Identifying Representative Weeks for the Finnish Power Sector with Time-Series Clustering

Ilona Ylikoski

22.11.2024

Advisor: *Afzal Siddiqui*

Supervisor: *Afzal Siddiqui*

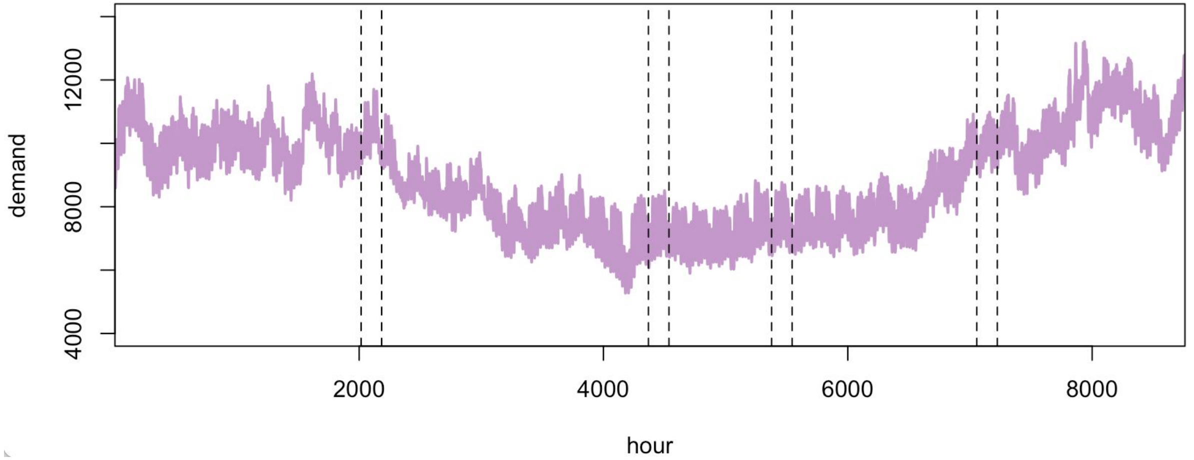
Aalto-yliopisto
Perustieteiden korkeakoulu

Background

- Power system models can be complex and computationally intractable problem instaces
- To simplify the models, their temporal resolution can be reduced
  - Time series constraints can be simplified by using representative periods instead of full datasets.

Goals

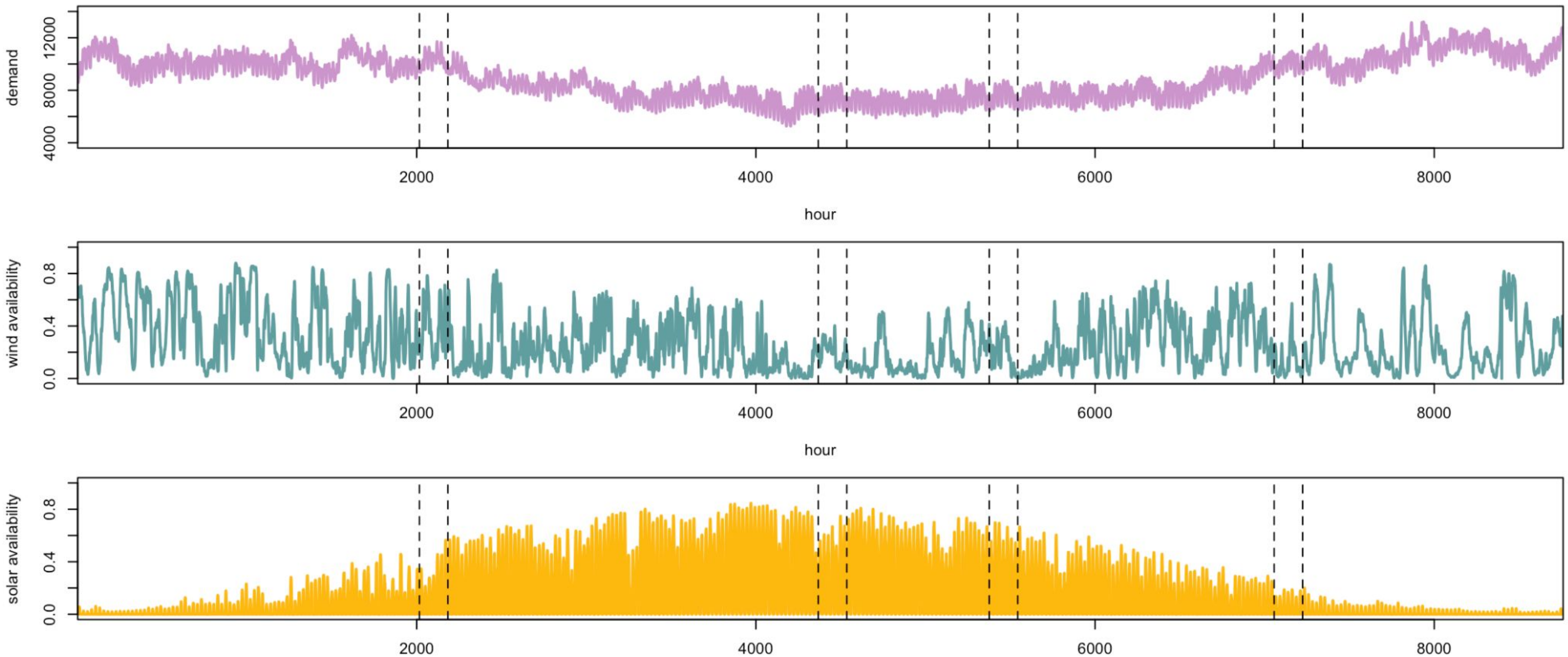
- Apply time-series aggregation techniques to Finnish demand, wind and solar data.
- Compare the results of two different approaches to aggregate data into four representative weeks
    1. Partitional approach
    2. Seasonal approach



Aalto-yliopisto
Perustieteiden
korkeakoulu

Data

- Hourly time-series for consumption [1] as well as wind [2], and solar [3] availability for 2023

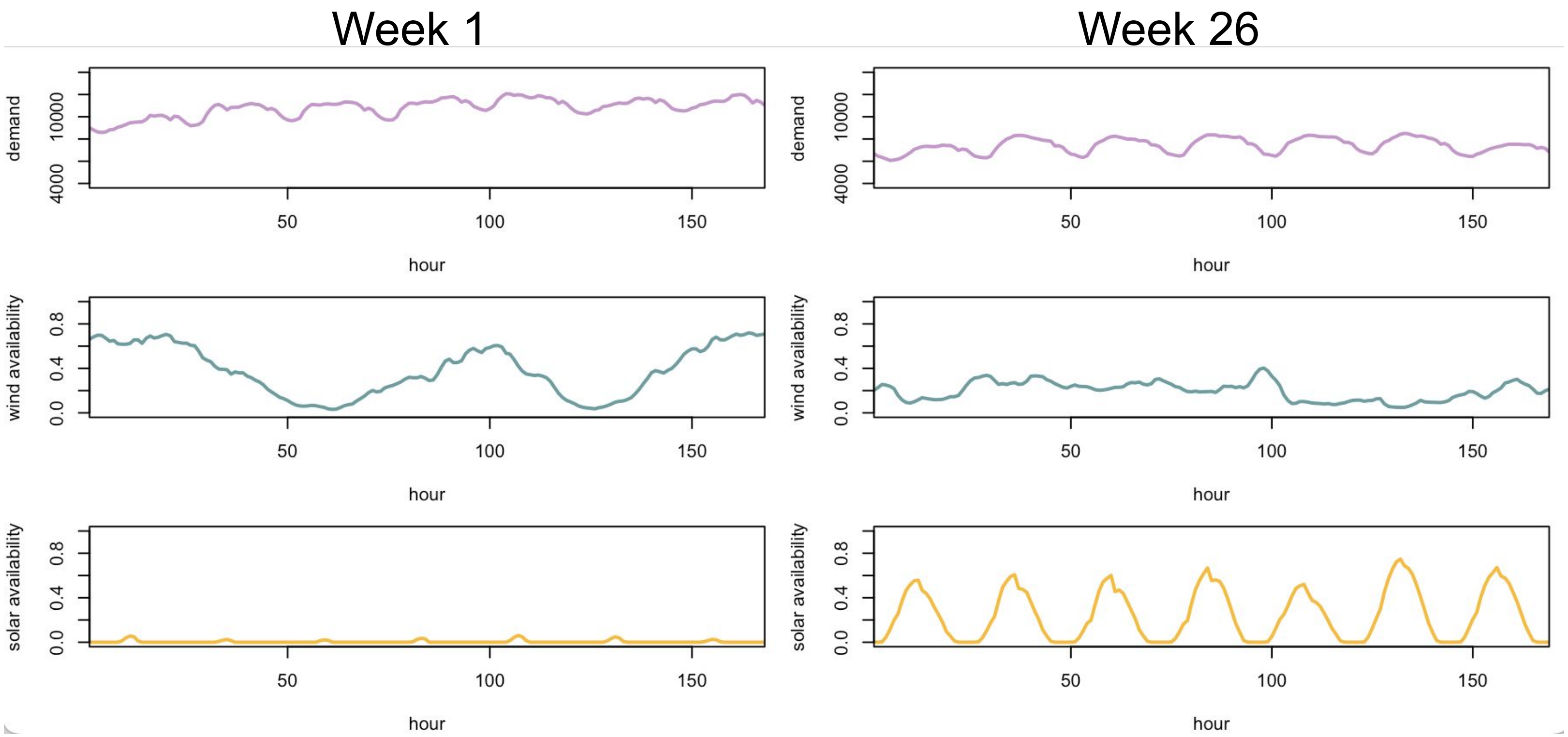


Aalto-yliopisto
Perustieteiden
korkeakoulu

[1] Nord Pool data portal: Consumption, https://data.nordpoolgroup.com/power-system/consumption
[2] Fingrid Open Data, Wind Power, https://www.fingrid.fi/en/electricity-market-information/wind-power-generation/
[3] Fingrid Open data, Solar Power, https://www.fingrid.fi/en/electricity-market-information/solar-power/

Data

- Data are split into 52 168-hour weeks
- The last day is represented by the first 24 hours of the previous week



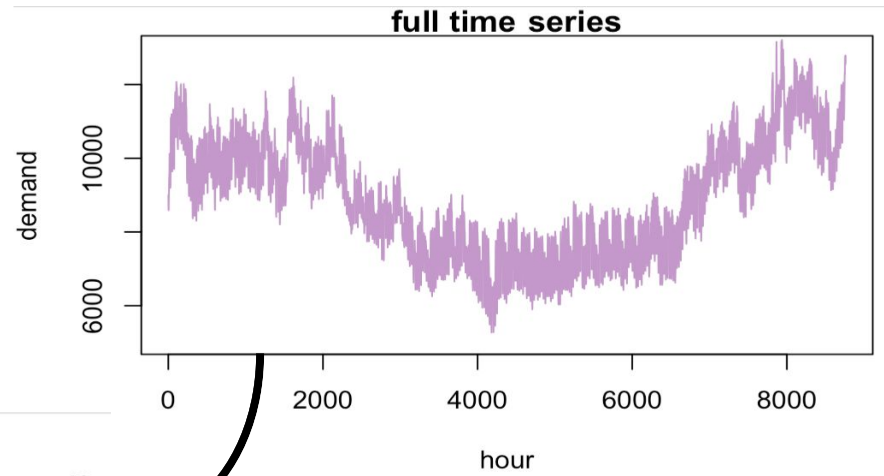
Aalto-yliopisto
Perustieteiden korkeakoulu

The Two Approaches

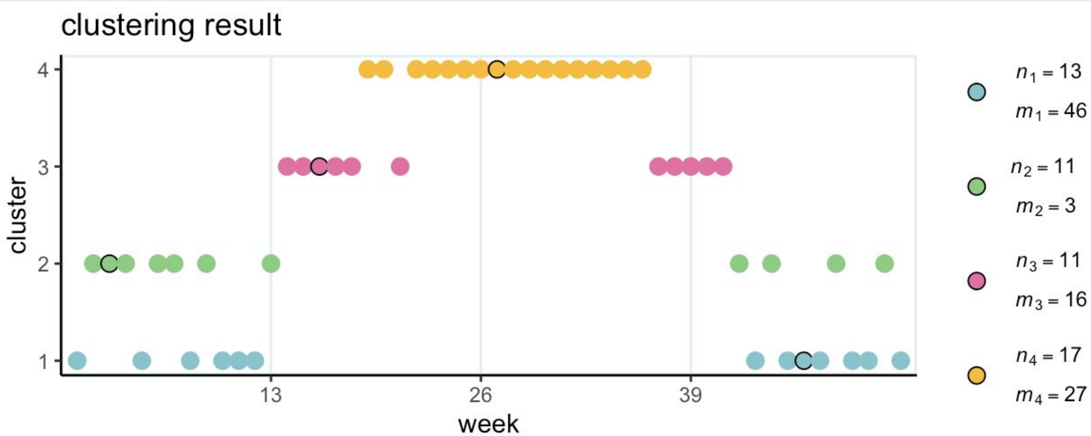
- Two different aggregation approaches for the data were compared
- Partitional approach
  - Use a clustering algorithm to determine groups in the data.
  - Choose a representative week as the medoid of each cluster.
- Seasonal approach
  - Take predetermined groups in data
  - Choose representative weeks in terms of a test statistic

Partitional Approach

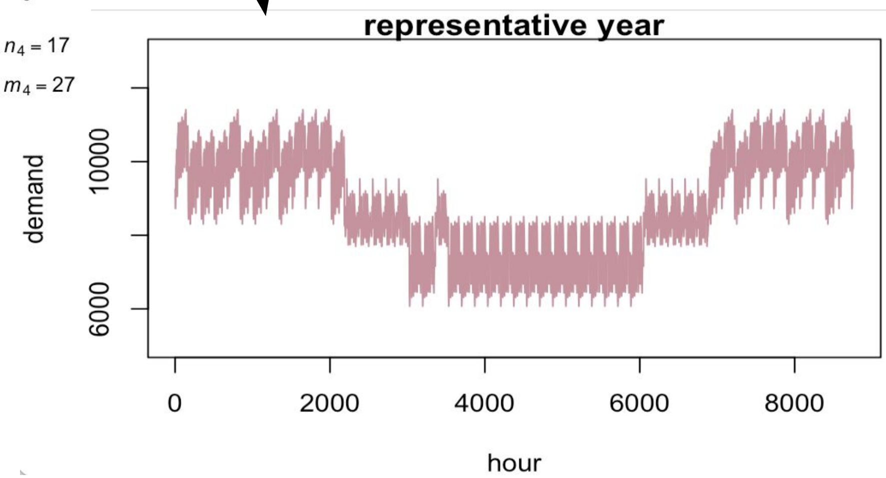
The clustering algorithm assigns each week to a cluster



distances between pairs of normalised weeks are computed



a representative year is formed for each variable

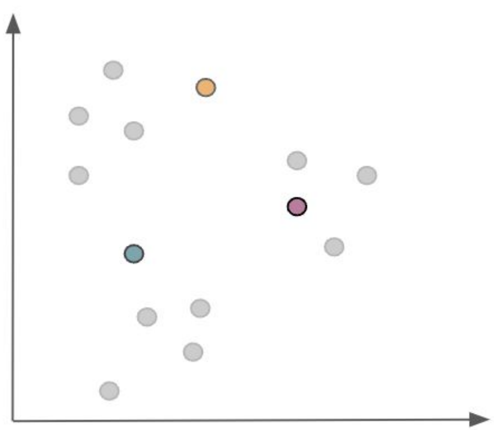


each week w is represented by the medoid m_c of the cluster c it is assigned to

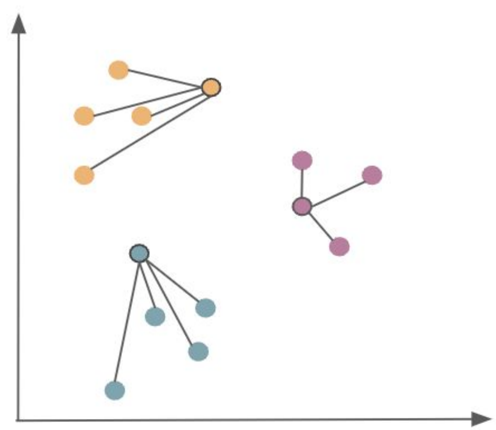
Partitioning Around Medoids (PAM)

- Partitional clustering algorithm, implemented with R-package “cluster” [4]
- Chosen because
  1. Medoids are real objects (weeks) and can be directly chosen as the representative weeks
  2. Greedy selection of starting points leads to semi-deterministic outcomes.

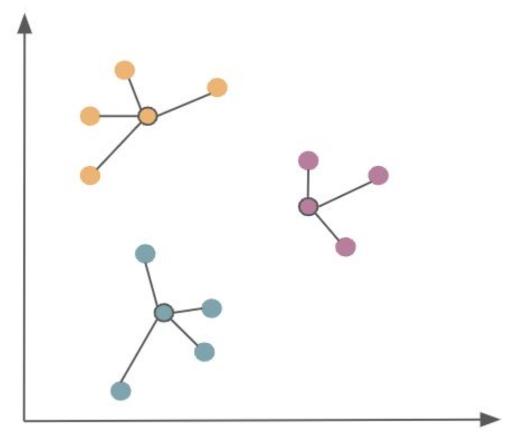
1. Greedy selection of initial medoids



2. Assign each remaining object to its nearest medoid



3. Swap medoids until total sum distances cannot be further minimised



Aalto-yliopisto
Perustieteiden korkeakoulu

[4] Martin Maechler, Peter Rousseeuw, Anja Struyf, Mia Hubert, and Kurt Hornik. cluster: Cluster Analysis Basics and Extensions, 2023a. R package version 2.1.6

Measuring Dissimilarity of Weeks

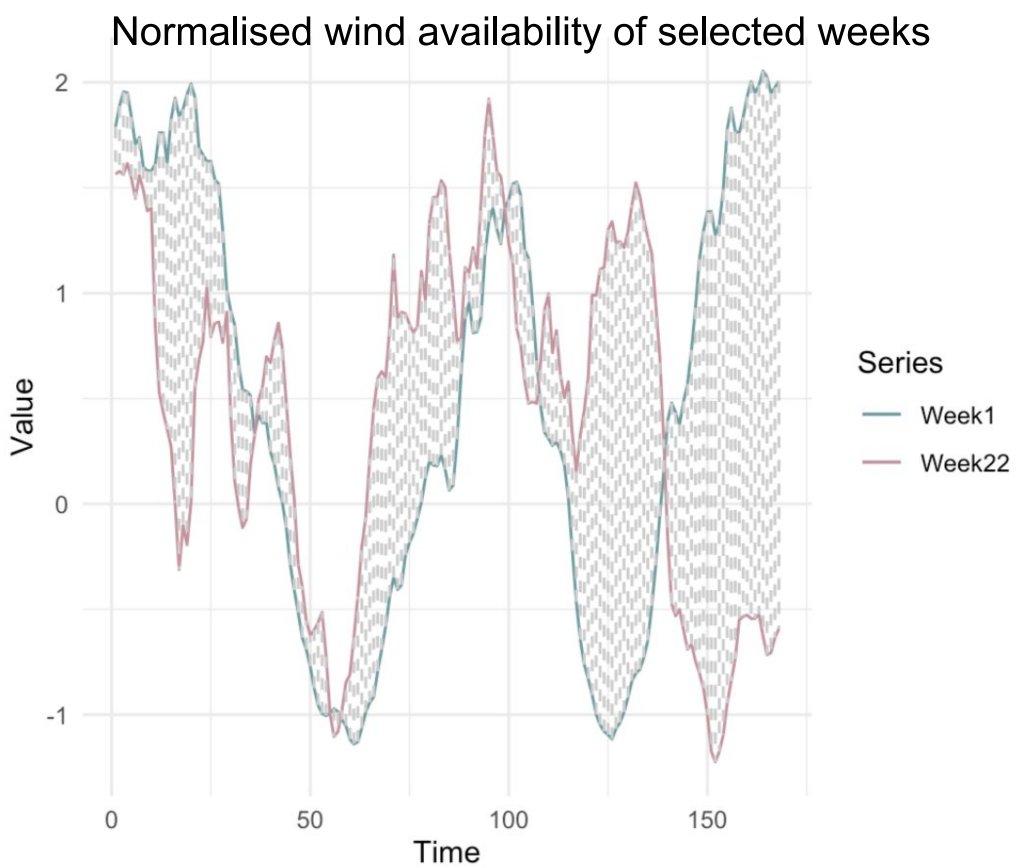
- Each time step *t* of week x is compared to the same time step of week y.
- L2 = Euclidean distance

$$d_{L2}\left(\overline{x},\overline{y}\right) = \sqrt{\sum_{t \in T} \left(x_t - y_t\right)^2}$$

- L1 = Manhattan distance

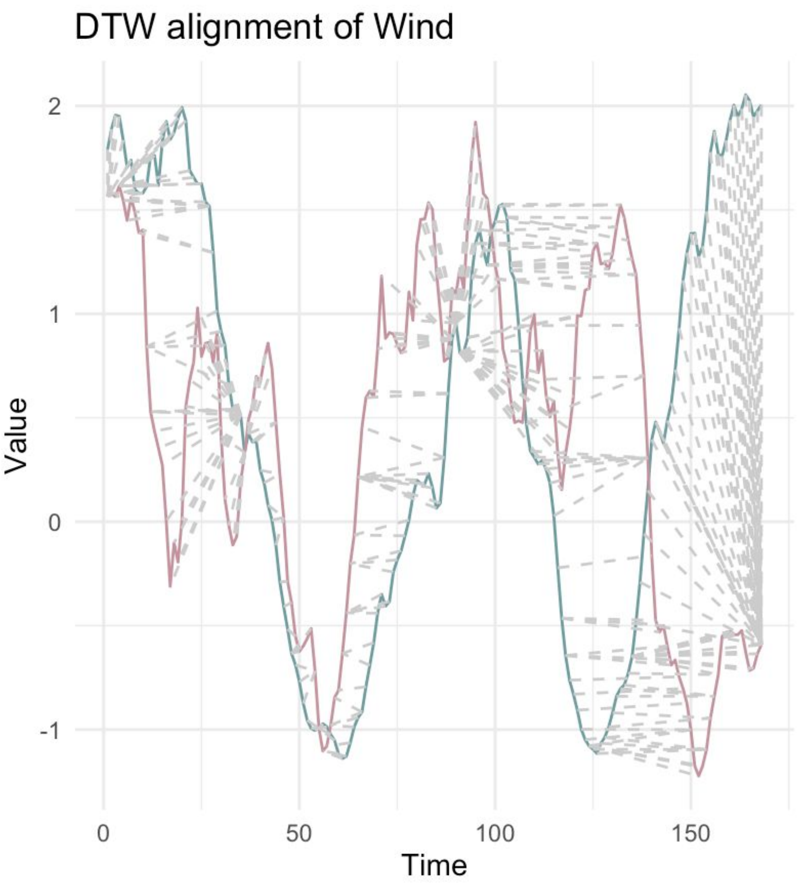
$$d_{L1}\left(\overline{x},\overline{y}\right) = \sum_{t \in T} \left|x_t - y_t\right|$$

$$T = \{\mathbb{Z} \cap [1,168]\}$$



Alignments

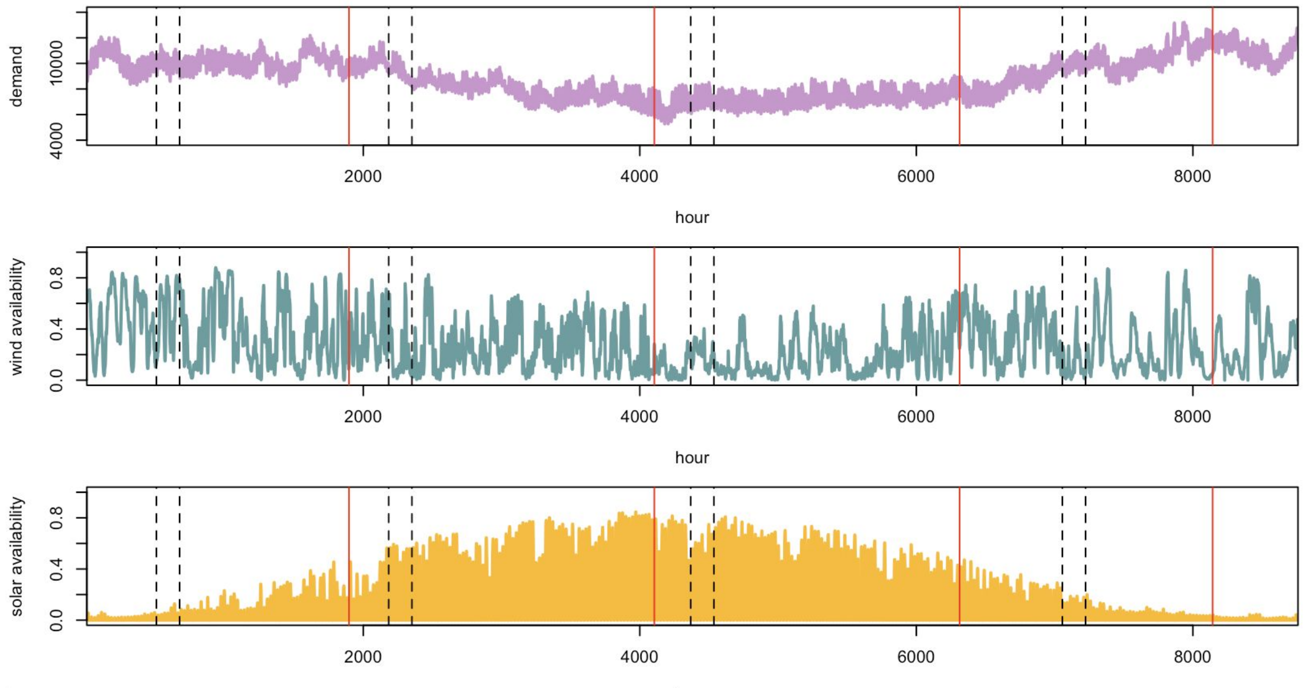
- In addition to pointwise comparison of time steps, Dynamic Time Warping (DTW) is also considered
  - Implemented with R-package "dtw" [5]
- Wind production is very stochastic
  - The peaks in generation do not necessarily align like for demand and solar
- Tests were done both and without a constraint on how far time can be warped
  - Itakura parallelogram defines a window for time warping [5]



[5] Toni Giorgino (2009). *Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package*. Journal of Statistical Software, 31(7), 1-24

Seasonal Approach

- Representative weeks are chosen from predetermined equal-sized seasons
- Implemented with a modified version of the GAMS script used in [6]

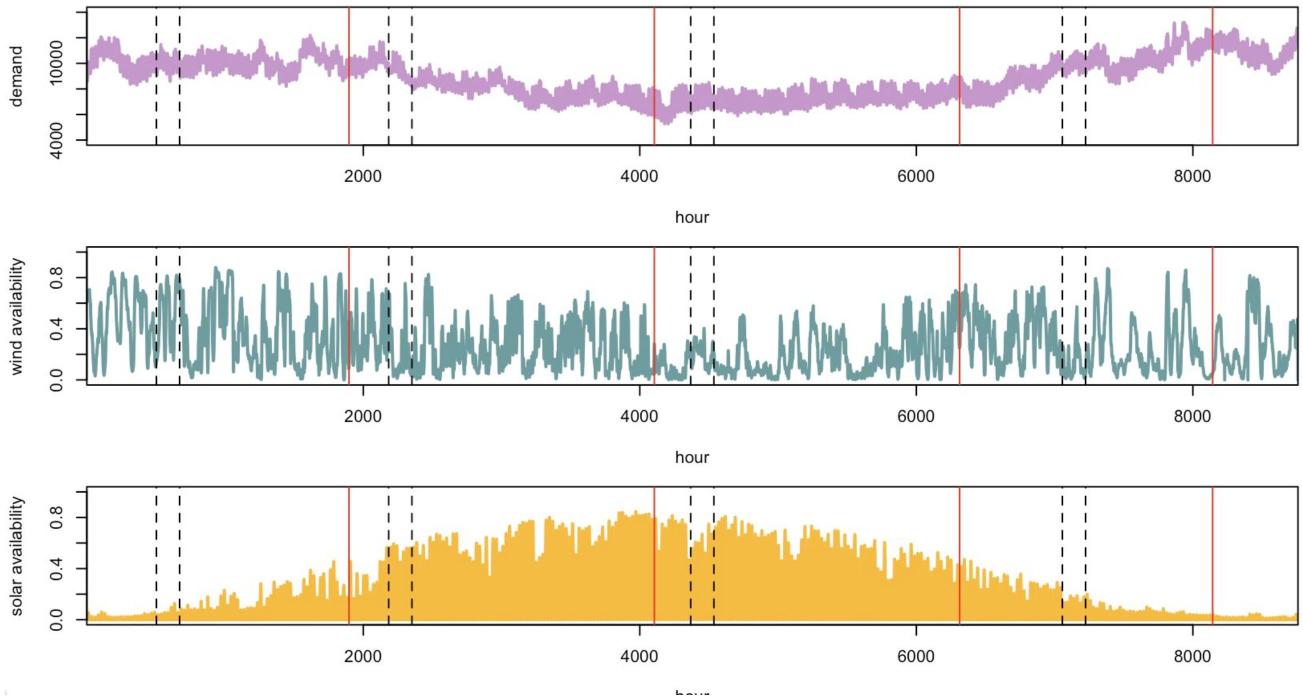


Aalto-yliopisto
Perustieteiden korkeakoulu

[6] Farzad Hassanzadeh Moghimi, Hanna Ek Fälth, Lina Reichenberg, and Afzal S. Siddiqui. Climate policy and strategic operations in a hydro-thermal power system. The Energy Journal, 44(5):67–94, 2023

Seasonal Approach

- Minimising the seasonal-weekly error in (1) the mean and (2) the standard error within each season
  - i.e. finding the week within each season that has mean and s.e. closest to the average over the whole season



Measuring Representativity

- Benchmark optimisation model
  - A capacity expansion model for Finland the year 2023
    - Modified from [6] and [7]
    - Implemented in GAMS
  - Minimise annual cost of operations, generation, and possible investment
    - Wind, solar and hydro generation has zero cost
    - Cost of thermal generation determined by fuel
  - Hydro reservoir creates dependency across time steps
- Selected statistical values representing the mean and the variance in the time series

[6] Farzad Hassanzadeh Moghimi, Hanna Ek Fälth, Lina Reichenberg, and Afzal S. Siddiqui. Climate policy and strategic operations in a hydro-thermal power system. The Energy Journal, 44(5):67–94, 2023
[7] Tatiana Tassi. Multi-criteria decision analysis for portfolio planning in power generation assessing economic and environmental tradeoffs in the finnish power sector.

Measuring Representativity

- Quality of representative weeks is analysed based on the error in selected statistics
  - Benchmark model - comparison to FTR
    - obj = Objective function value of benchmark model
    - W% = Share of total demand generated with wind power
  - Demand
    - μ_Q = Mean of hourly demand
    - s_Q = Standard deviation of demand
  - Wind
    - μ_W = Mean of hourly wind availability
    - s_W = Standard deviation of wind availability

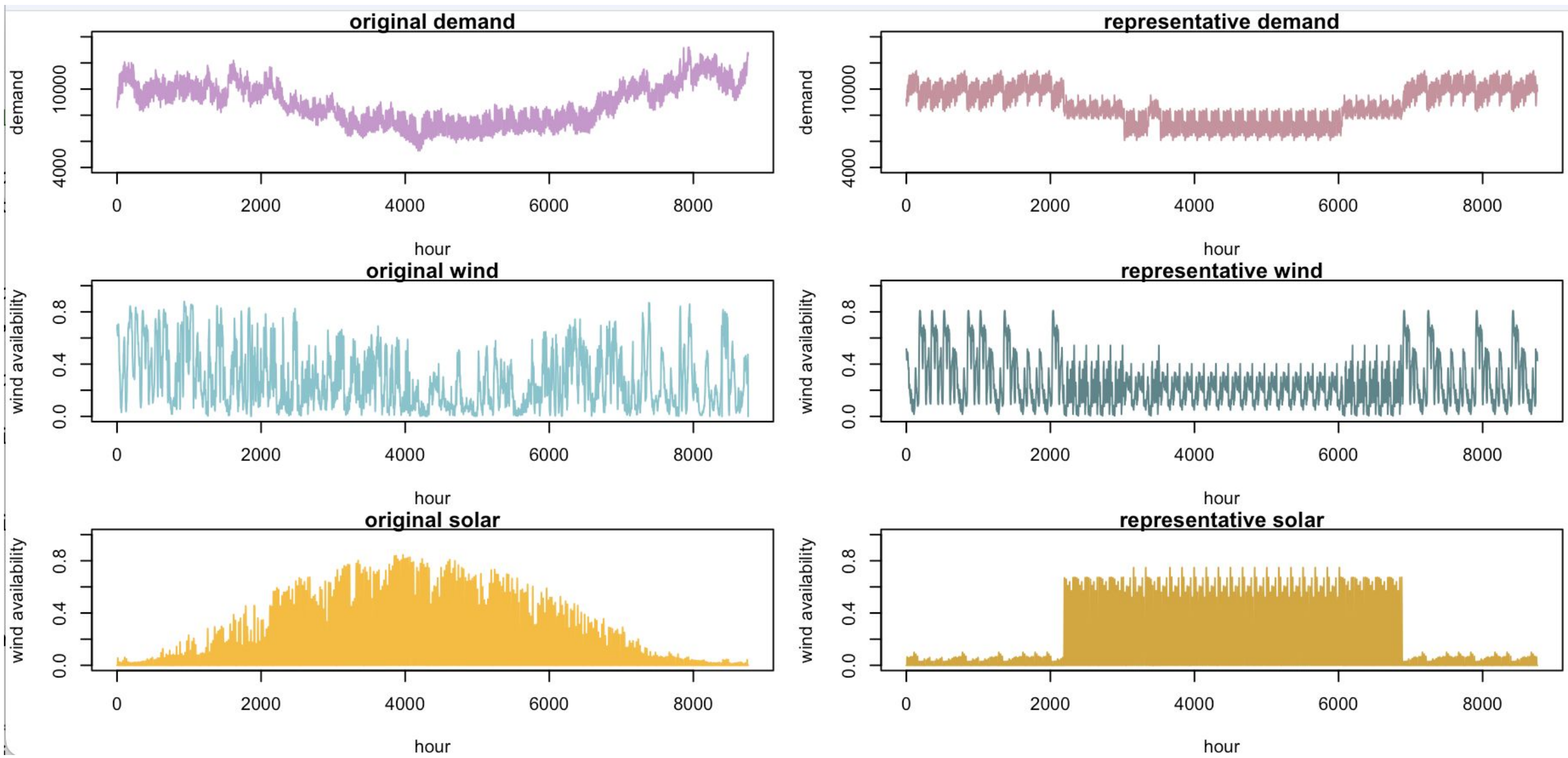
Results with the Partitional Approach

Wind + Demand	obj	W%	μ_Q	s_Q	μ_W	s_W
L2	-2,54 %	-8,12 %	-1,31 %	-7,37 %	-9,39 %	-20,47 %
L1	-0,78 %	-11,05 %	0,33 %	-2,46 %	-11,55 %	-26,51 %
DTW L1	-5,34 %	-1,61 %	-0,74 %	-0,74 %	-3,25 %	-13,95 %
DTW L1 window	-2,38 %	-6,81 %	1,12 %	-10,68 %	-6,86 %	-18,60 %
DTW L2	-5,72 %	-2,16 %	-1,09 %	-1,50 %	-4,33 %	-14,88 %
DTW L2 window	-2,07 %	-6,81 %	1,12 %	-10,68 %	-6,86 %	-18,60 %

- Objective function value lower than FTR for all variants
    - Likely due to lower share of outlier hours in representative weeks than in the original time series

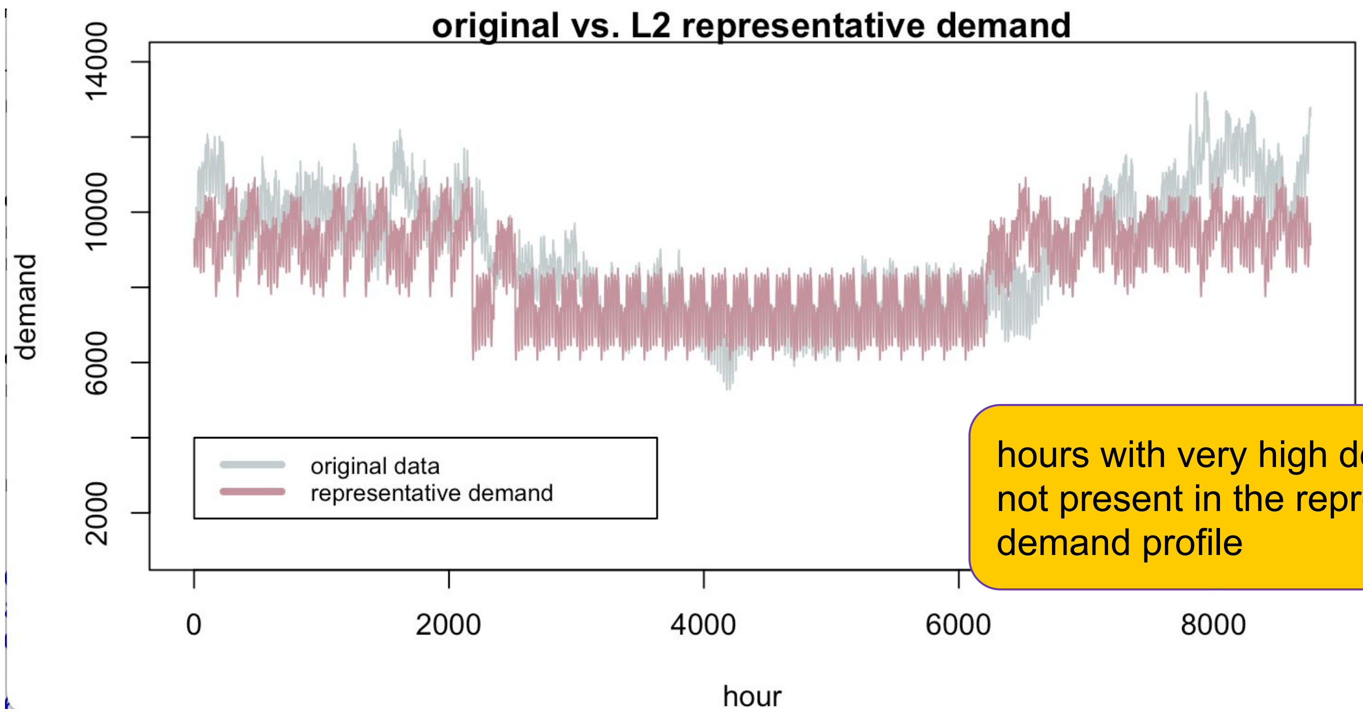
Results with L2 Distance

Wind + Demand	obj	W%	μ_Q	s_Q	μ_W	s_W
L2	-2,54 %	-8,12 %	-1,31 %	-7,37 %	-9,39 %	-20,47 %



Results with L2 distance

Wind + Demand	obj	W%	μ_Q	s_Q	μ_W	s_W
L2	-2,54 %	-8,12 %	-1,31 %	-7,37 %	-9,39 %	-20,47 %



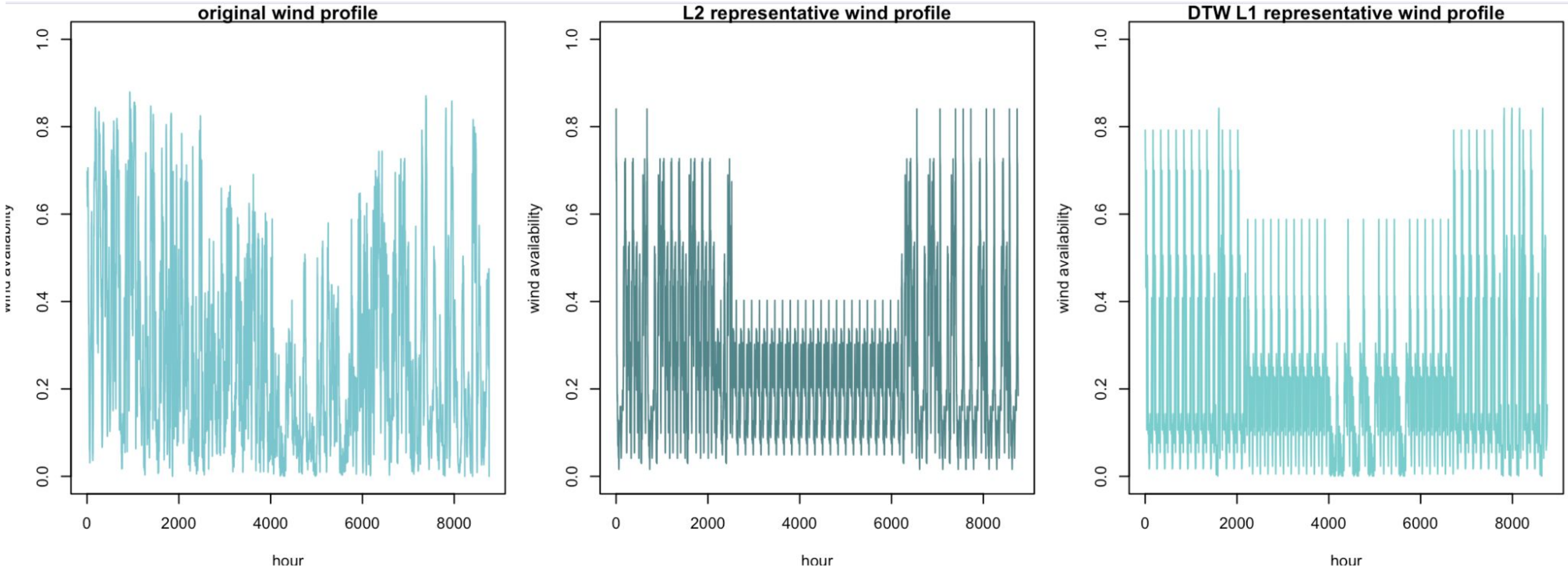
hours with very high demand are not present in the representative demand profile

Results with the Partitional Approach

Wind + Demand	obj	W%	μ_Q	s_Q	μ_W	s_W
L2	-2,54 %	-8,12 %	-1,31 %	-7,37 %	-9,39 %	-20,47 %
L1	-0,78 %	-11,05 %	0,33 %	-2,46 %	-11,55 %	-26,51 %
DTW L1	-5,34 %	-1,61 %	-0,74 %	-0,74 %	-3,25 %	-13,95 %
DTW L1 window	-2,38 %	-6,81 %	1,12 %	-10,68 %	-6,86 %	-18,60 %
DTW L2	-5,72 %	-2,16 %	-1,09 %	-1,50 %	-4,33 %	-14,88 %
DTW L2 window	-2,07 %	-6,81 %	1,12 %	-10,68 %	-6,86 %	-18,60 %

- Mean demand tracked well with all variants
  - Standard deviation however is lower
- Wind mean and generation lower than expected
  - DTW alignments perform better in this regard

Representative Wind Profiles



Wind + Demand	obj	W%	μ_Q	s_Q	μ_W	s_W
L2	-2,54 %	-8,12 %	-1,31 %	-7,37 %	-9,39 %	-20,47 %
DTW L1	-5,34 %	-1,61 %	-0,74 %	-0,74 %	-3,25 %	-13,95 %

Results with the Partitional Approach

Wind + Demand + Solar	obj	W%	μ_Q	s_Q	μ_W	s_W
L2	-4,37 %	-2,31 %	-2,49 %	-5,65 %	-5,05 %	-15,81 %
L1	-1,80 %	-8,61 %	-0,49 %	-5,01 %	-9,03 %	-14,42 %
DTW L1	-1,96 %	-0,59 %	-0,66 %	-2,38 %	-2,17 %	-1,40 %
DTW L1 window	-3,03 %	-0,67 %	-1,28 %	-3,87 %	-2,89 %	-13,02 %
DTW L2	-0,33 %	-11,93 %	-0,98 %	-2,63 %	-13,72 %	-6,98 %
DTW L2 window	-3,03 %	-0,67 %	-1,28 %	-3,87 %	-2,89 %	-13,02 %

- After adding solar
  - Error in the objective function value decreases with the unconstrained DTW alignments and increases with the others

Results with the Seasonal Approach

- Performs well in terms of the variable means
  - Expected as it is a directly in the minimised objective criteria
- However, the error in the objective function value is higher
  - Likely due to the lower obtained variance in the demand
  - Adding a higher weight to consumption (Q) leads to some improvement

Variables	obj	W%	μ_Q	s_Q	μ_W	s_W
Equal weights (WQ, WQS)	-8,14 %	+0,44 %	-0,14 %	-30,40 %	-0,72 %	-2,79 %
1,5xQ, 2xQ and 1xW (WQ)	-3,72%	+0,26%	-1,10%	-12,28%	-1,08%	-5,58%

Conclusions

- No one method clearly outperformed the others
- Both approaches dampen outliers
  - Reflected in the lower objective function values than the FTR
  - Number of representative weeks was fixed at four, which likely aggravated the effect
- Low error in the mean and standard deviation did not guarantee low error in the objective function or vice versa

References

Maximilian Hoffmann, Leander Kotzur, Detlef Stolten, and Martin Robinius. A review on time series aggregation methods for energy system models. Energies, 13(3):641, 2020.

Holger Teichgraeber and Adam R Brandt. Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison. Applied energy, 239:1283–1293, 2019.

Lina Reichenberg, Afzal S Siddiqui, and Sonja Wogrin. Policy implications of downscaling the time dimension in power system planning models to represent variability in renewable output. Energy, 159:870–877, 2018.

Tatiana Tassi. Multi-criteria decision analysis for portfolio planning in power generation assessing economic and environmental tradeoffs in the finnish power sector.

Farzad Hassanzadeh Moghimi, Hanna Ek Fälth, Lina Reichenberg, and Afzal S. Siddiqui. Climate policy and strategic operations in a hydro-thermal power system. The Energy Journal, 44(5):67–94, 2023