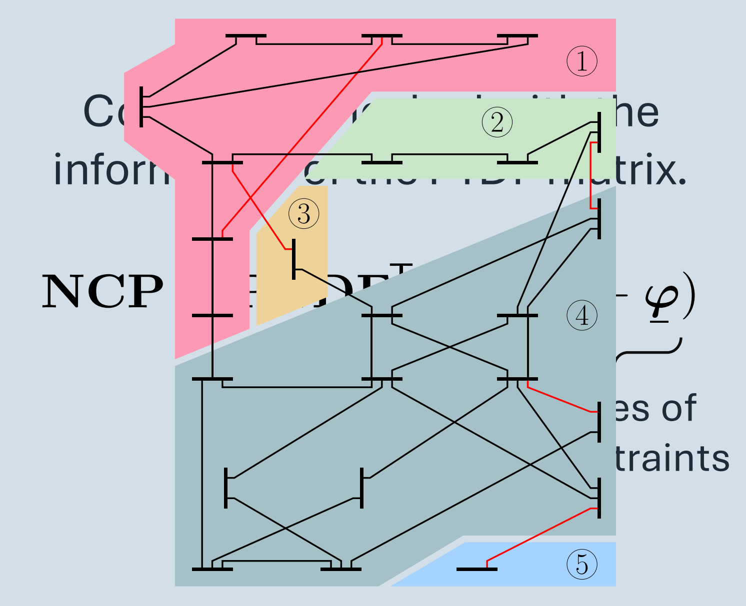


ANAC



Network Congestion Price (NCP)

KMeans

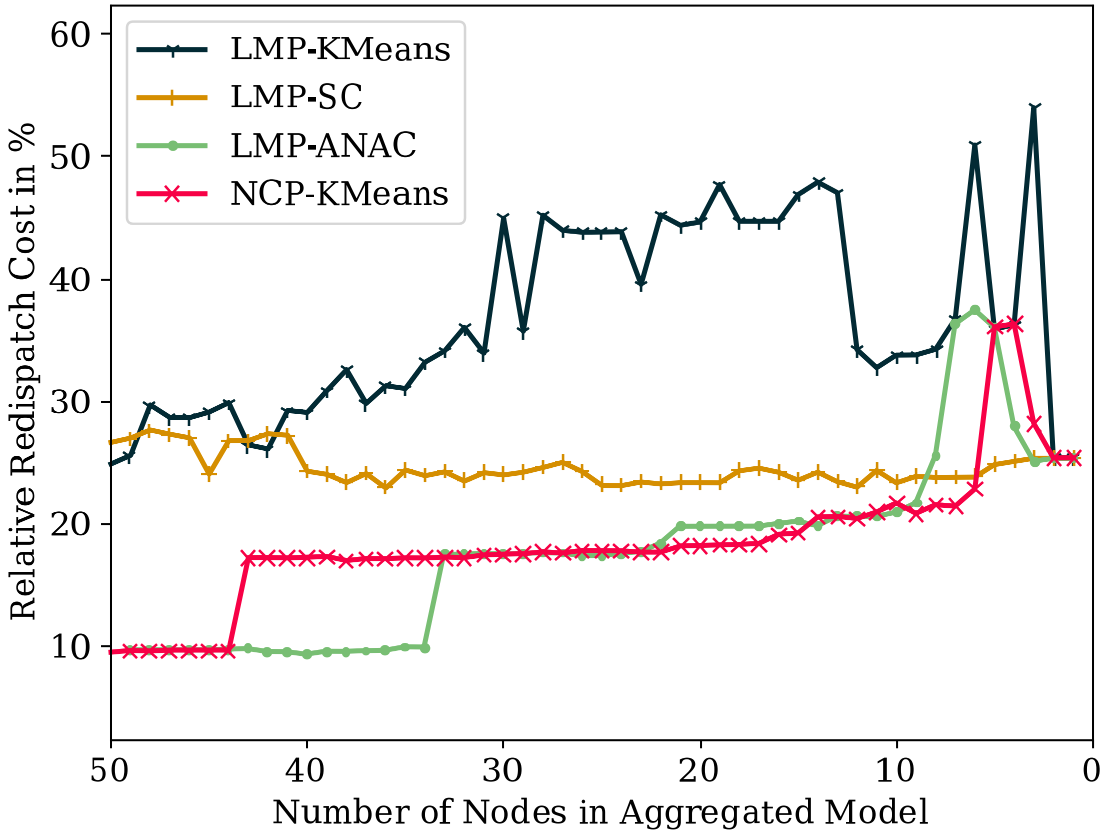


ANAC ... Adjacent Node Agglomerative Clustering

Source: **B. Stöckl** et al. "Congestion-Sensitive Grid Aggregation for DC Optimal Power Flow." 2025 IEEE Kiel PowerTech.

Relative Redispatch Cost

IEEE 300 Bus System



High deviation for **not direct connection** enforced clustering (LMP-KMeans)!

Spectral Clustering **fails** to aggregate the grid well.

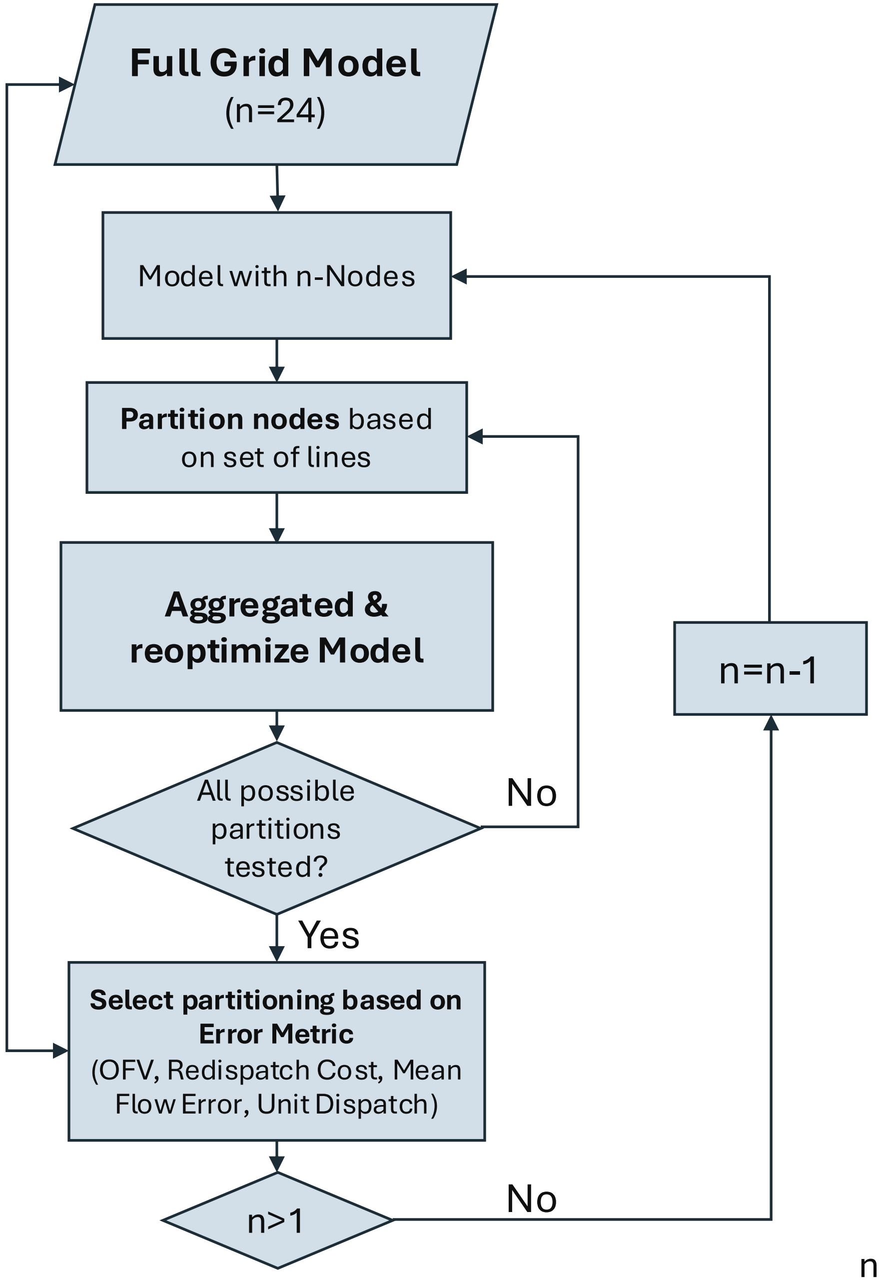
NCP-KMeans & ANAC Clustering achieve **good approximation** of OFV.

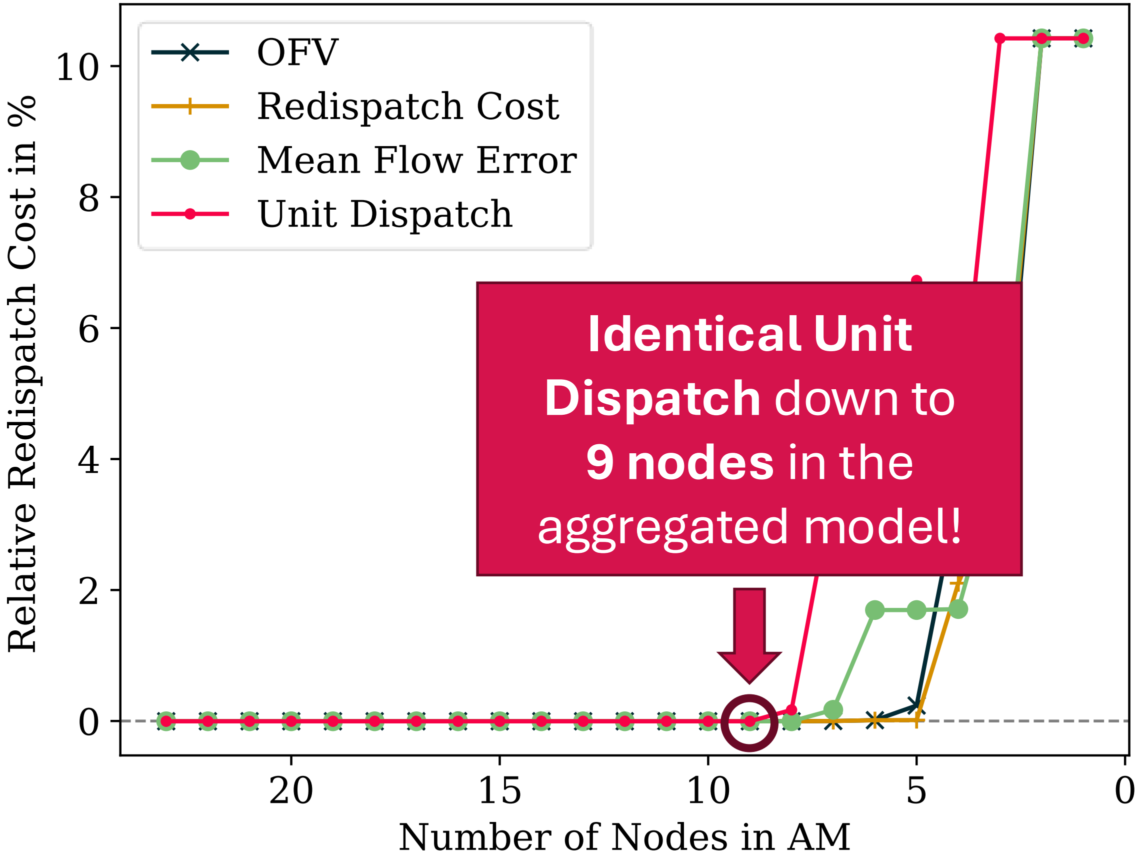
NCP-KMeans features **30 times faster clustering** compared to ANAC methods!

ANAC...Adjacent Node Agglomerative Clustering

Source: **B. Stöckl** et al. "Congestion-Sensitive Grid Aggregation for DC Optimal Power Flow." 2025 IEEE Kiel PowerTech.

Greedy Exhaustive Enumeration





n…number of nodes in AM

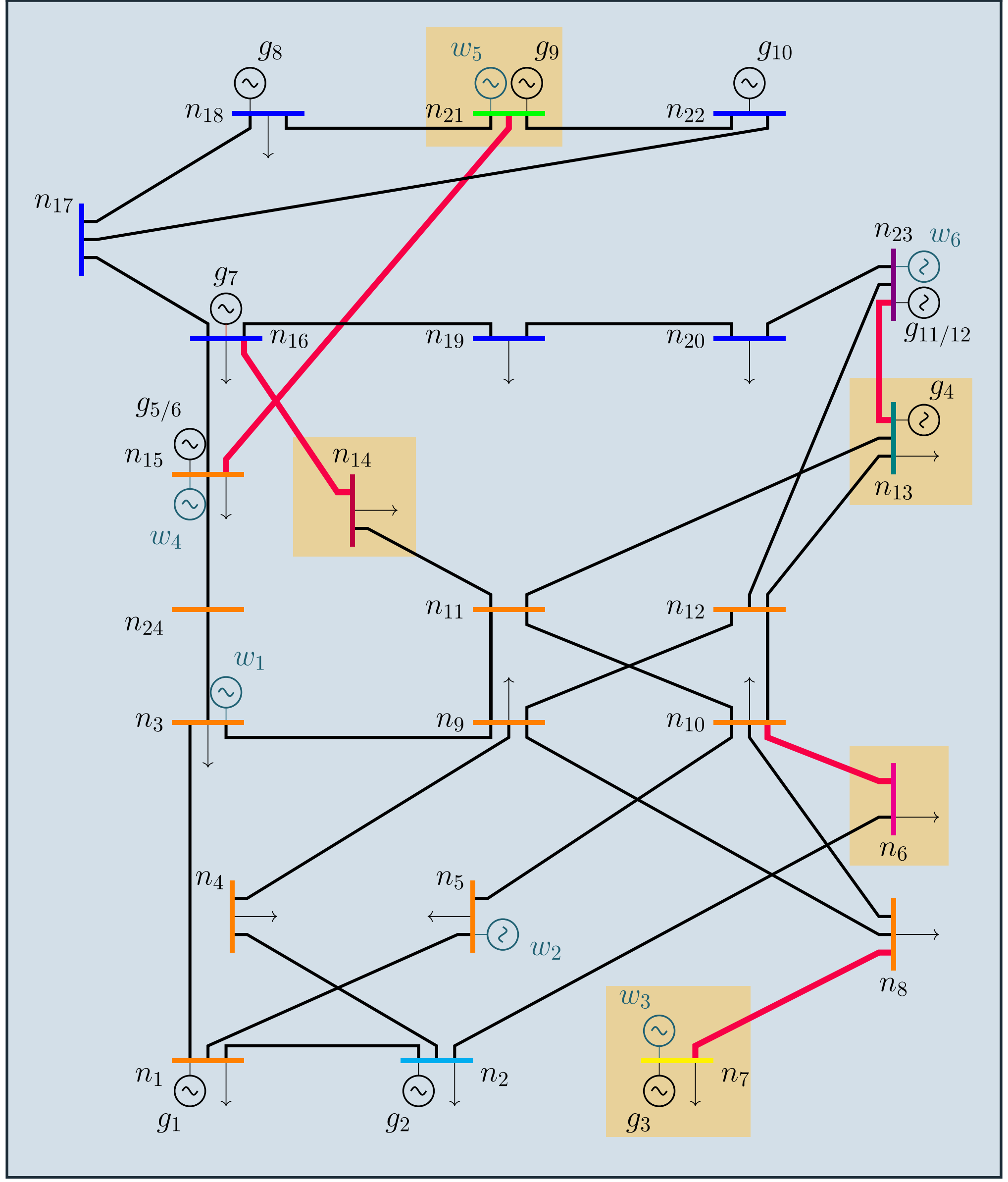
Optimal Partitioning

Typical partitioning methods **fail** to aggregate networks **without deviations** in unit dispatch.

Generators which are **restricted by the grid** must be taken into account.

Partitioning method which better accounts for the **restricted generators**!

IEEE 24-Bus RTS



Key Messages

Grid Aggregation

Clustering-Based Partitioning

- Considering the **network topology** and the **connection of nodes** is crucial for **node clustering**!
- Partitioning solely based on line-congestion information is **not sufficient**!

Future research

- Partitioning method which better accounts for the **restricted generators**!

🚀 Check out our paper

Congestion-Sensitive Grid Aggregation for DC Optimal Power Flow

Benjamin Stöckl, Yannick Werner, and Sonja Wogrin
Institute of Electricity Economics and Energy Innovation, Graz University of Technology, Graz, Austria
{benjamin.stoeckl,yannick.werner,wogrin}@tugraz.at

***Abstract*—The vast spatial dimension of modern interconnected electricity grids challenges the tractability of the DC optimal power flow problem. Grid aggregation methods try to overcome this challenge by reducing the number of network elements. Many existing methods use Locational Marginal Prices as a distance metric to cluster nodes. In this paper, we show that prevalent methods adopting this distance metric fail to adequately capture the impact of individual lines when there is more than one line congested. This leads to suboptimal outcomes for the optimization of the aggregated model. To overcome those issues, we propose two methods based on the novel Network Congestion Price metric, which preserves the impact of nodal power injections on individual line congestions. The proposed methods are compared to several existing aggregation methods based on Locational Marginal Prices. We demonstrate all methods on adapted versions of the IEEE RTS 24- and 300-Bus systems. We show that the proposed methods outperform existing approaches both in terms of objective function value error and maximum line limit violation, while exhibiting faster node clustering. We conclude that aggregation methods based on the novel Network Congestion Price metric are better at preserving the essential physical characteristics of the network topology in the grid aggregation process than methods based on Locational Marginal Prices.**

challenge [2]. These methods try to conserve the essential physical properties of a given network topology with fewer network elements [4].

The aggregation process can be classified into two groups: spatial aggregation (e.g., nodes and lines) and technological aggregation (e.g., generators and demands). Spatial aggregation can be further divided into identifying groups of nodes, which we refer to as *grid partitioning*, and determining parameters of the lines in the aggregated grid, which we refer to as *grid aggregation*. In this paper, we neglect technological aggregation and focus on grid partitioning only, enabling an unbiased and rigorous comparison of different methods. We refer the interested reader to [2] for further information on technological aggregation and provide a brief summary of existing grid aggregation methods in Appendix A.

In recent years, several methods combining various distance metrics with clustering algorithms to aggregate electrical grids have been proposed in the literature [2], [5]. In the context of DC-OPF, prevalent distance metrics based solely on the network topology include the geographical [6] or electrical

Source: **B. Stöckl** et al. "Congestion-Sensitive Grid Aggregation for DC Optimal Power Flow." 2025 IEEE Kiel PowerTech.

Conclusions

- Energy system models are **complex**, but data aggregation can help.
- Many aggregation methods (e.g., TSA) are purely input-based and leave a lot of aggregation potential untapped.
- **Our research:**
  - Leverages this potential to get **more bang for your buck**,
  - shows that **<u>exact</u> TSA is possible with network, ramping, and storage constraints**
  - uses **ML** to leverage aggregation potential in practice, and
  - provides **performance guarantees**

Thank you for the attention!

Institute of Electricity Economics and Energy Innovation

Research Center energetic

TU Graz

Sonja Wogrin

Univ.-Prof. Dipl.-Ing. Dr. M.Sc.

Graz University of Technology

Institute of Electricity Economics and Energy Innovation
Inffeldgasse 18
8010 Graz, Austria

Mail: wogrin@tugraz.at
Web: iee.tugraz.at

Check out our website!

NETZERO-OPT

instagram.com/iee.tugraz

linkedin.com/company/iee-tugraz

Funded by the European Union

European Research Council
Established by the European Commission

Funded by the European Union (ERC, NetZero-Opt, 101116212). *Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.*