Cracking the Complexity Barrier: Towards Exact Data Aggregation for High-Performance Energy System Models (Part 1)

DTU PES Summer School 2026

Sonja Wogrin
Institute of Electricity Economics and Energy Innovation
Graz University of Technology

20.05.2026

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

My Team

Head of Institute

Univ.-Prof. DI Dr Sonja Wogrin MSc

Institut für Elektrizitätswirtschaft und Energieinnovation (IEE)
Technische Universität Graz

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

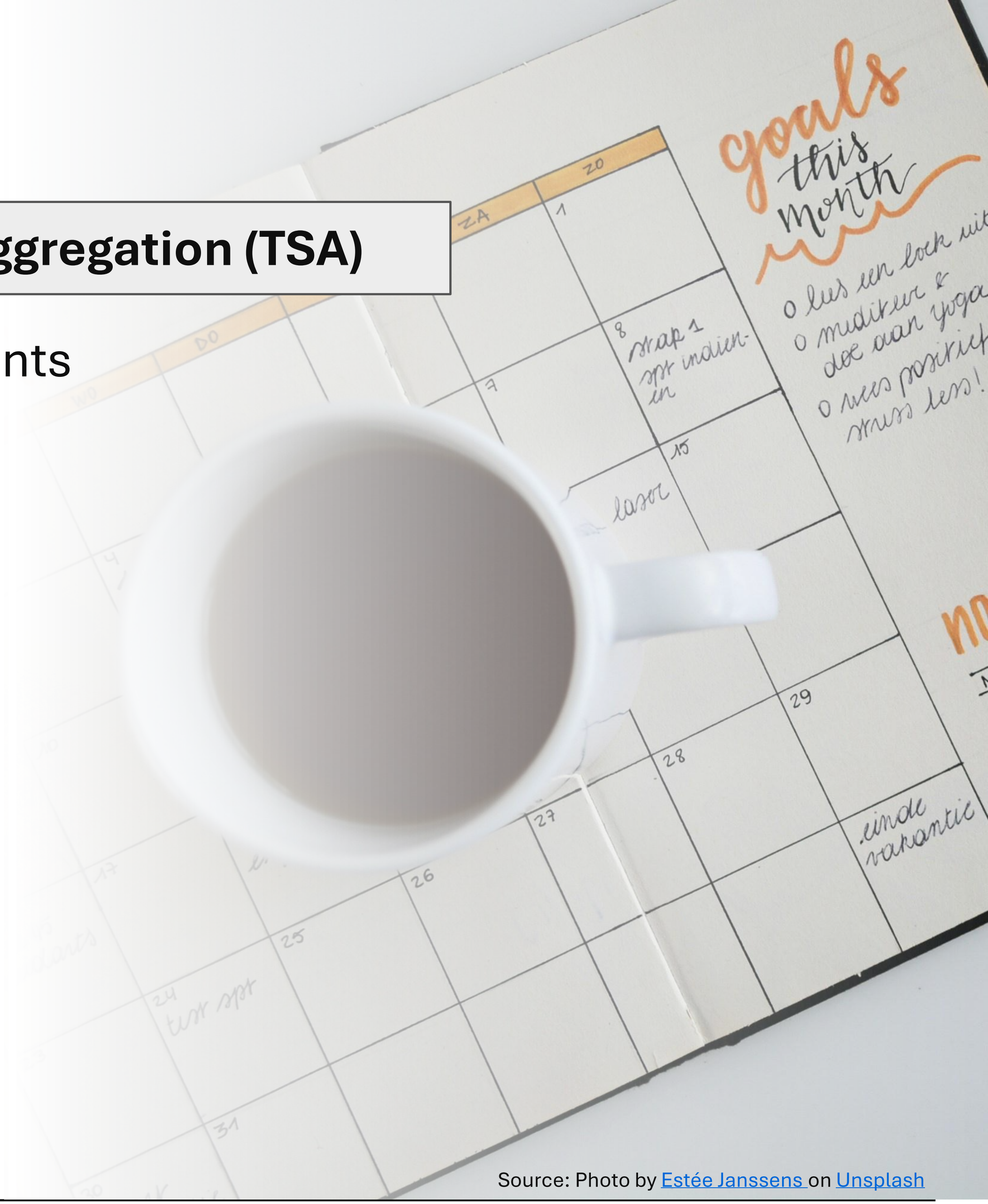
Spokesperson

Research Center energetic

E-Mail: energetic@tugraz.at
Web: energetic.tugraz.at

Agenda

- **I** Data Aggregation. **Motivation & Time Series Aggregation (TSA)**
- **II** TSA Extension to **Network** & **Ramping** Constraints
- **III** TSA Extension to **Storage** Constraints
- **IV** TSA with **Bounded Error**
- **V** Data Aggregation. **Spatial Aggregation**
- **VI** Conclusions



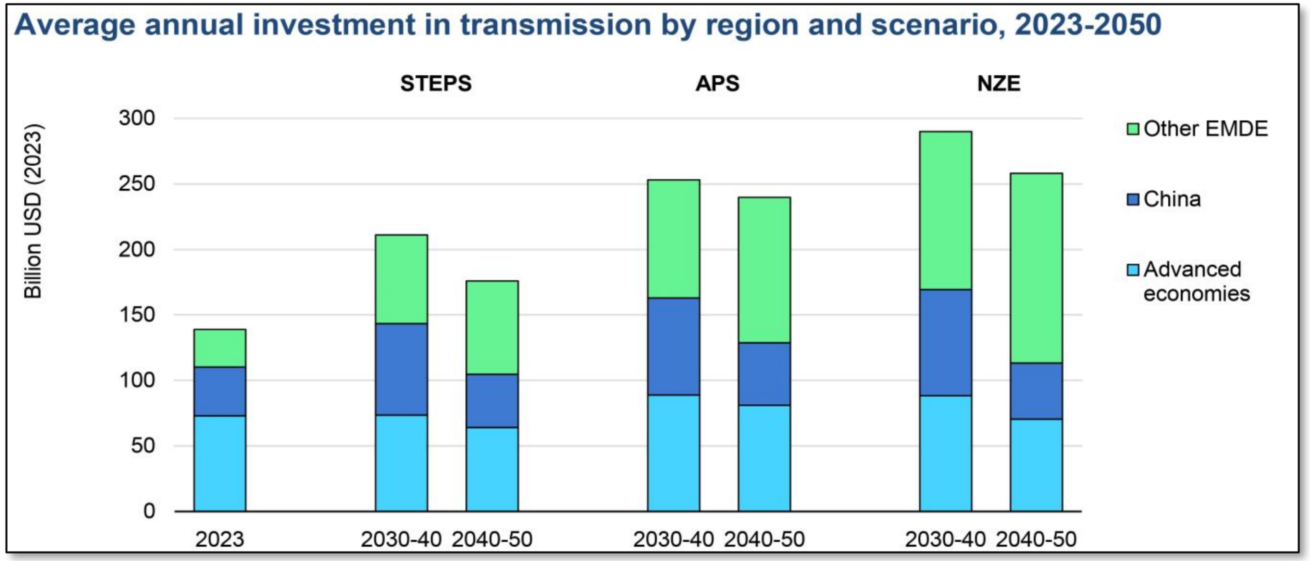
Source: Photo by Estée Janssens on Unsplash

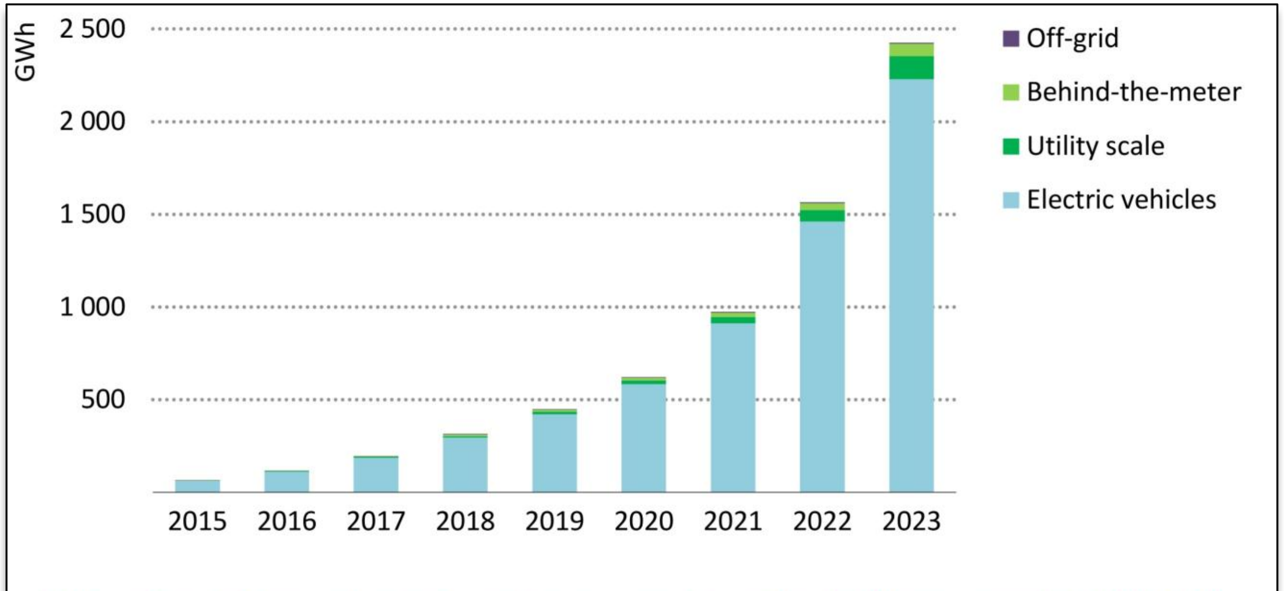
ARE WE MODELING THE INTRACTABLE?

The energy grid is the largest and most intricate machine ever built.

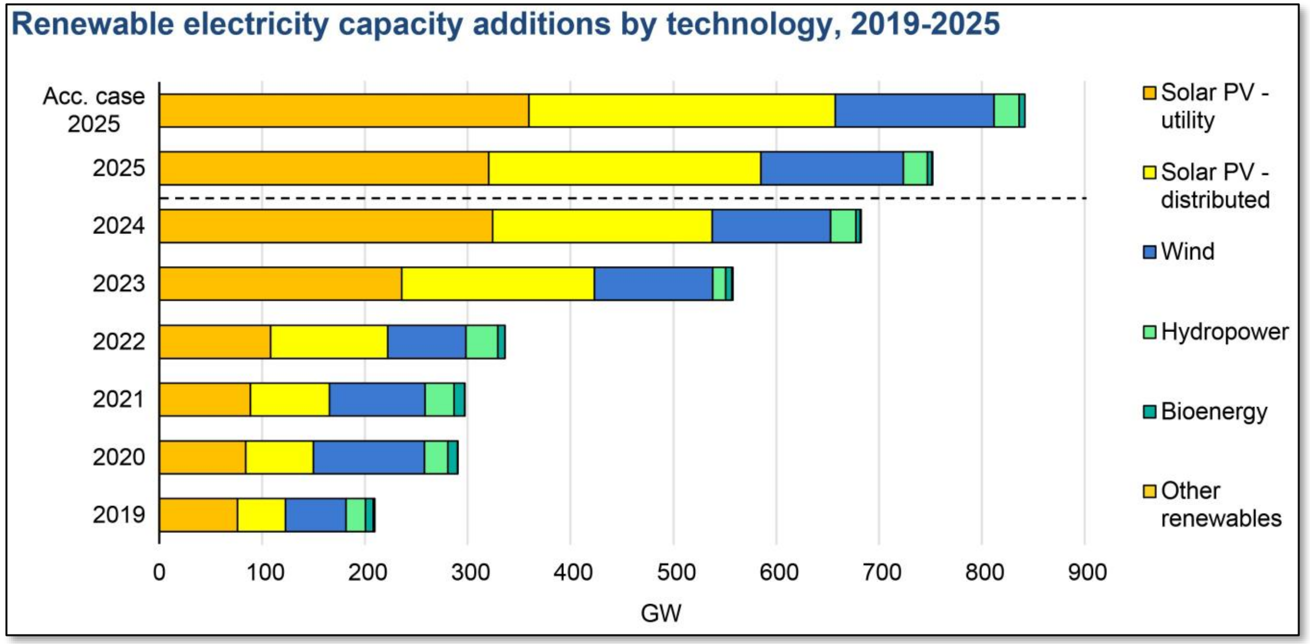
The price of accuracy is complexity!

Sources: IEA Reports, 2023–2025 (Building the Future Transmission Grid, Renewables 2025, Batteries and Secure Energy Transitions).



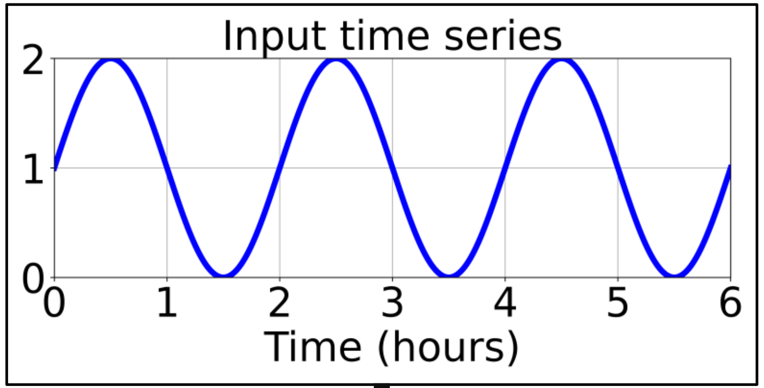


Lithium-ion battery volumes in use have surged over the last three years to 2 400 GWh



GENERATION EXPANSION PLANNING

Determine the optimal mix and sizing of generation units to minimize investment and operational costs while meeting the future energy demand.



FULL-SCALE OPTIMIZATION MODEL

$$
\begin{aligned}
\min_{\boldsymbol{z}} \quad & J(\boldsymbol{z}) \\
\text{s.t.} \quad & \sum_{g \in \boldsymbol{G}} p_{g,t}\Delta + d_t^{\mathrm{ns}} = D_t, \quad \forall t, \\
& 0 \le p_{g,t} \le F_{g,t}\, x_g, \quad \forall g, \forall t, \\
& b_g\, \underline{X}_g \le x_g \le b_g\, \overline{X}_g, \quad \forall g.
\end{aligned}
$$

Operational + investment costs

Operational and **investment** constraints

Mixed-integer formulations of the generation expansion planning problem are **NP-hard***!

***Source**: S. Goderbauer, M. Comis, and F. J. Willamowski, "The synthesis problem of decentralized energy systems is strongly NP-hard," *Computers & Chemical Engineering*, vol. 124, pp. 343-349, 2019.

Motivation. Dimensions of complexity of energy system models

.... and how to tackle those challenges

Uncertainty
E.g., scenario reduction.

Yannick

Technical
E.g., decompositions, better formulations.

Thomas

Spatial
E.g., node partitions and aggregations.

Benjamin

Marco

NPAP

Temporal
E.g., time series aggregation.

Sonja

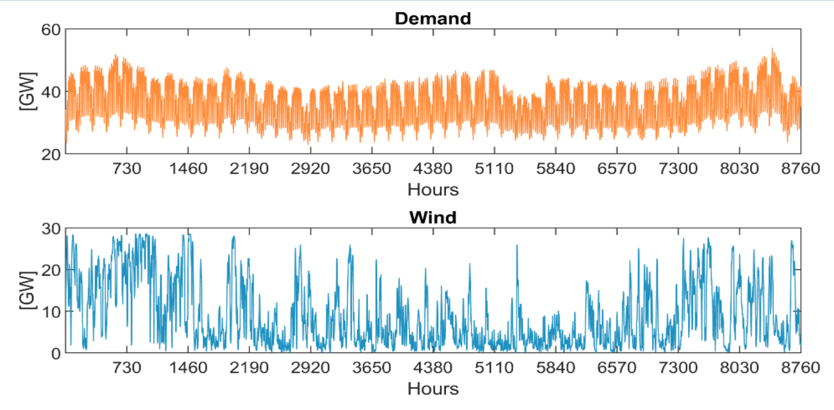
Jakub

Luca

European Research Council
Established by the European Commission

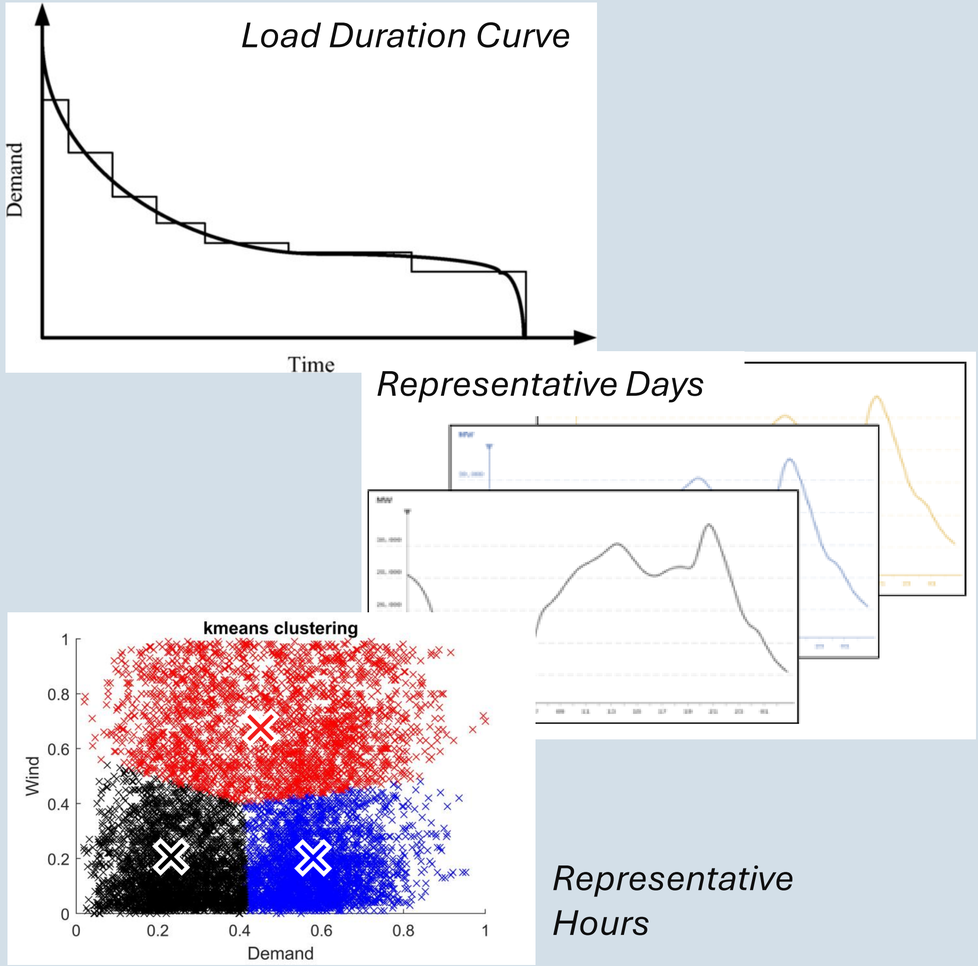
Time Series Aggregation (TSA)

Original Data Time Series



TSA

Times Series Aggregation



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

Sources:
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).

Computational Challenge

EU Power System Model

Variables: 660 M

NetZero-Opt

 # Variables: 0.5 M

Funded by the European Union

European Research Council
Established by the European Commission

3 orders of magnitude

Frontier Supercomputer*
600 M$

Typical Workstation**
5000 €

	Hourly Resolution	NetZero-Opt
Frontier 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€.

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

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

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

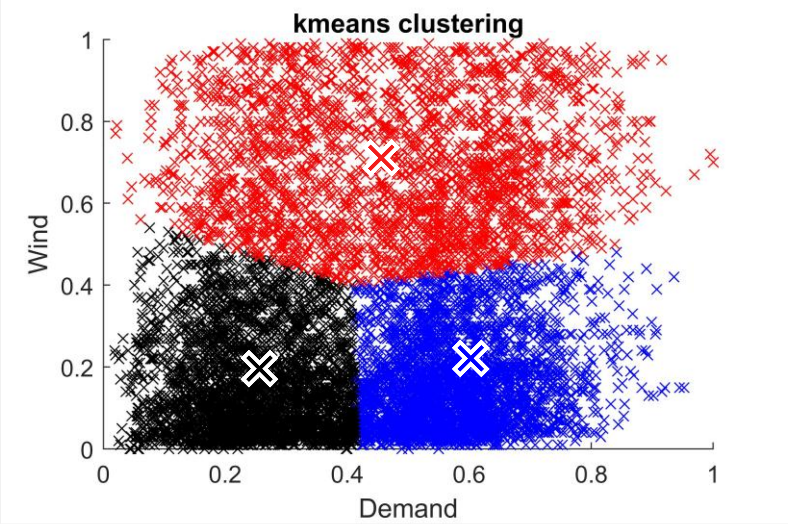
$$\underline{P}_g \leq p_{g,r} \leq \overline{P}_{g,r} \quad \forall g, r$$

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 k-means clustering:

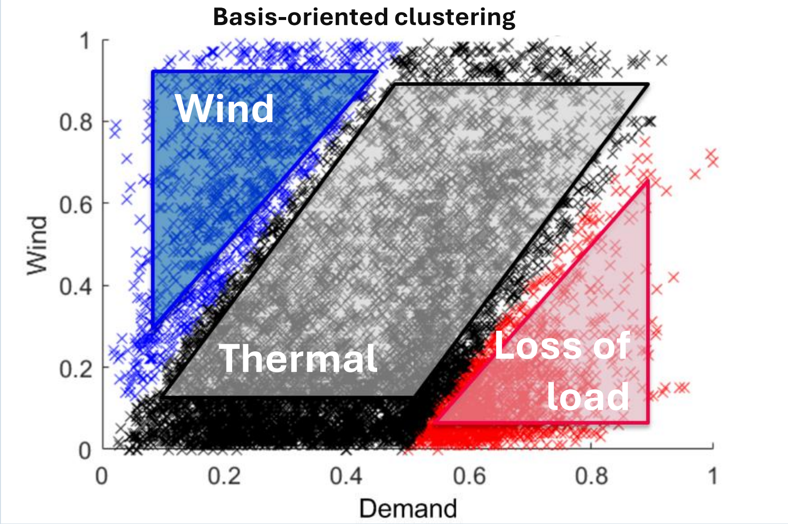


Results of aggregated economic dispatch problem:

- 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).

Open Questions ...

... we seek to answer in this lecture

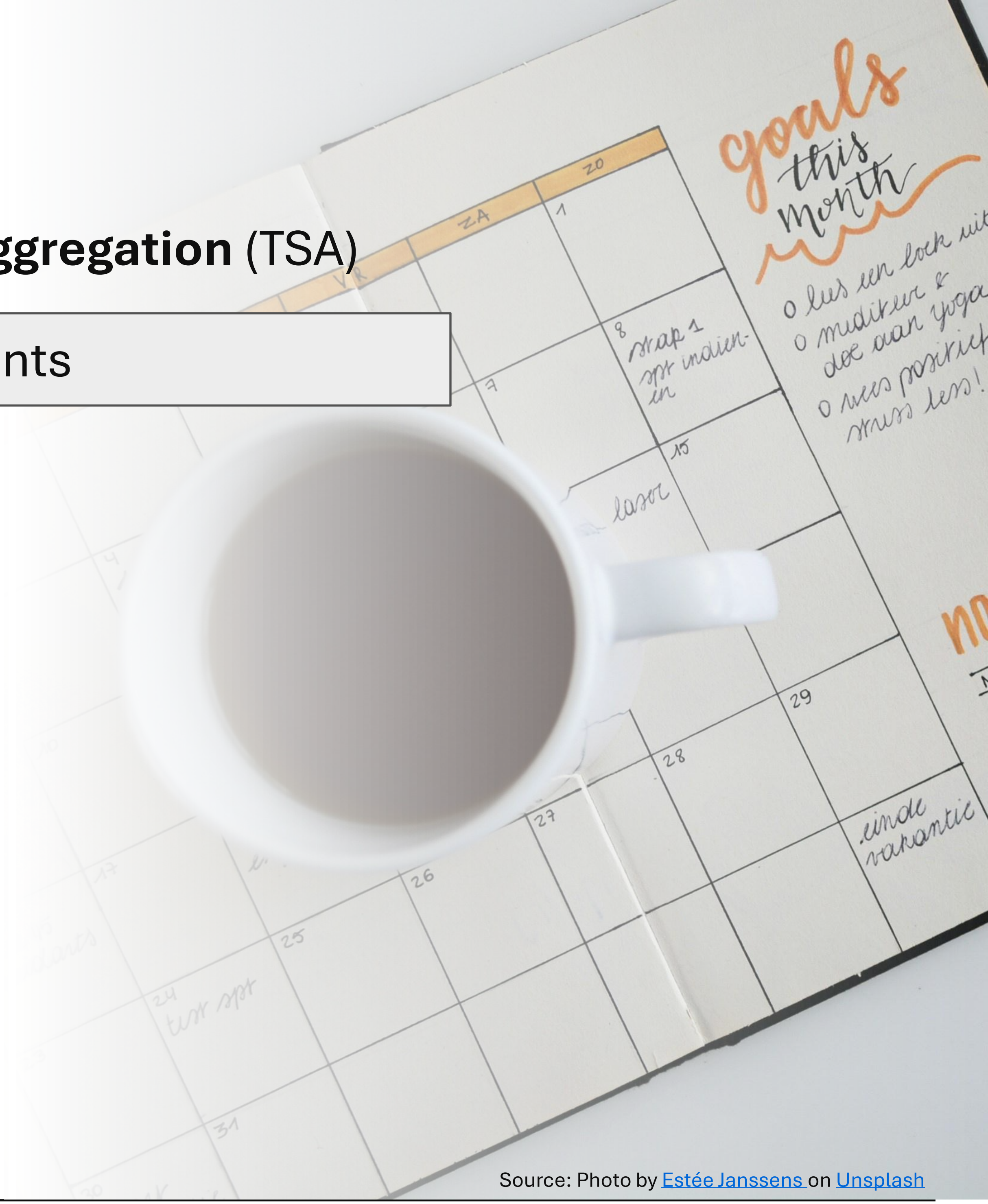
Can this methodology be extended to more complex power system models, e.g., including **network** and **time-linking constraints**?

Is **exact TSA** possible with time-linking constraints (i.e., **ramping or storage**) when model structure is accounted for?

Can we establish valid **performance bounds** for TSA aggregated models?

Agenda

- **I** Data Aggregation. **Motivation** & **Time Series Aggregation** (TSA)
- **II** TSA Extension to **Network** & **Ramping** Constraints
- **III** TSA Extension to **Storage** Constraints
- **IV** TSA with **Bounded Error**
- **V** Data Aggregation. **Spatial Aggregation**
- **VI** Conclusions



Source: Photo by Estée Janssens on Unsplash

The methodology extends* to transport problems

Formulation

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

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

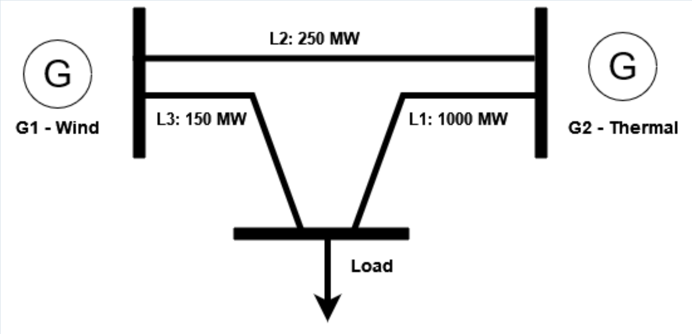
$$\underline{P_t} \leq p_{t,k} \leq \overline{P_t} \quad \forall k, t$$

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

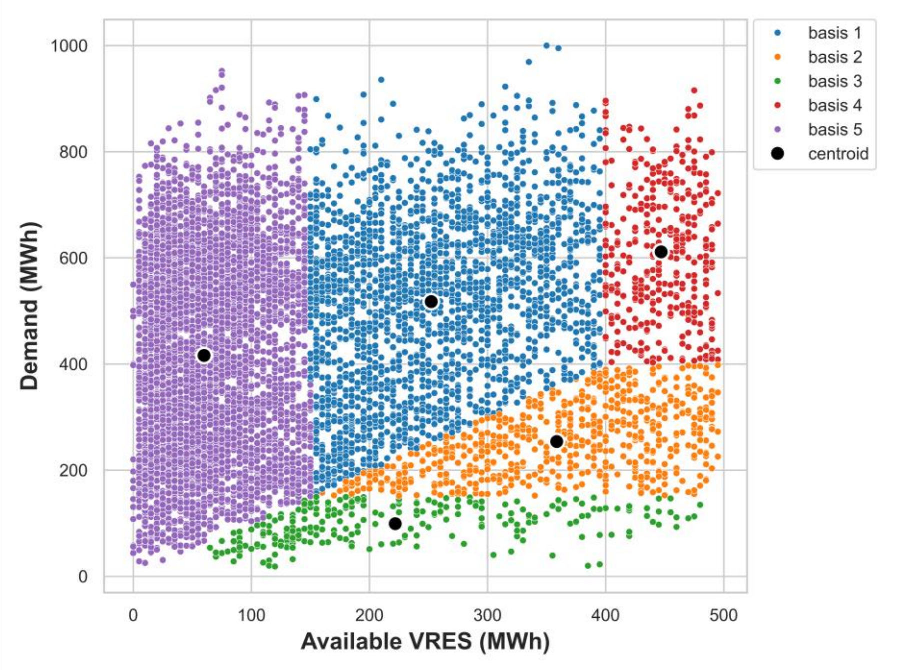
$$f_{k,i,j} \leq \overline{F_l} \quad \forall k, i, j$$

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

Add network constraints



Aggregation Results



- More active constraint sets because of **line constraints**
- Only **five active constraint sets** of one-hour length (centroids) 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

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

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

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

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

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

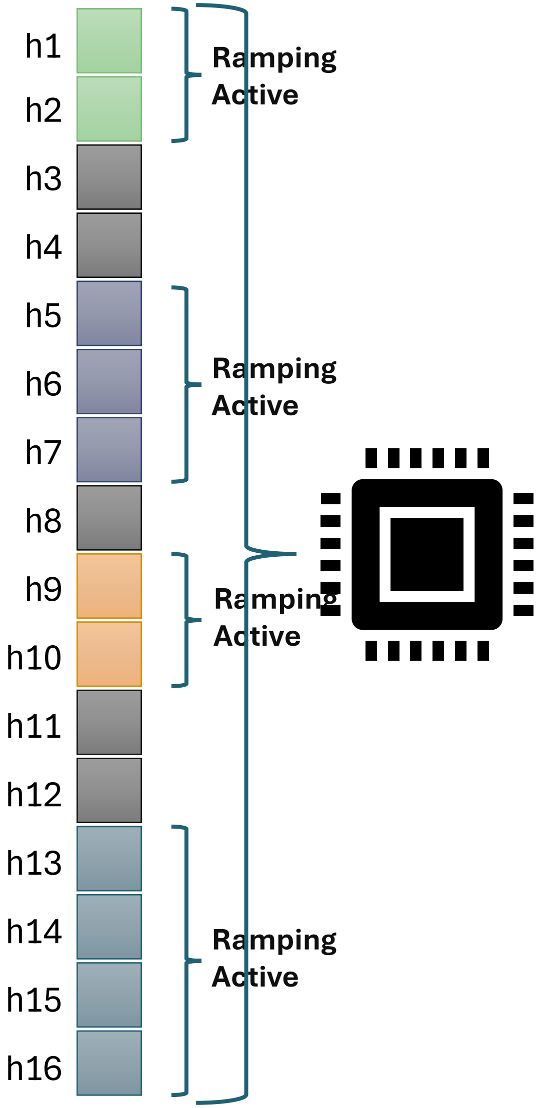
$$f_{k,j,i} \le \overline{F_l} \quad \forall k, i, j$$

Model Structure

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

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

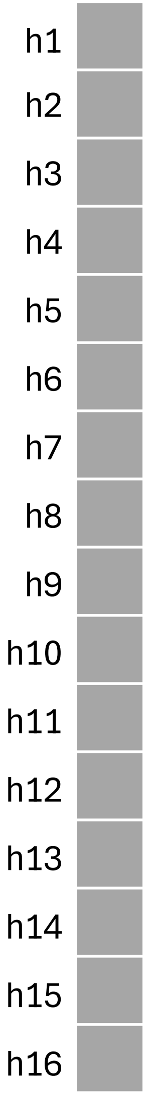
Identifying “active constraint sets” which are time-linked



1. Identify **active ramping constraints**

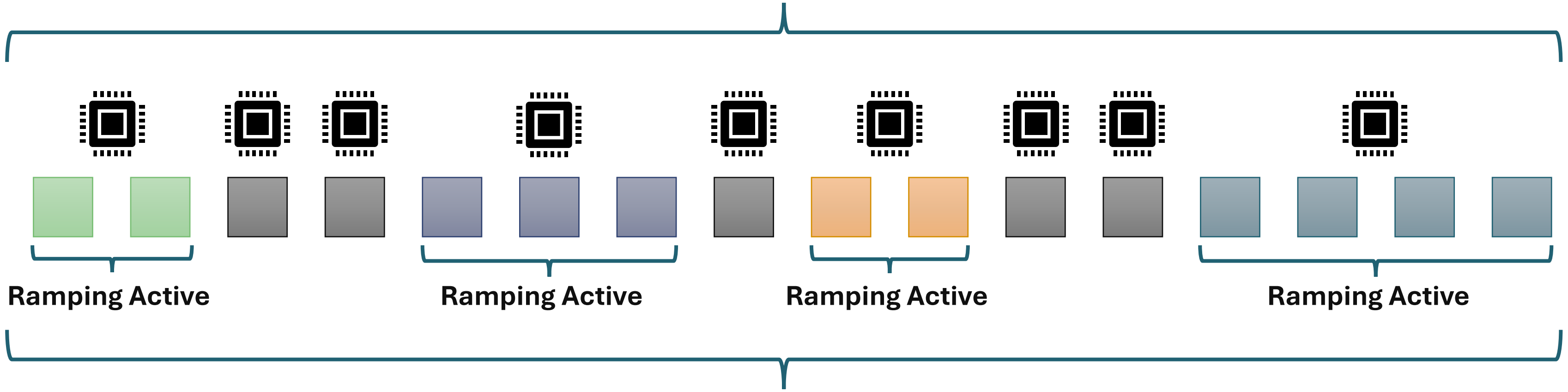
Source*: **D. Cardona-Vasquez** et al. "Disaggregation of energy system optimization models using machine learning for identification of active constraints." Sustainable Energy, Grids and Networks (2025).

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

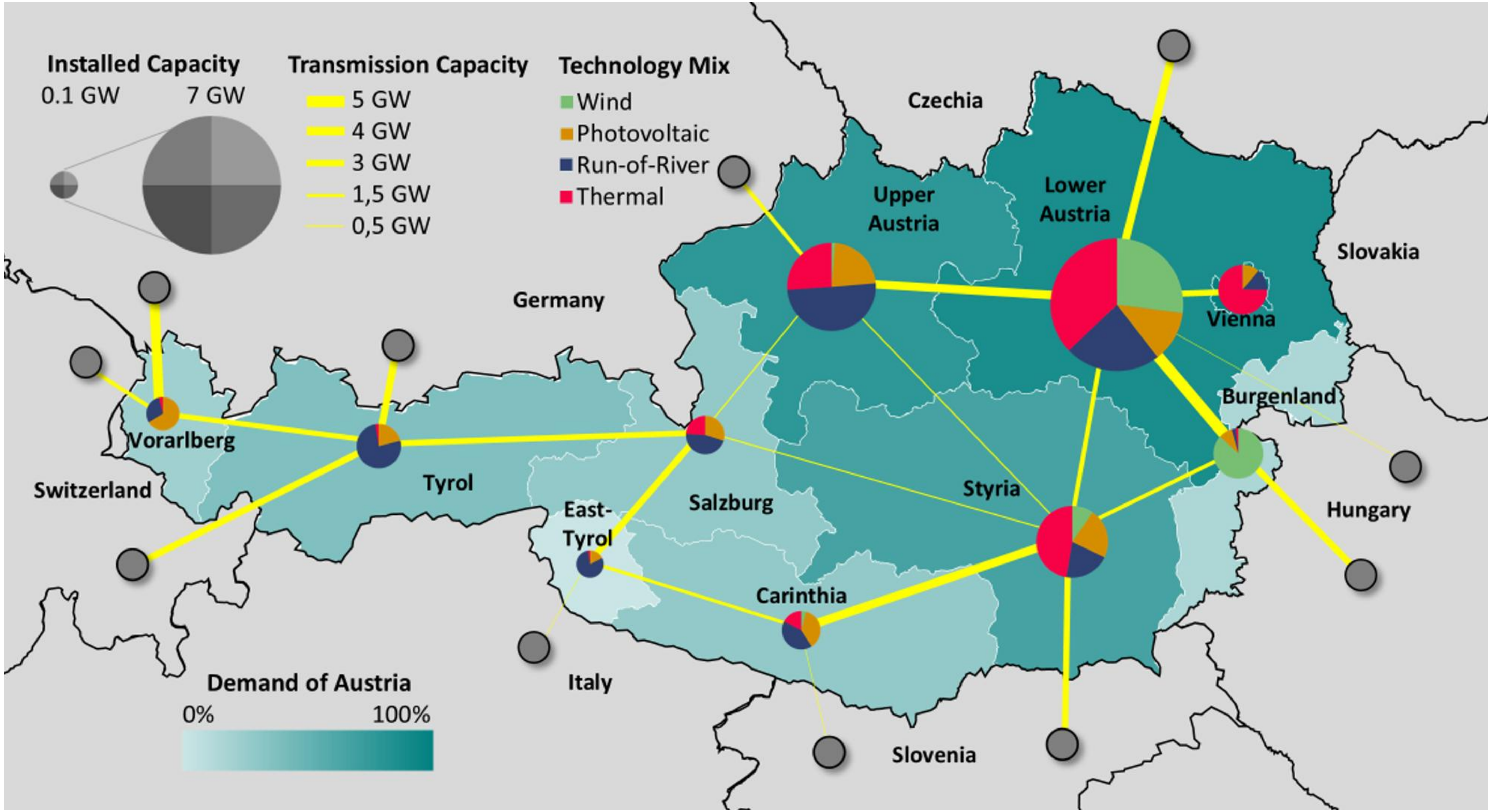
Disaggregate full model into submodels and solve in parallel



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 (2025).

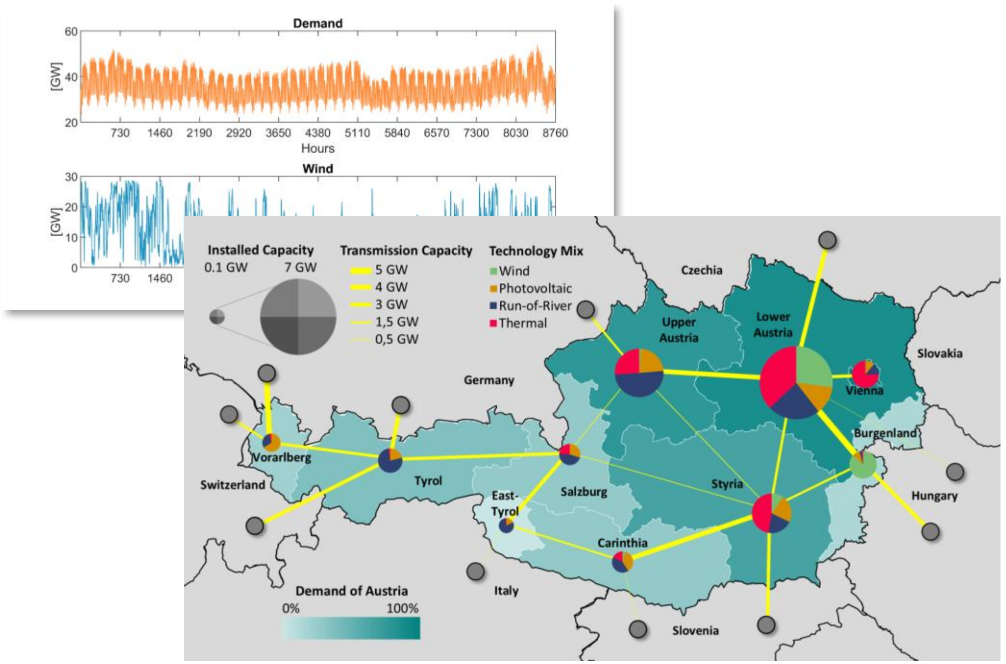
Case Study: Austrian Power System

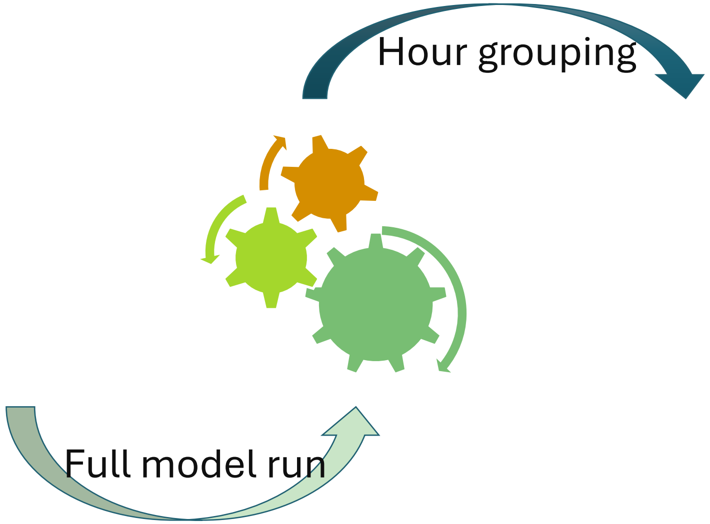


Source*: **D. Cardona-Vasquez** et al. "Disaggregation of energy system optimization models using machine learning for identification of active constraints." Sustainable Energy, Grids and Networks (2025).

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 (2025).

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%

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

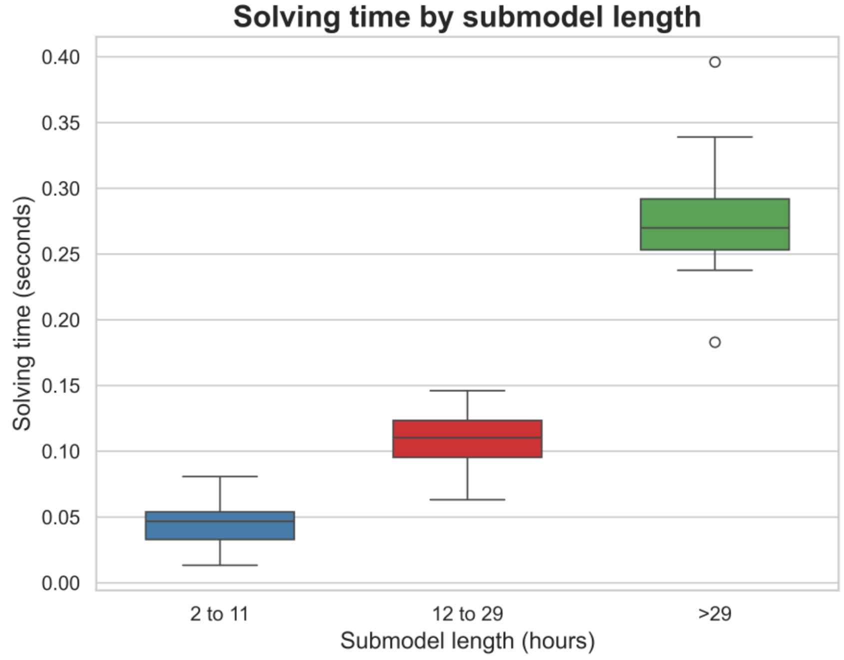
Source*: **D. Cardona-Vasquez** et al. "Disaggregation of energy system optimization models using machine learning for identification of active constraints." Sustainable Energy, Grids and Networks (2025).

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 (2025).

Key Messages: Network & Ramping Constraints

Exact TSA based on active constraint sets extends to include 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.

Agenda

I Data Aggregation. **Motivation** & **Time Series Aggregation** (TSA)

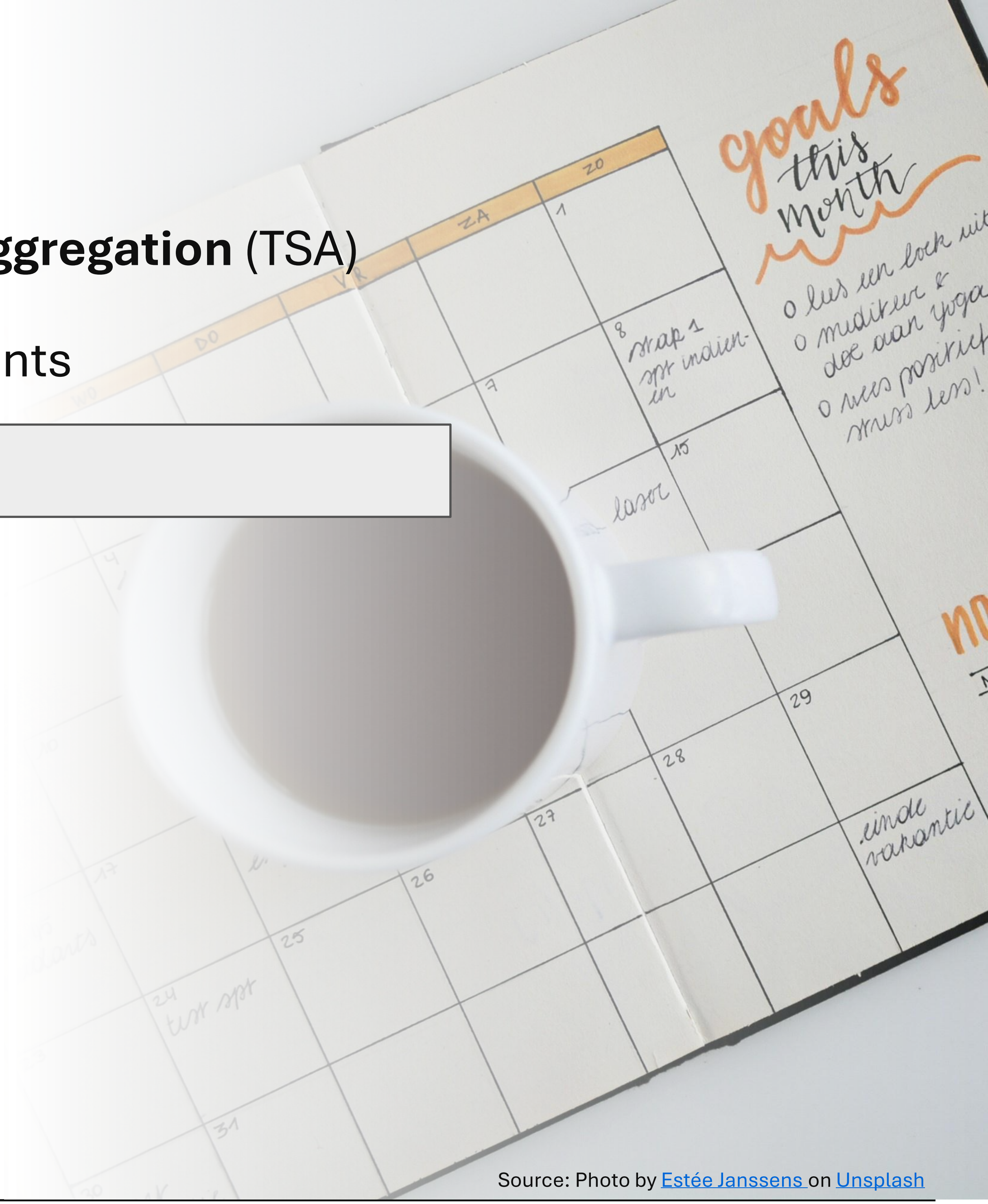
II TSA Extension to **Network** & **Ramping** Constraints

III TSA Extension to **Storage** Constraints

IV TSA with **Bounded Error**

V Data Aggregation. **Spatial Aggregation**

VI Conclusions



Source: Photo by Estée Janssens on Unsplash

Ramping versus Storage Time-Coupling Constraints

Ramping

- **Inequality time-coupling** constraints
- **Active** or **inactive**, i.e., ramping is only active during a **limited number of periods**

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

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

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

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

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

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

Storage constraints

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

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

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

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

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

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

Model Structure

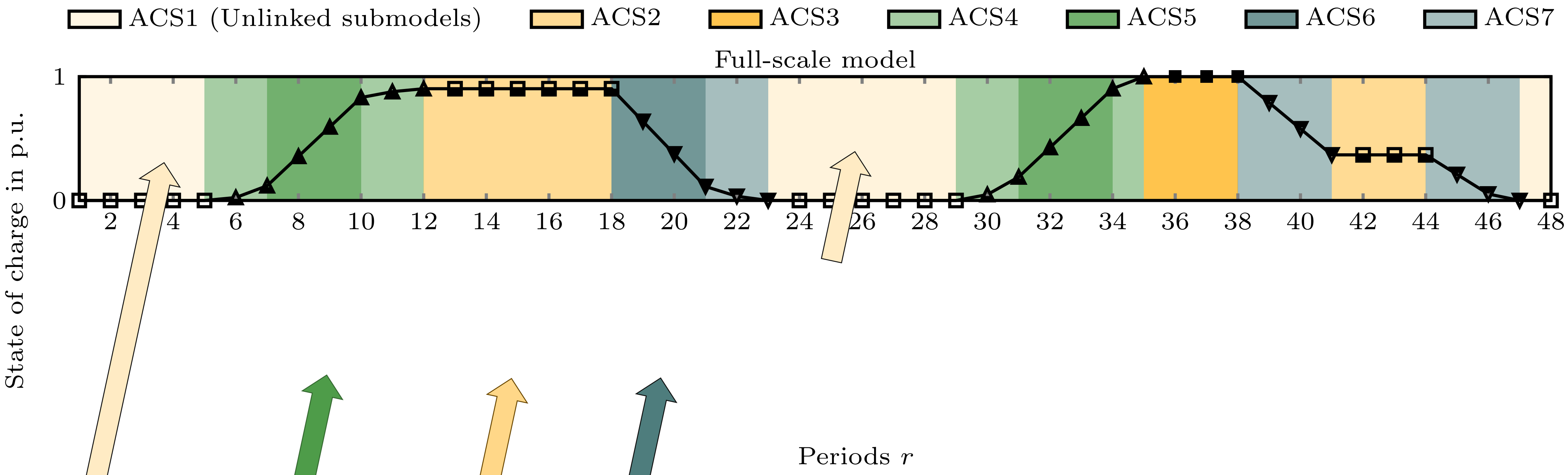
$$\begin{pmatrix} \underbrace{\begin{matrix} 1 & -W_{r-1}\eta^{\mathrm{c}}_s & W_{r-1}/\eta^{\mathrm{d}}_s \\ 0 & 0 & 0 \end{matrix}}_{r-1} & \Bigg| & \underbrace{\begin{matrix} -1 & 0 & 0 \\ 1 & -W_r\eta^{\mathrm{c}}_s & W_r/\eta^{\mathrm{d}}_s \end{matrix}}_{r} \end{pmatrix} \begin{pmatrix} e_{r-1,s} \\ p^{\mathrm{c}}_{r-1,s} \\ p^{\mathrm{d}}_{r-1,s} \\ e_{r,s} \\ p^{\mathrm{c}}_{r,s} \\ p^{\mathrm{d}}_{r,s} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

- **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)

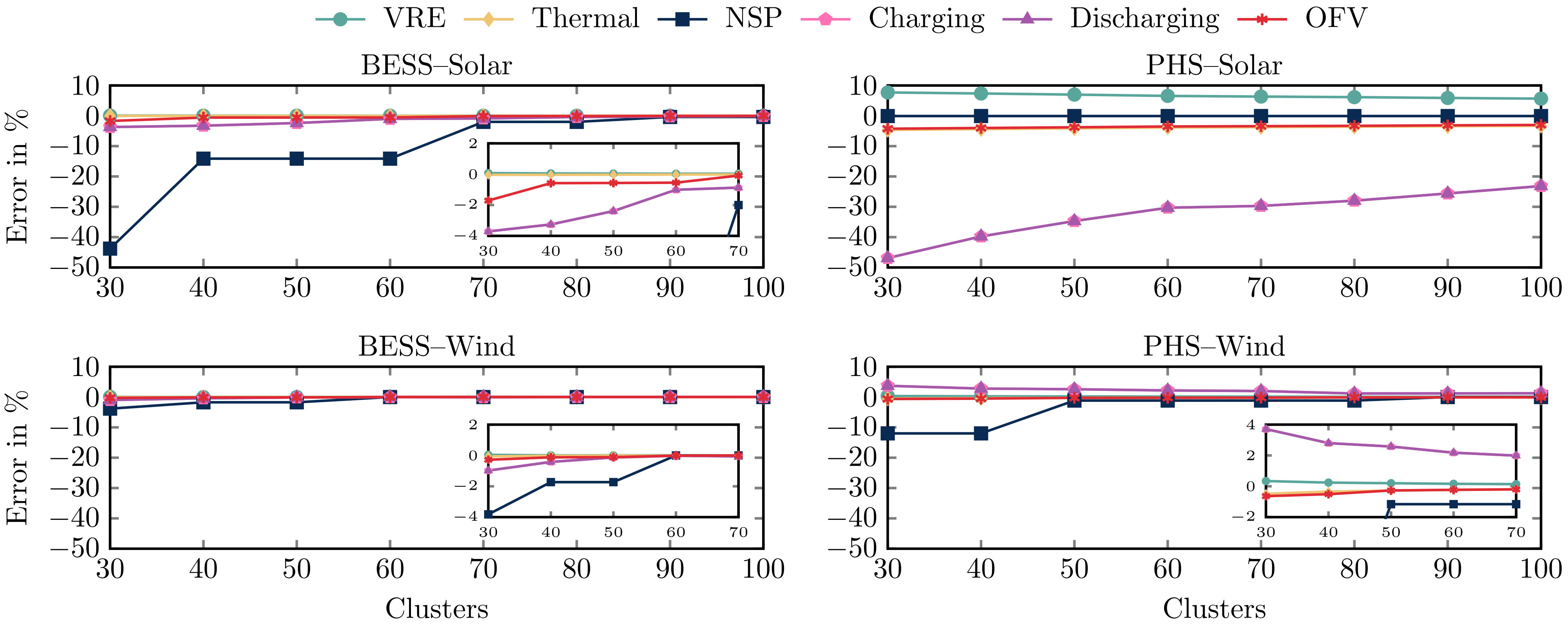


1. **Disaggregation** into unlinked submodels happens when storage is empty (ACS1)
2. **Aggregation** within a submodel: consecutive time periods can be **aggregated perfectly** within the **same active constraint sets** (e.g., **ACS2 storage idle**; **ACS4 storage charging**; **ACS6 storage discharging 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).

Exact TSA with storage in practice

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

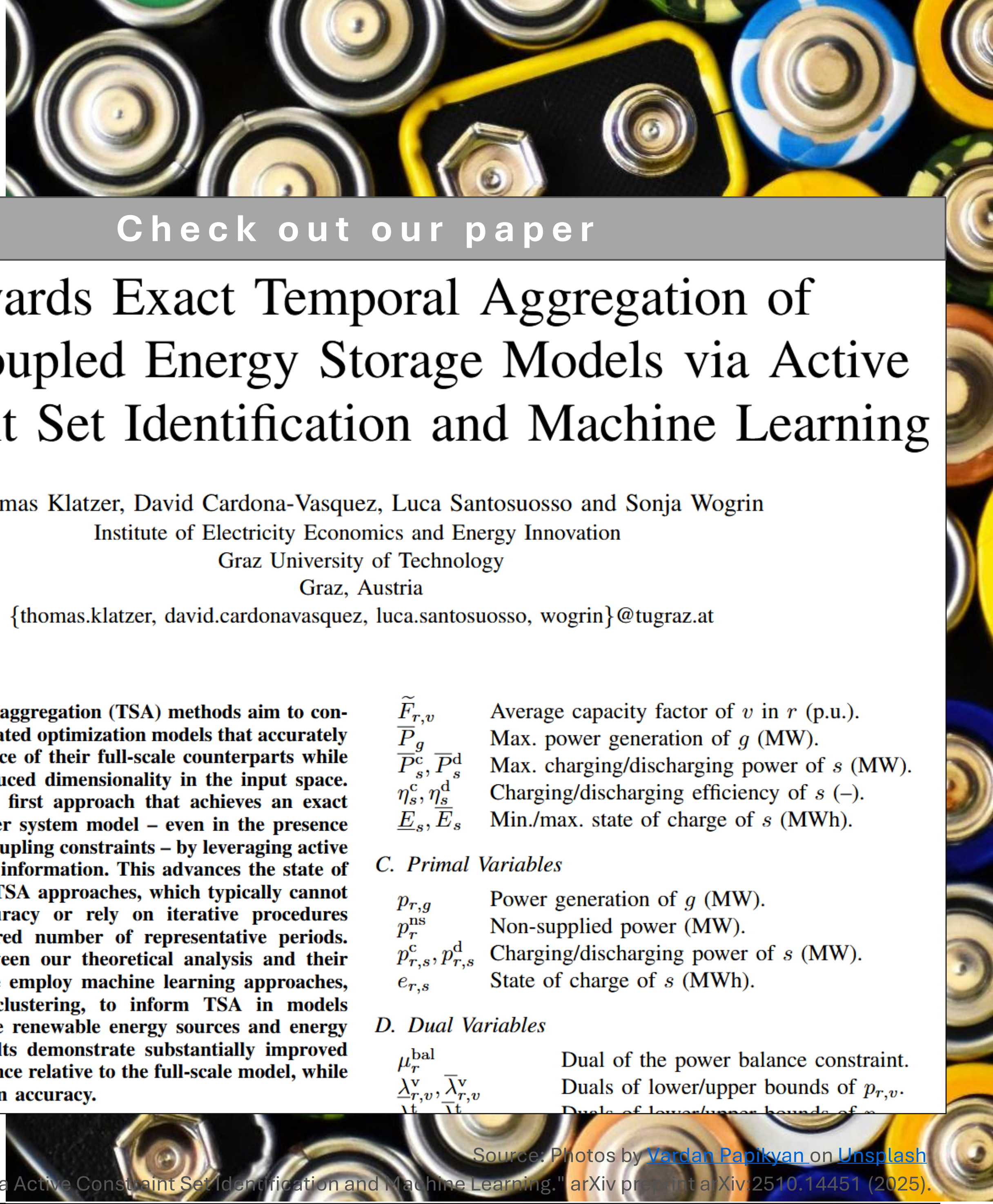


Results

- **Small relative output error** depends on the number of clusters in aggregation
- Up to 486-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



Exact TSA with storage

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

Ongoing 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).
$\overline{P}^{\mathrm{c}}_s, \overline{P}^{\mathrm{d}}_s$	Max. charging/discharging power of s (MW).
$\eta^{\mathrm{c}}_s, \eta^{\mathrm{d}}_s$	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).
$p^{\mathrm{c}}_{r,s}, p^{\mathrm{d}}_{r,s}$	Charging/discharging power of s (MW).
$e_{r,s}$	State of charge of s (MWh).

D. Dual Variables

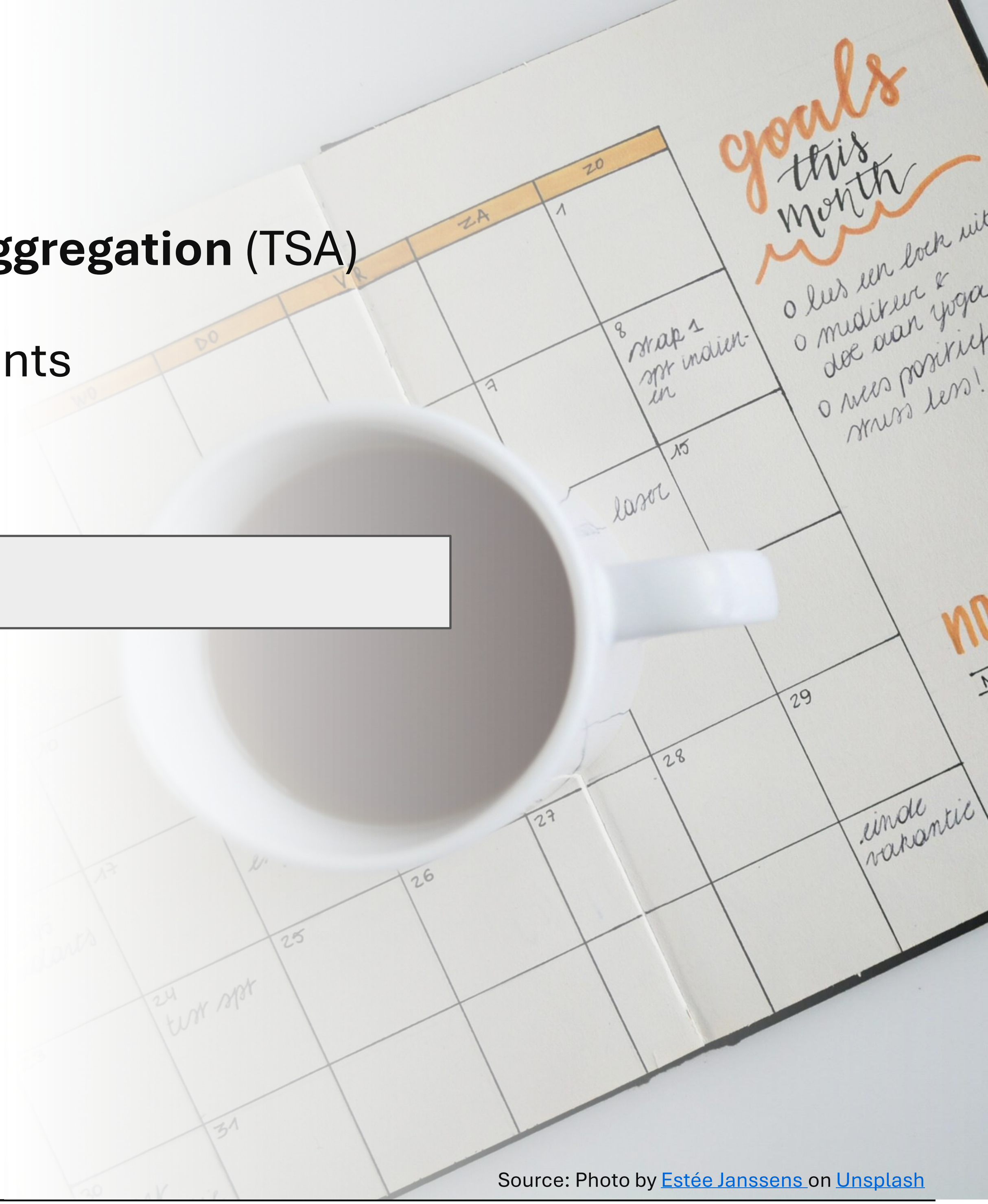
μ^{bal}_r	Dual of the power balance constraint.
$\underline{\lambda}^{\mathrm{v}}_{r,v}, \overline{\lambda}^{\mathrm{v}}_{r,v}$	Duals of lower/upper bounds of $p_{r,v}$.

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	Data Aggregation. **Motivation** & **Time Series Aggregation** (TSA)
II	TSA Extension to **Network** & **Ramping** Constraints
III	TSA Extension to **Storage** Constraints
IV	TSA with **Bounded Error**
V	Data Aggregation. **Spatial Aggregation**
VI	Conclusions

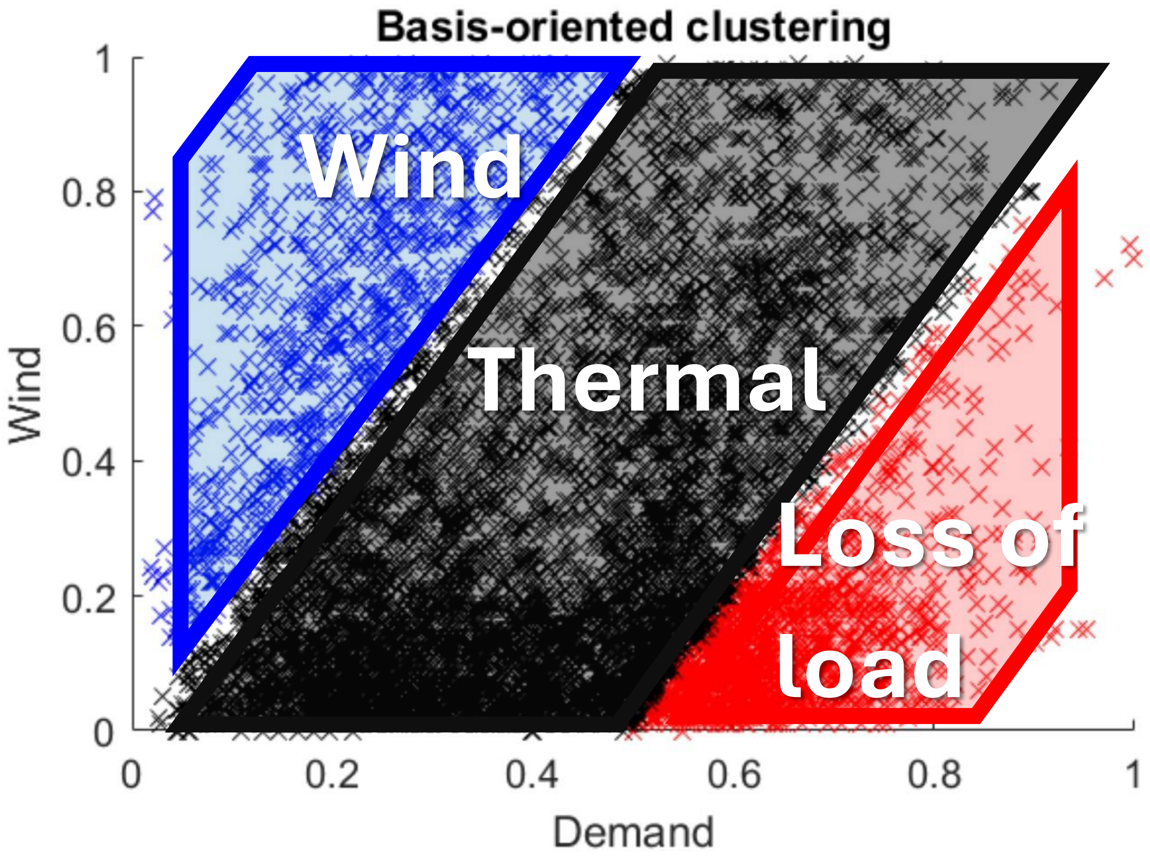


Source: Photo by Estée Janssens on Unsplash

Déjà vu

From previous research, we know that **exact TSA is possible***!

But..



- ... clusters were derived from the **active constraints** of the full-scale model.
- And we don’t know the active constraints of the model a priori 🤔

***Source**: S. Wogrin, "Time series aggregation for optimization: one-size-fits-all?," *IEEE Transactions on Smart Grid*, vol. 14, no. 3, pp. 2489-2492, 2023.

TIME SERIES AGGREGATION WITH BOUNDED ERROR

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

TIME SERIES AGGREGATION WITH BOUNDED ERROR

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.

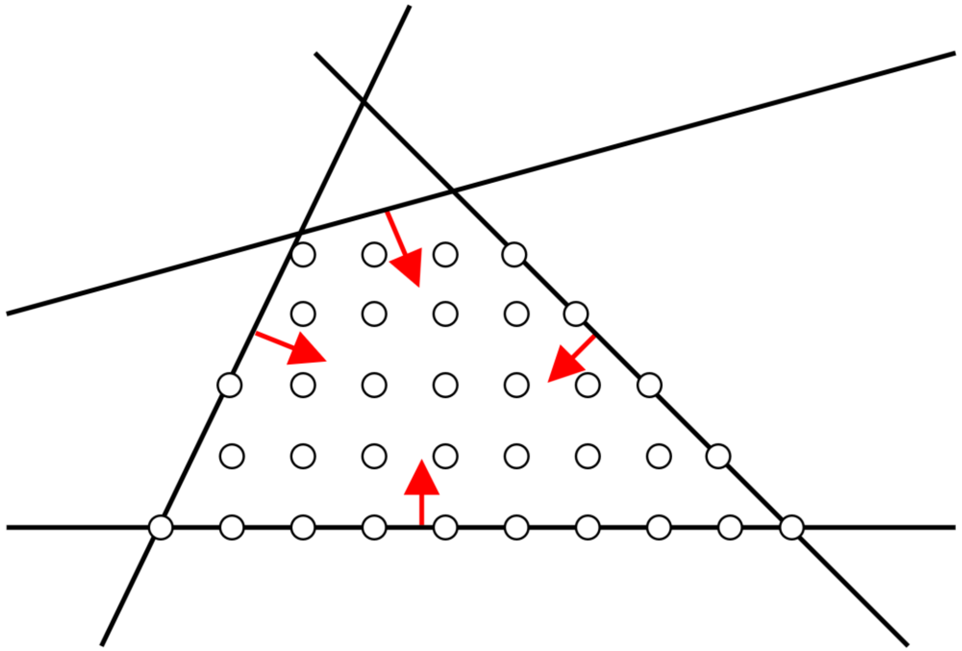
Proposition 1. Let $\boldsymbol{z}$ be a feasible solution to the full-scale model (2). Let $\hat{\boldsymbol{z}}$ be derived from $\boldsymbol{z}$ accordingly to (3)–(6). Then, $\hat{\boldsymbol{z}}$ is a feasible solution to the aggregated model (8) and it holds that

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

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," 2025 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe).

Full-scale model feasible region



TIME SERIES AGGREGATION WITH BOUNDED ERROR

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

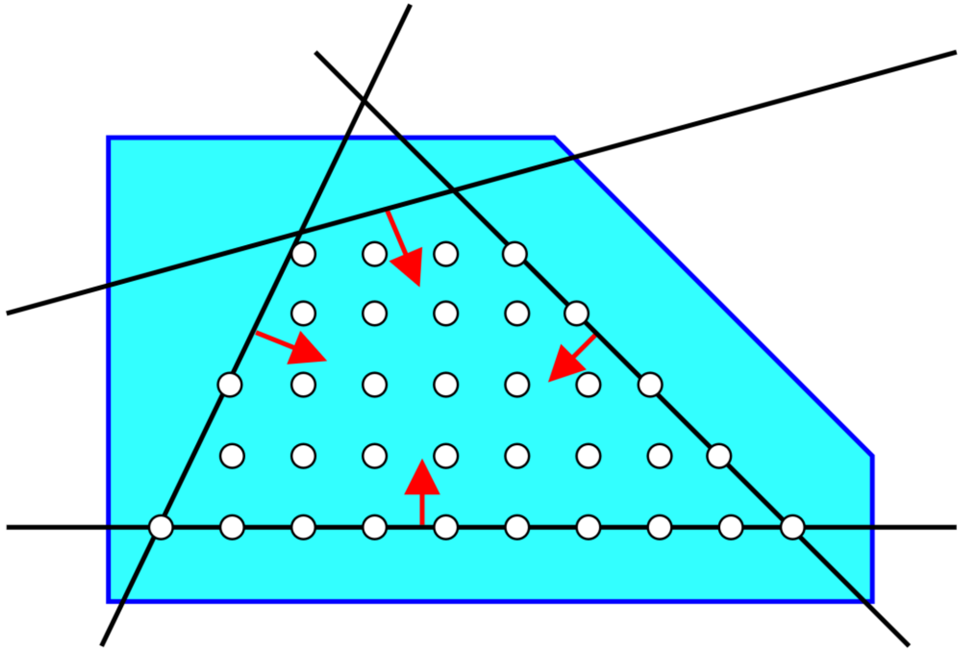
Proposition 1. Let z be a feas
model (2). Let $\hat{z}$ be derived fr
Then, $\hat{z}$ is a feasible solution to the aggregated model (8) and it holds that

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

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

Full-scale obj.
=
Aggregated obj.

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



***Source**: **L. Santosuosso** and S. Wogrin, "Optimal virtual power plant investment planning via time series aggregation with bounded error," 2025 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe).

TIME SERIES AGGREGATION WITH BOUNDED ERROR

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

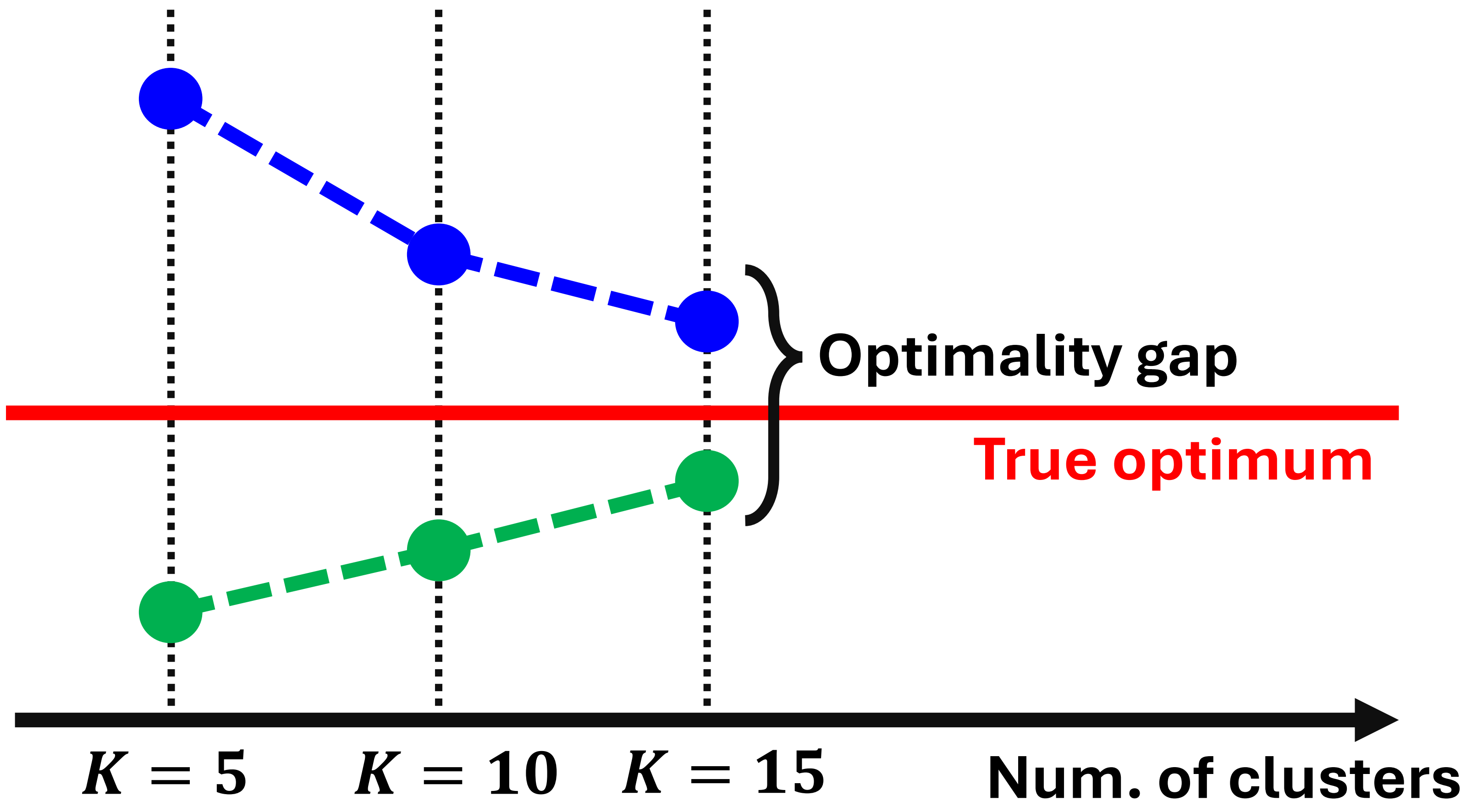
Input: Parameters $\{F_{g,t}, D_t, \overline{X}_g, \underline{X}_g \,|\, g \in \boldsymbol{G}, t \in \boldsymbol{T}\}$, initial number of clusters K^0, step size α, optimality threshold

(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,“ 2025 IEEE PES ISGT Conference Europe (ISGT Europe).



TIME SERIES AGGREGATION WITH BOUNDED ERROR

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 $J^{\text{UB}^{i+1}}$;
$K^{i+1} \leftarrow K^i + \alpha \lfloor \epsilon^{i+1} \rfloor$;
$i \leftarrow i + 1$;
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," 2025 IEEE PES ISGT Europe.

Lower bound:

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

Upper bound:

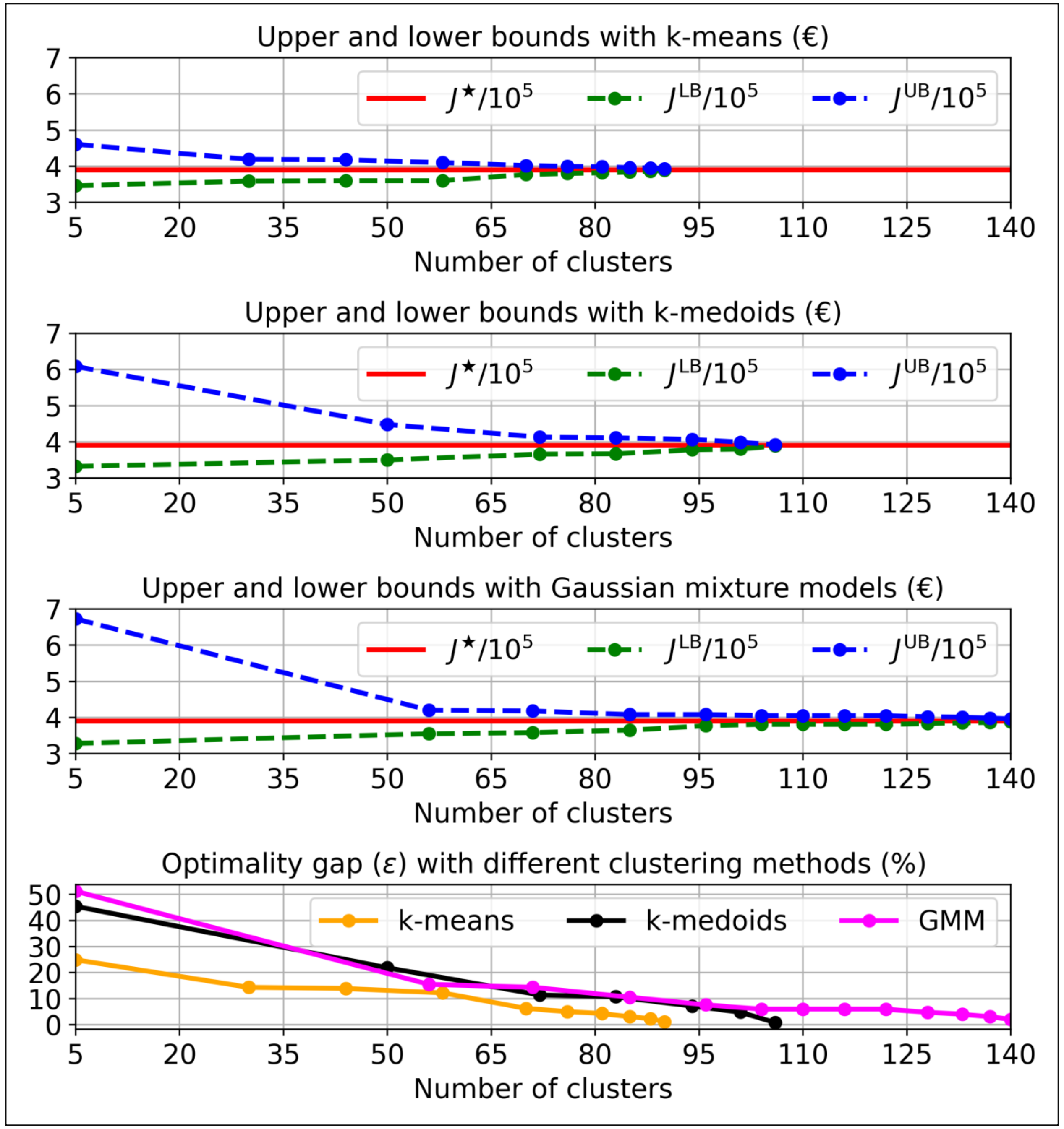
- From an **LP** model.
- Provides a **feasible solution** at each iteration.

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
$\|T\| = 8760$ $\|G\| = 100$	273	431 (**+58%**)	516 (+89%)	689 (+152%)
$\|T\| = 17520$ $\|G\| = 100$	793	956 (**+21%**)	1360 (+72%)	1892 (+139%)
$\|T\| = 8760$ $\|G\| = 1000$	4032	3346 (**−17%**)	5113 (+27%)	7629 (+89%)
$\|T\| = 17520$ $\|G\| = 1000$	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

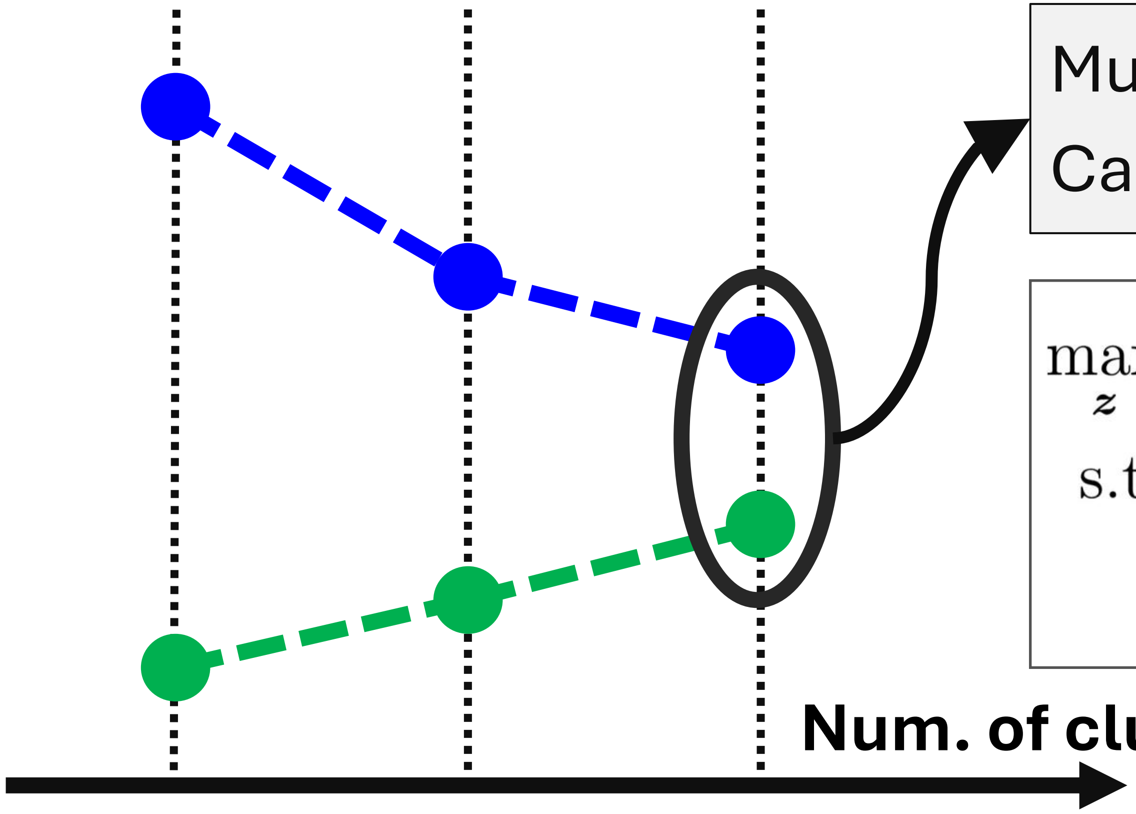
Sonja Wogrin
Institute of Electricity Economics and Energy Innovation
Graz University of Technology
Graz, Austria
wogrin@tugraz.at

***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 $M(\boldsymbol{z})$?

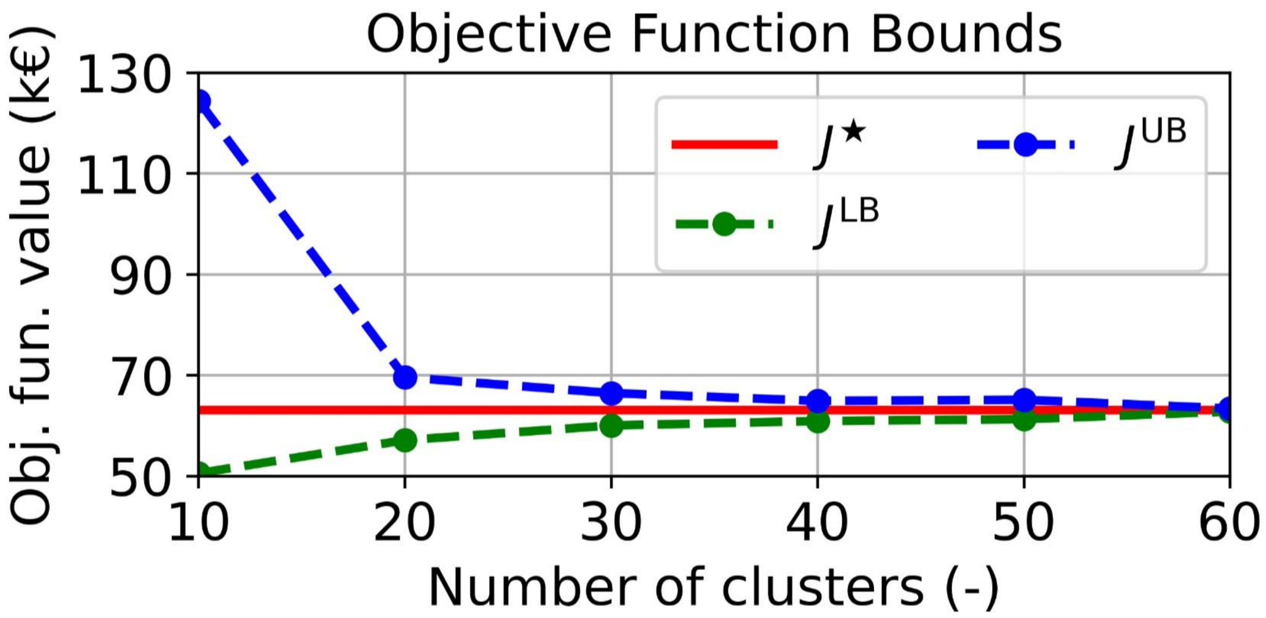
$$\max_{\boldsymbol{z}} \quad M(\boldsymbol{z})$$
$$\text{s.t.} \quad \text{Full-scale dynamics}$$
$$J^{\mathrm{LB}} \leq J(\boldsymbol{z}) \leq J^{\mathrm{UB}}$$

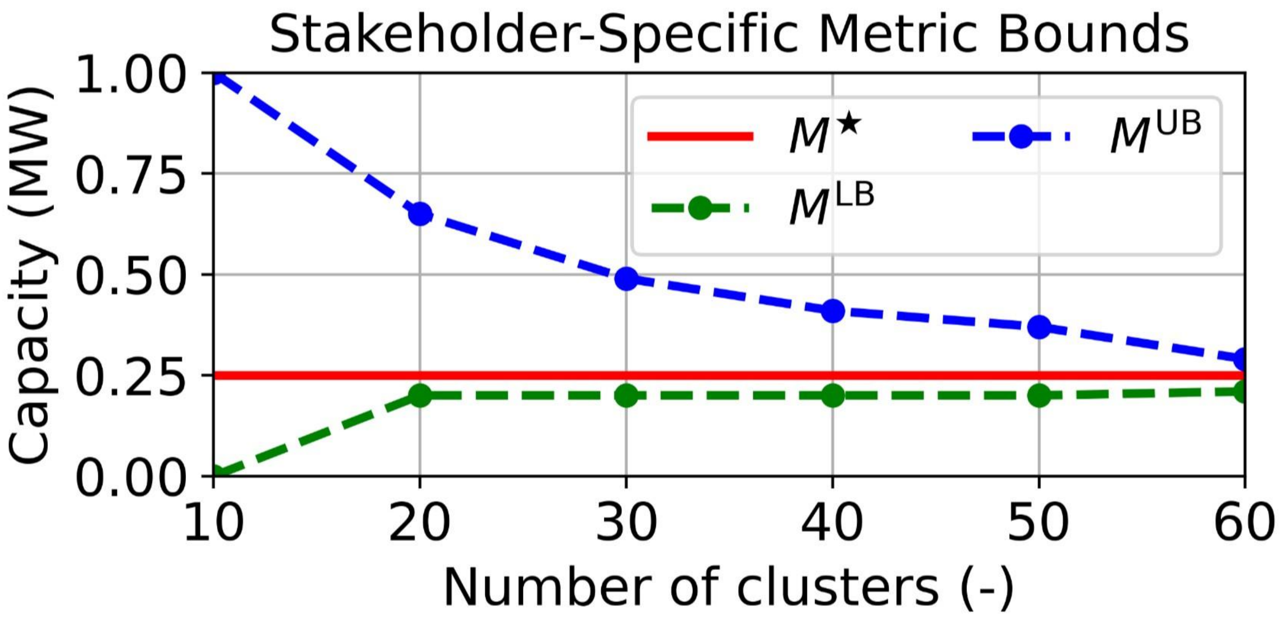
$$\min_{\boldsymbol{z}} \quad M(\boldsymbol{z})$$
$$\text{s.t.} \quad \text{Full-scale dynamics}$$
$$J^{\mathrm{LB}} \leq J(\boldsymbol{z}) \leq J^{\mathrm{UB}}$$

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.

Key Messages: Time Series Aggregation

- Modern energy system models are **complex and often intractable**, but TSA can help.
- However, traditional TSA methods often rely on **heuristics** and leave a lot of aggregation potential untapped.
- Remember the 3 Open Questions we posed:

> Can this methodology be extended to more complex power system models, e.g., including **network** and **time-linking constraints**?

> Is **exact TSA** possible with time-linking constraints (i.e., **ramping or storage**) when model structure is accounted for?

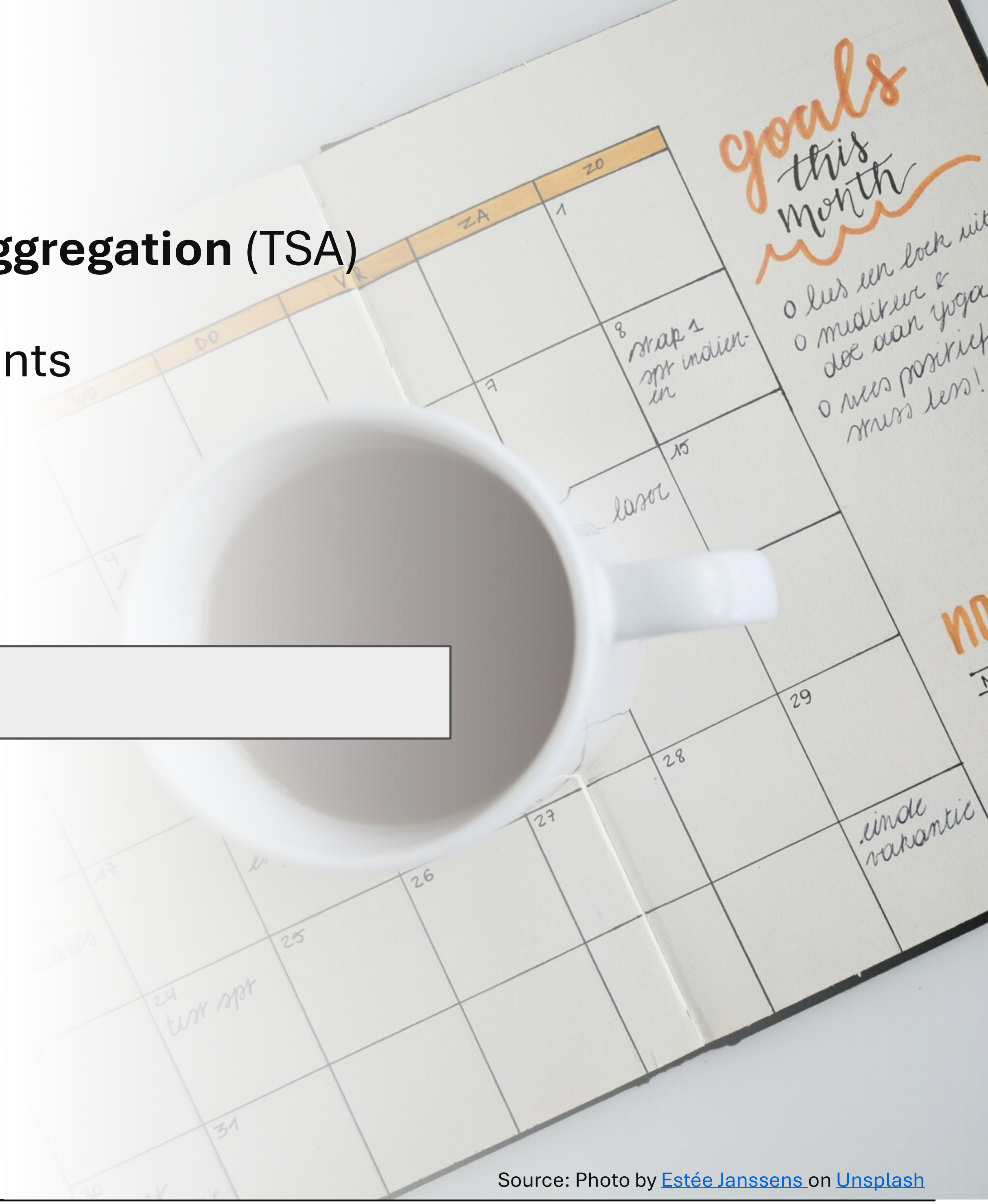
> Can we establish valid **performance bounds** for TSA aggregated models?

- **Our research:**
  - **Leverages aggregation potential to get more bang for your buck.**
  - Shows that **exact TSA is possible with network, ramping, and storage constraints.**
  - Delivers TSA methods with **strong performance guarantees** via formal error bounds.

Source: Photo by Markus Spiske on Unsplash

Agenda

- **I** Data Aggregation. **Motivation** & **Time Series Aggregation** (TSA)
- **II** TSA Extension to **Network** & **Ramping** Constraints
- **III** TSA Extension to **Storage** Constraints
- **IV** TSA with **Bounded Error**
- **V** Data Aggregation. **Spatial Aggregation**
- **VI** Conclusions



Source: Photo by Estée Janssens on Unsplash

Network Partitioning and Aggregation Package

Designed for general graph-based structures with a focus on Networks for Optimization Models

Partition

A **partition** of a network is a division of its nodes into disjoint groups (indicated by colors) such that every node belongs to exactly one group.

Aggregation

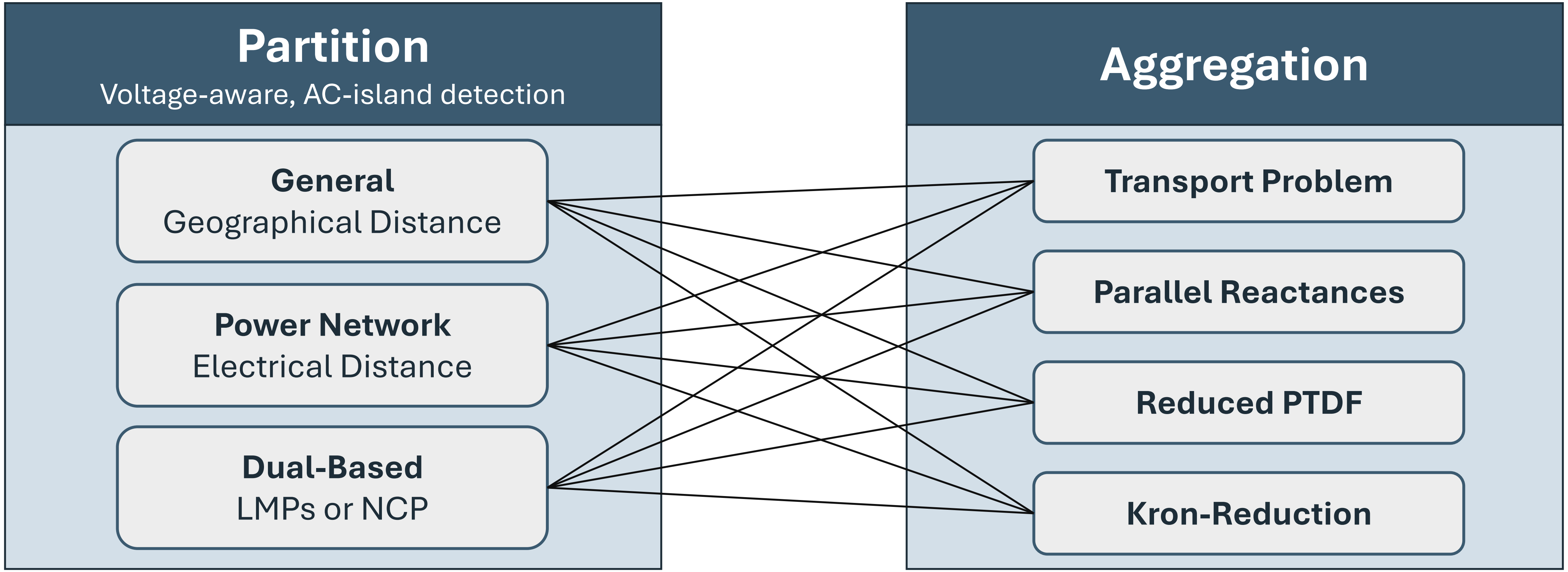
An **aggregation** reduces the network topology based on a given partition by aggregating nodes, edges, and their associated properties.

Scan it and **check out** our GitHub Repository 🚀

NPAP

Network Partitioning and Aggregation Package

Designed for general graph-based structures with a focus on
Networks for Optimization Models

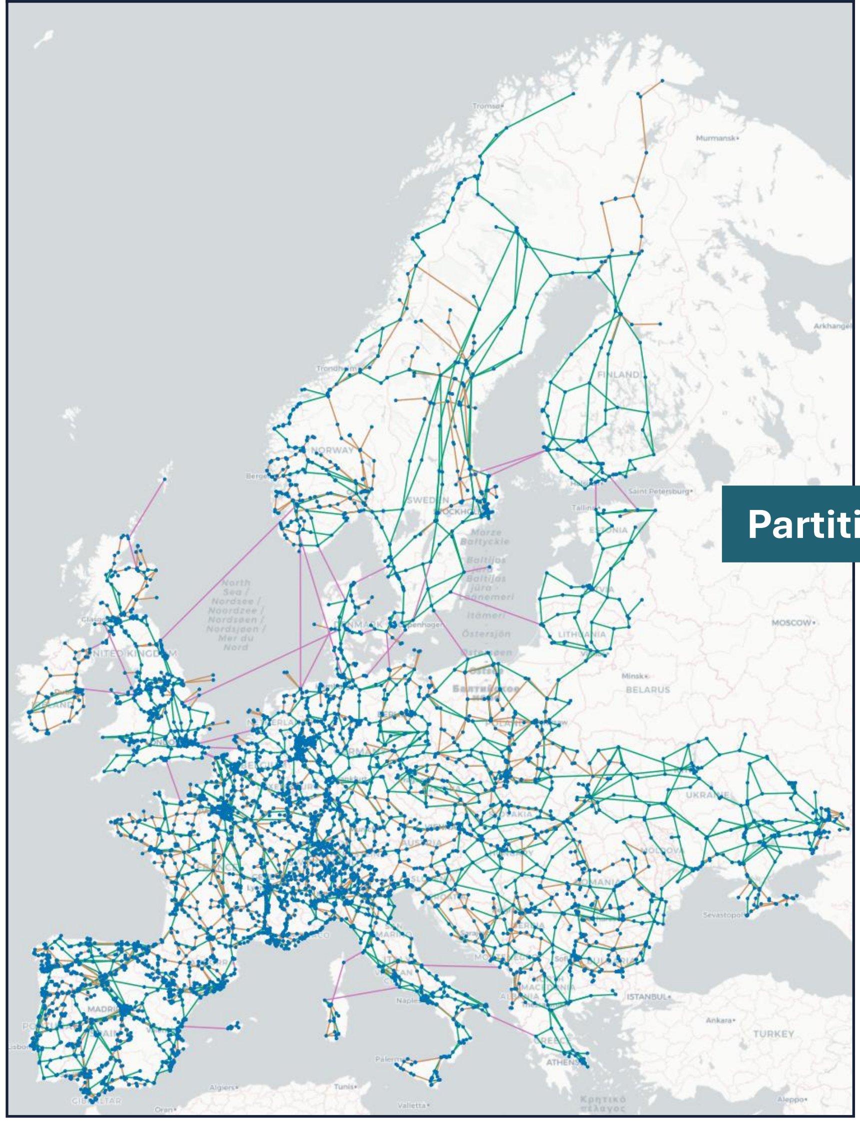


Geographical Partitioning

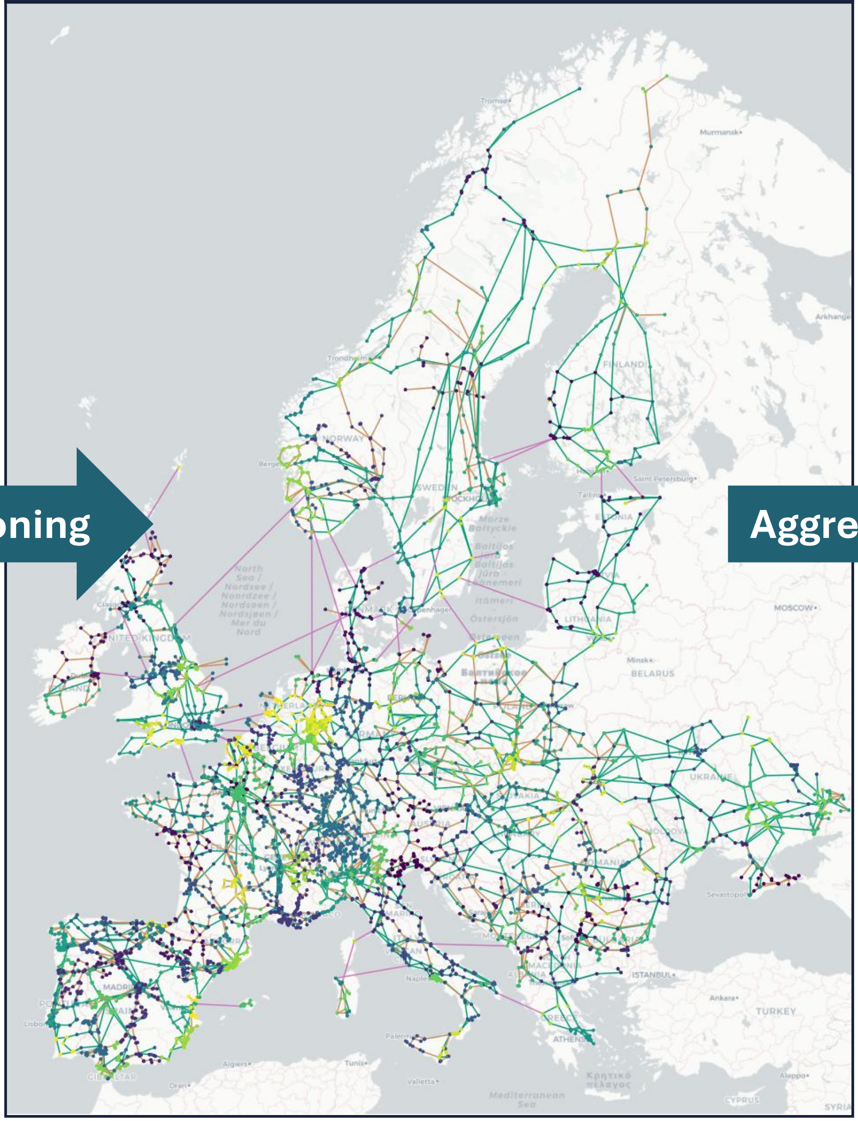
250 Clusters, Voltage-Aware, PyPSA Europe dataframe

NPAP www.tugraz.at

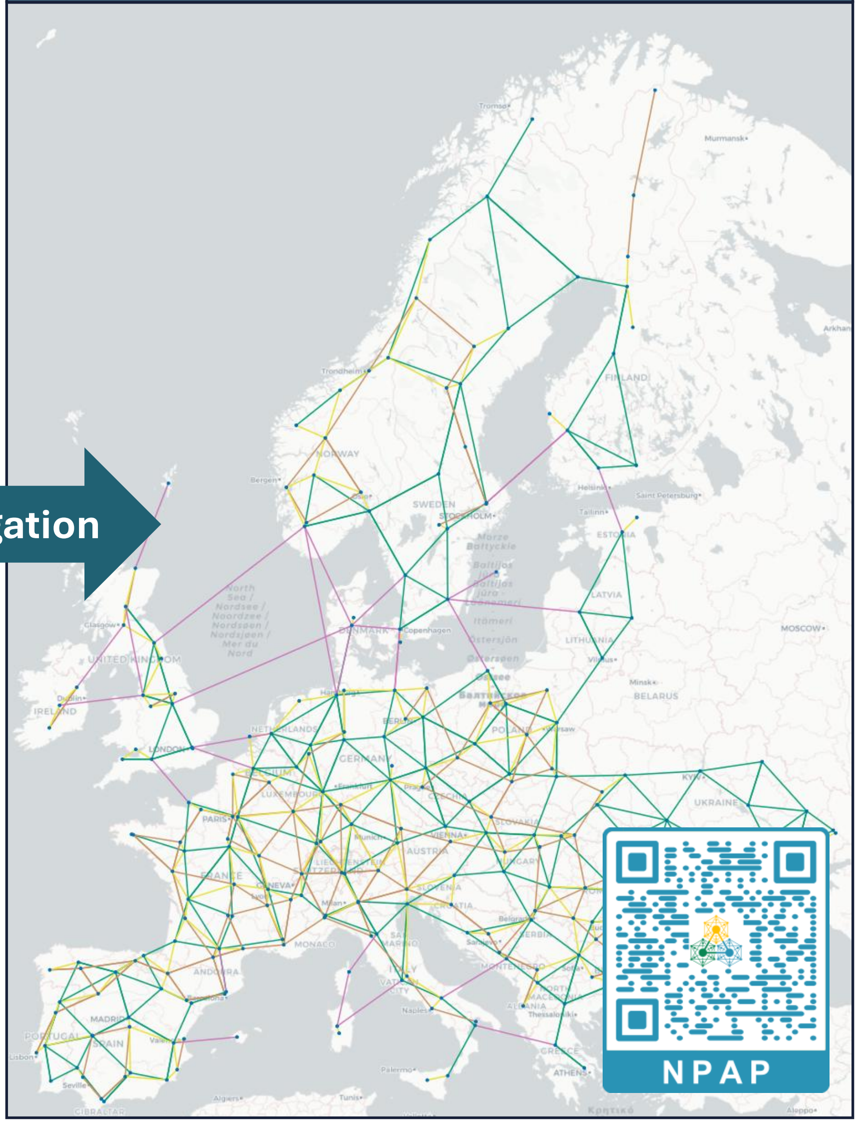
Full European HV Network	Partitioned Network	Aggregated Network



Partitioning

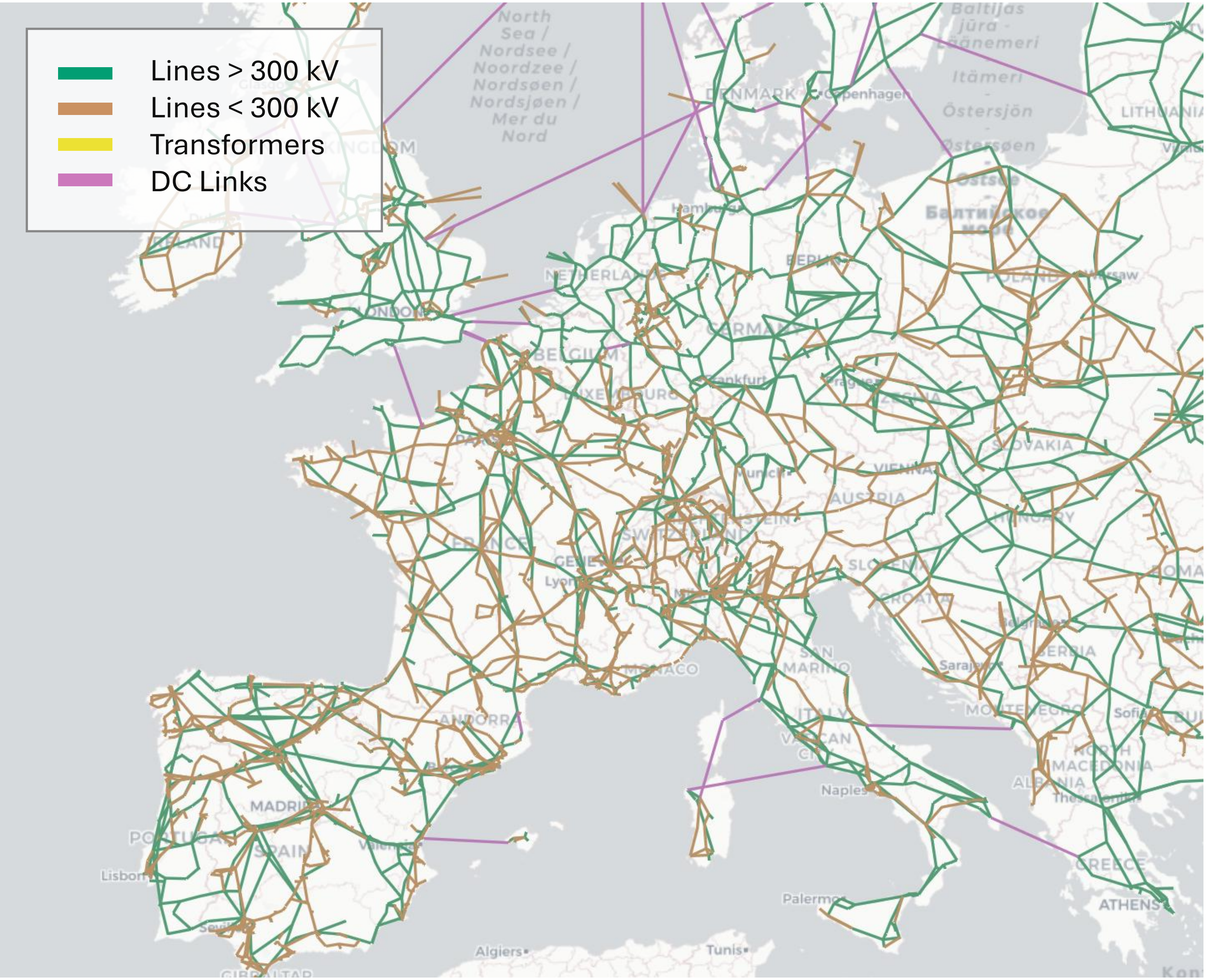


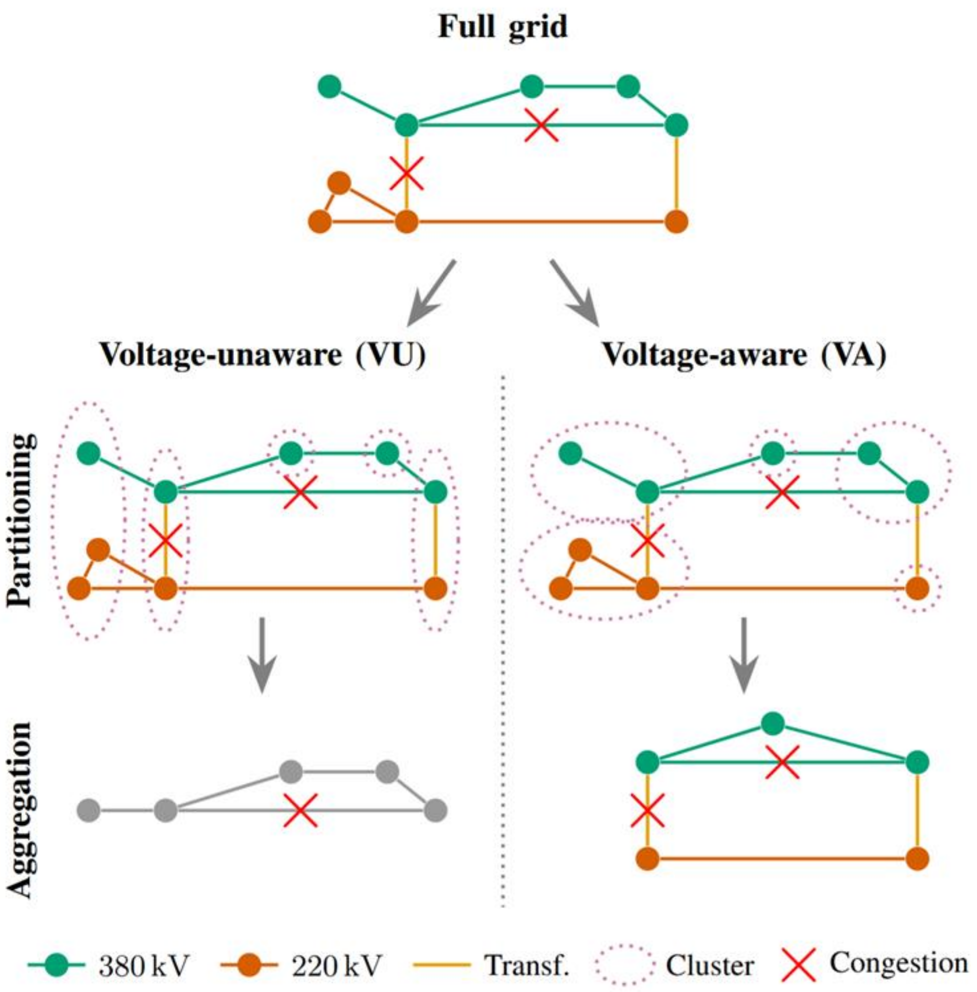
Aggregation



https://doi.org/10.5281/zenodo.7646728

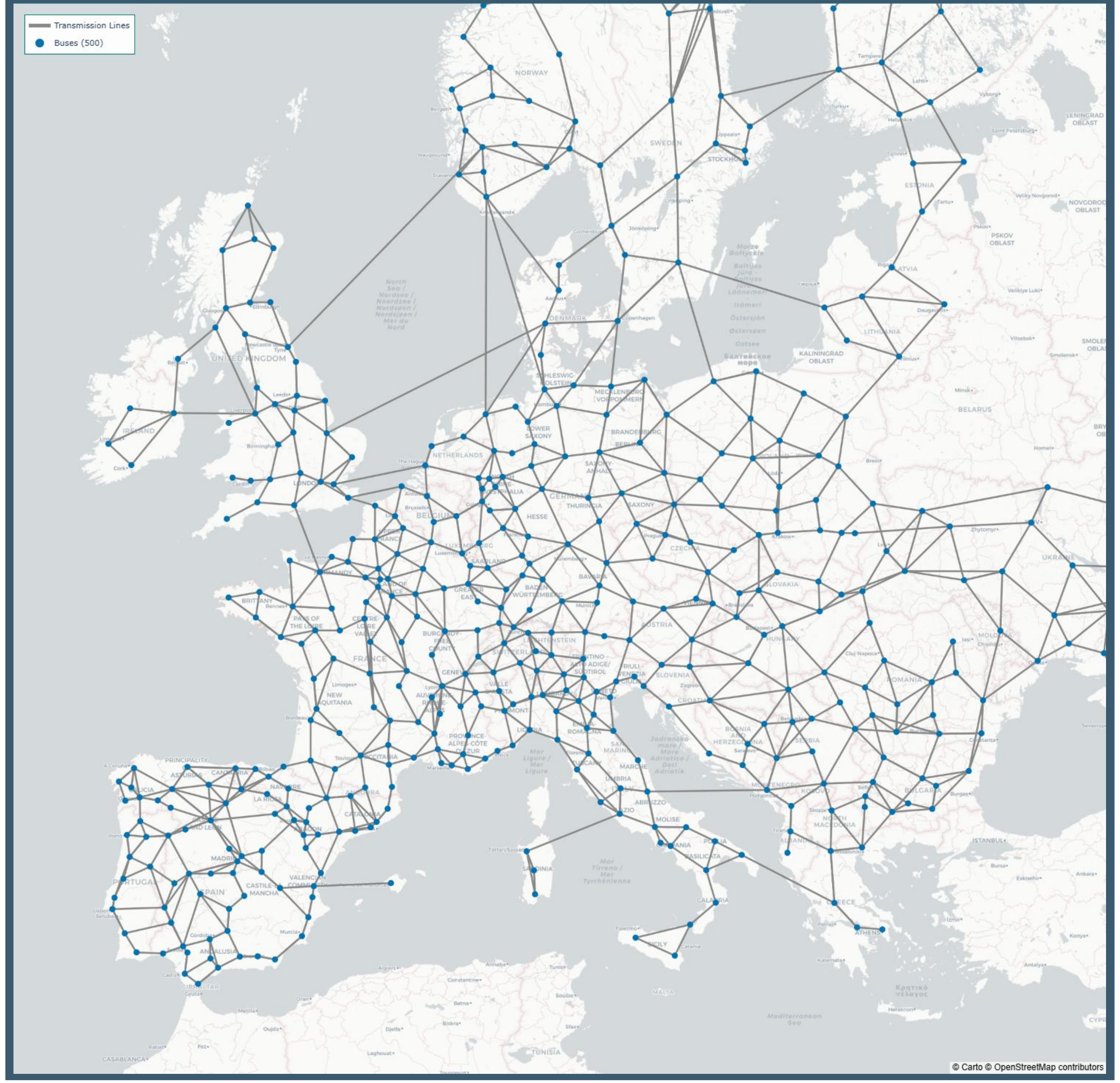
Why voltage-aware?



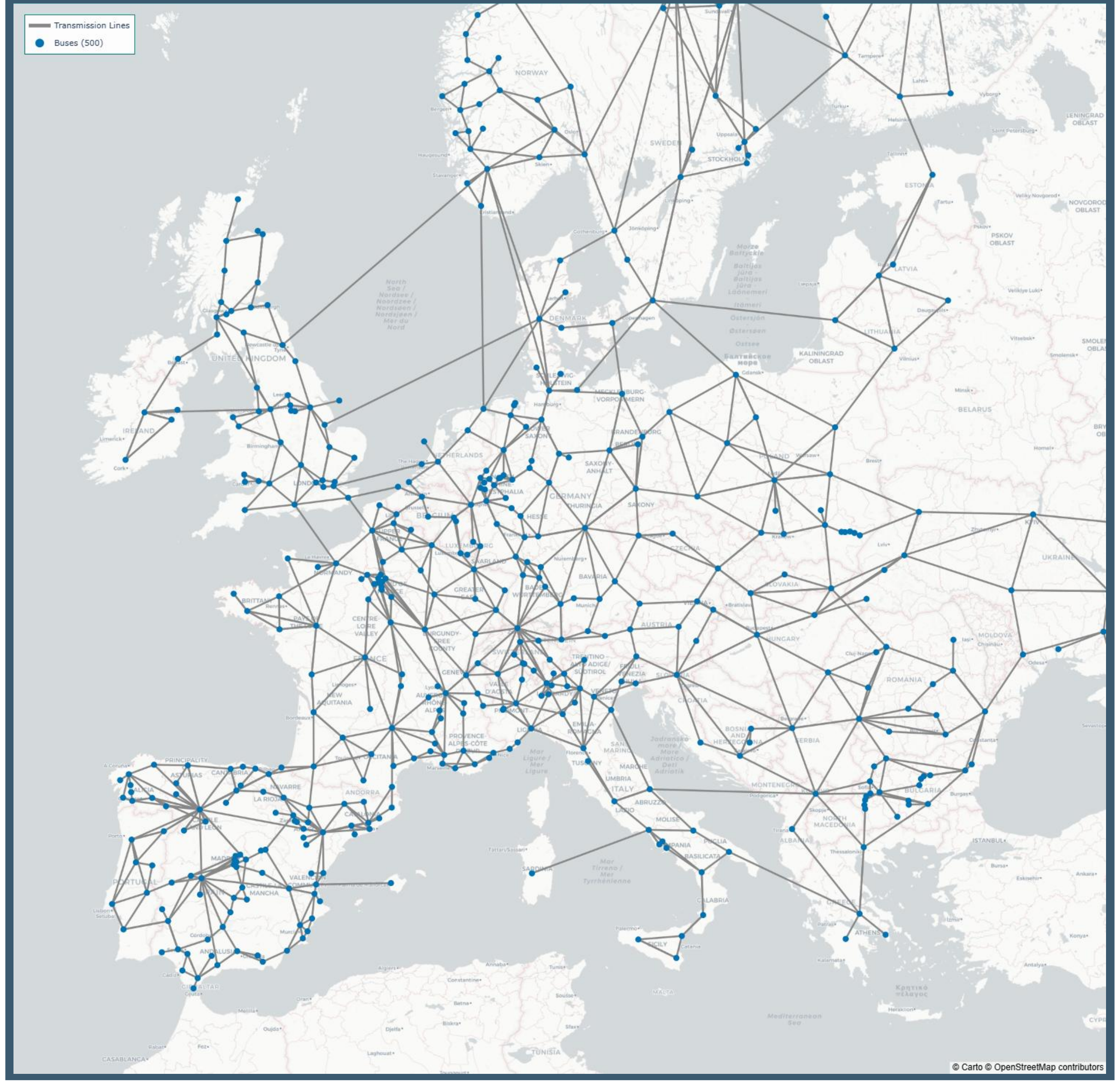


Voltage-unaware Partitioning

Geographical distance k-medoids

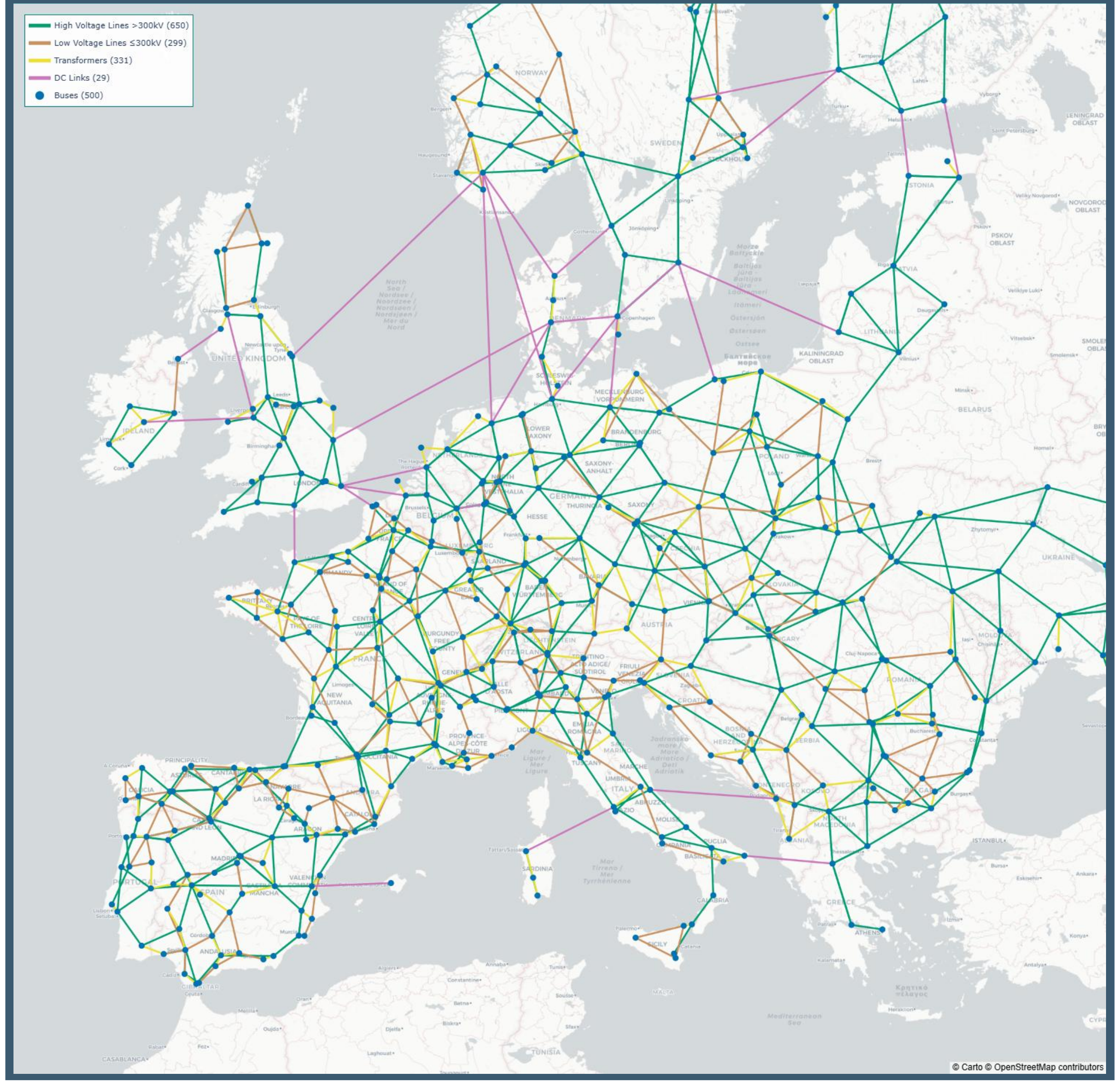


Electrical distance k-medoids

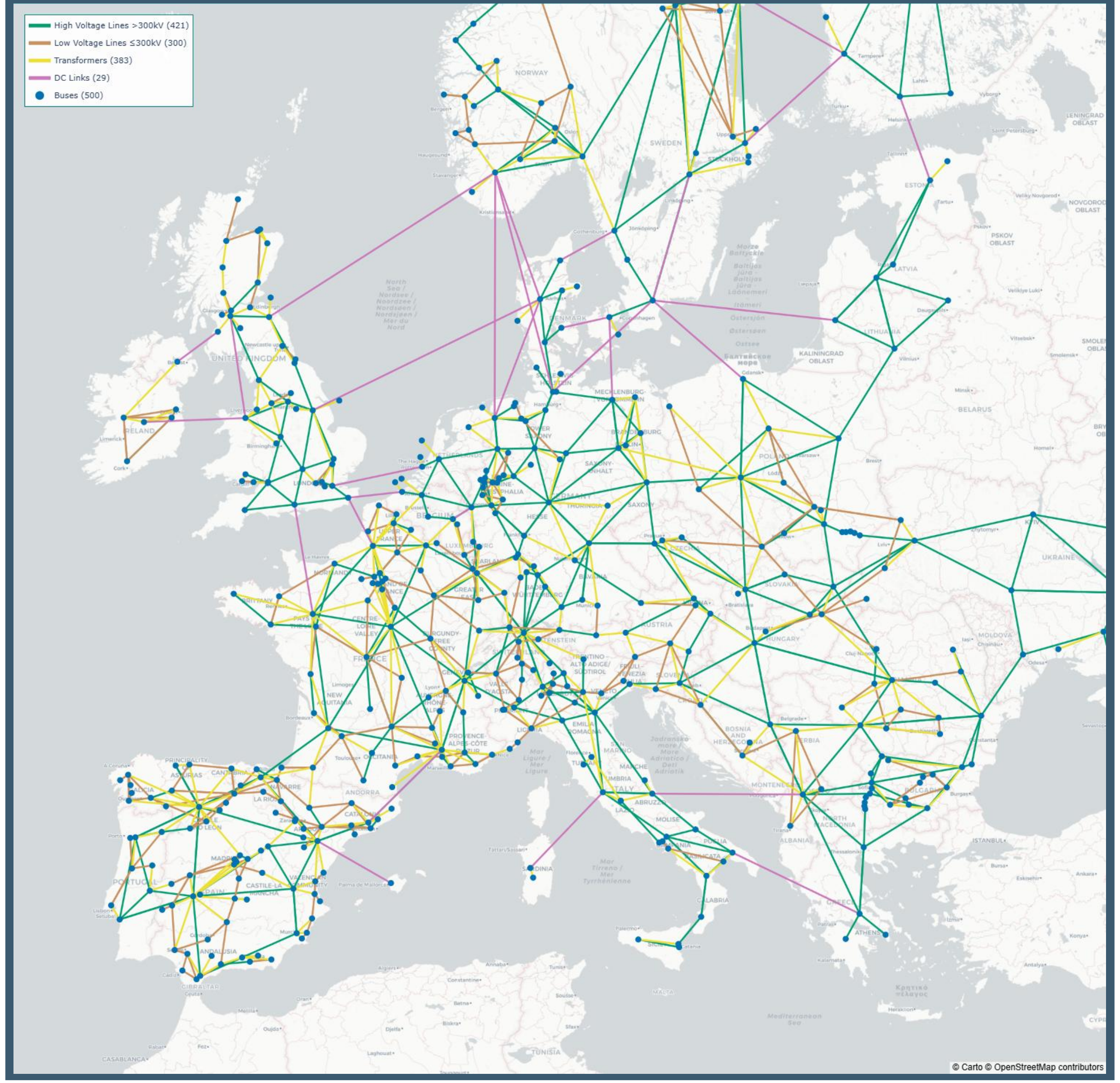


Voltage-aware Partitioning

Geographical distance k-medoids

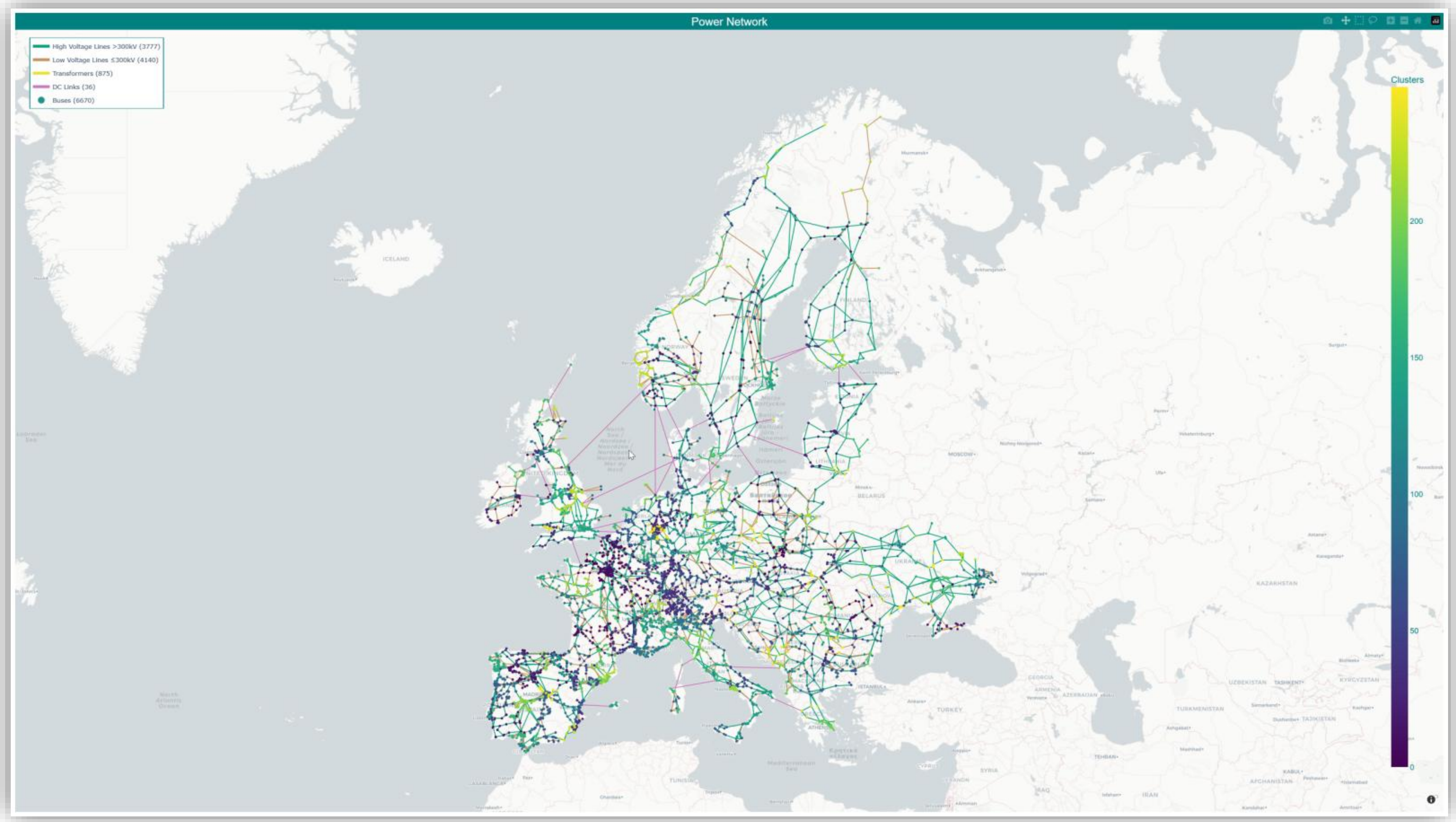


Electrical distance k-medoids



Quick NPAP Tutorial

- Check out a short NPAP tutorial on Colab.
    - Install package & load (PyPSA) data
    - Partition and aggregate European grid



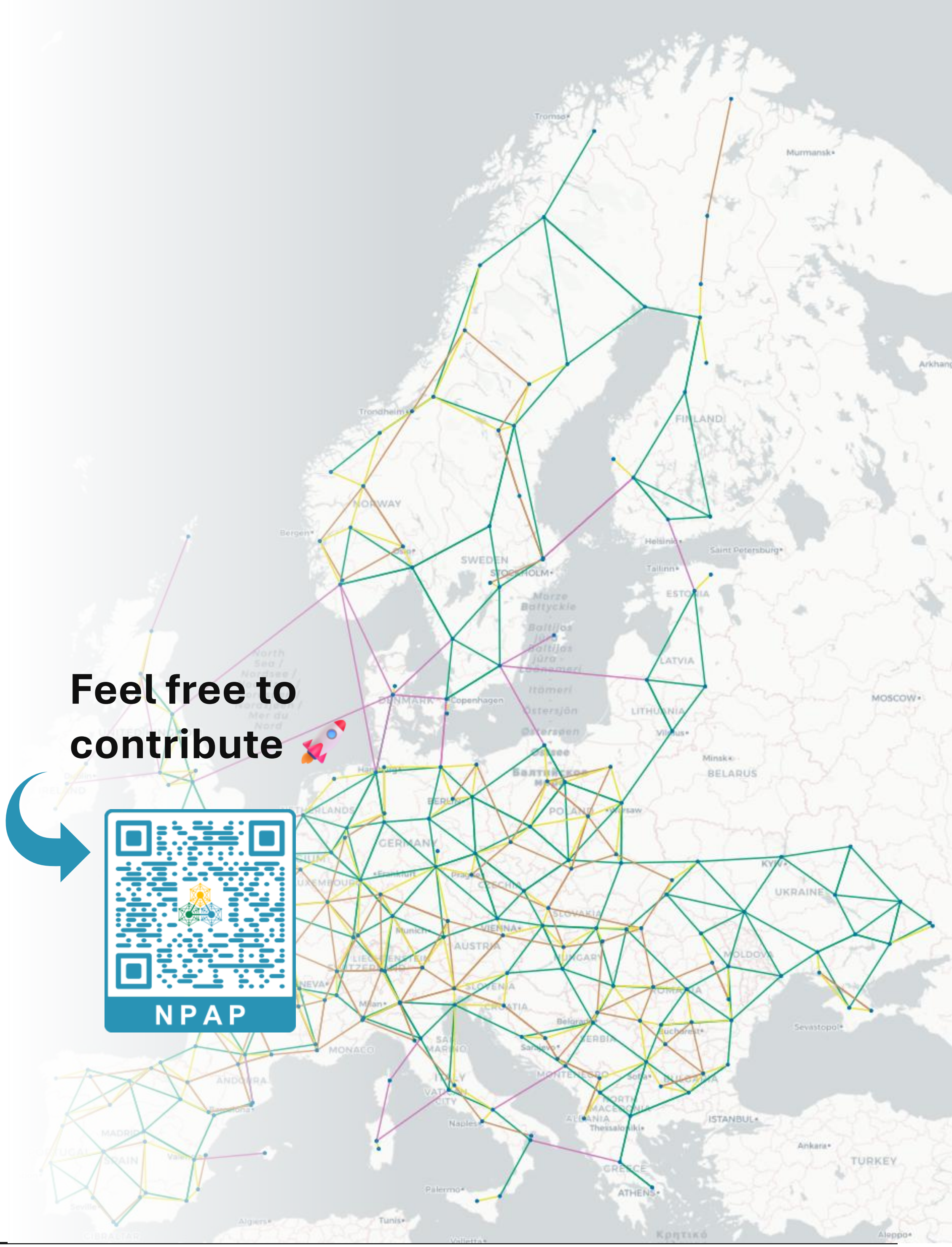
Scan it and check out a short **NPAP tutorial** 🚀

https://colab.research.google.com/drive/1_ror9ie3HfMWhJOttzUS-XeTtEXa2-FK

Conclusions

- **Computational complexity** of energy system models can be tackled by data aggregation: in time or/and space.
- **Exact aggregation** is possible when model structure and active constraint sets are exploited — even with networks, ramping, and storage.
- **Disaggregation, parallelization, and aggregation** deliver orders-of-magnitude speedups with negligible error. **Bounded-error methods** enable scalable, reliable decision-making for future energy systems.
- The **NPAP software package** supports knowledge transfer.
- **Take-home message**: *The energy transition needs not only better models — but smarter methods behind them.*

Feel free to contribute 🚀



Thank you for the attention!

Sonja Wogrin

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

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

QR code to our website!

NETZERO-OPT

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.*