Martin Densing, Yi Wan, Energy Economics Group, Paul Scherrer Institute, Switzerland

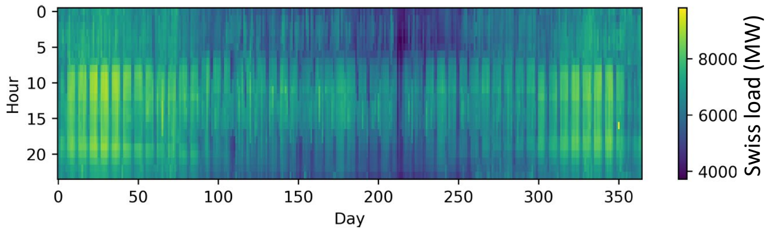
Aggregation of intermittent renewables in energy market models

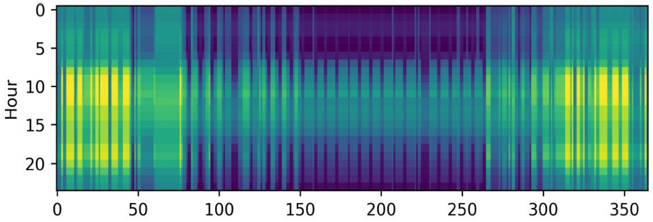
Capturing correlations and extreme events

IAEE Milano, 25. July 2023

Energy time series aggregation

- Why time series aggregation? → Numerical tractability of energy system models
- Majority of works (and ours): Aggregation on criteria “inside” input data (Hoffman,2020)
  - Minority: “Energy system structure” aggregation: Pöstges & Weber, 2019; Teichgräber et al., 2019; Wogrin, 2022
- Frequently used are clustering method: K-means, k-medoids, hierarchical clustering, with vector-norm of differences, e.g. $\sqrt{\sum_i (x_i - y_i)^2}$

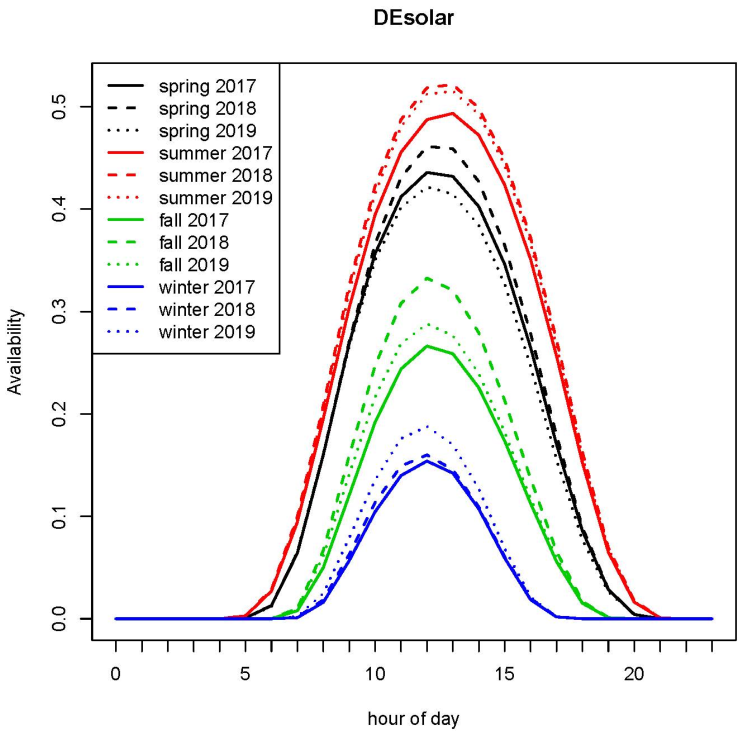


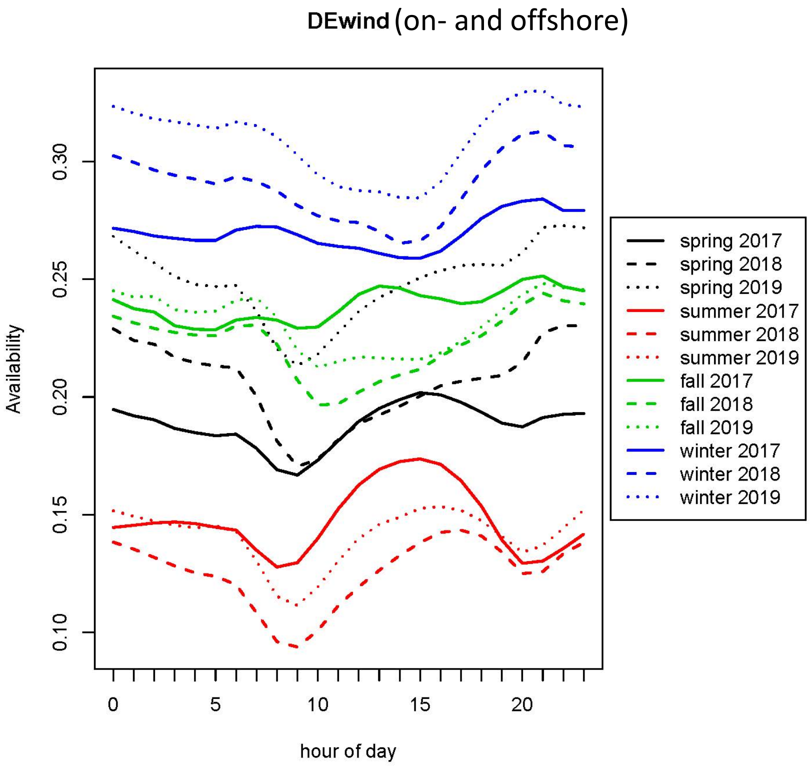


Boddu, S. (2021): (PSI MSc Thesis)

- Clustering ensures that amounts (energy, availabilities, etc.) are similar “on average”
- What about the correlations between hours of day, and to other time series (esp. wind & solar)?
- In this talk: **Capturing correlations of wind & solar availability per seasons**

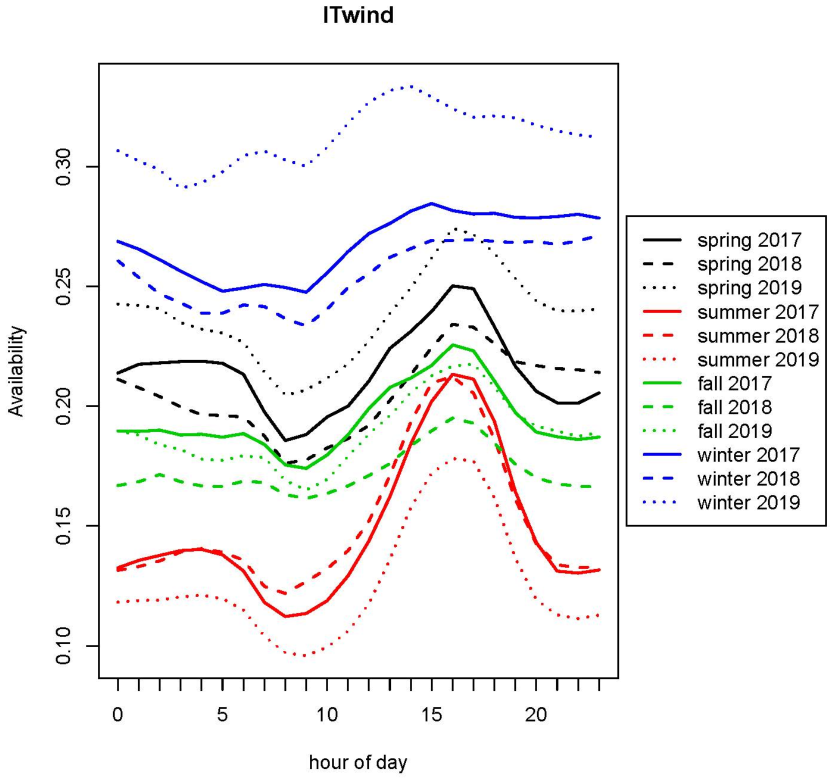
Wind & solar availability: Average days per season





- Solar-wind correlation

Different wind pattern: Example Italy



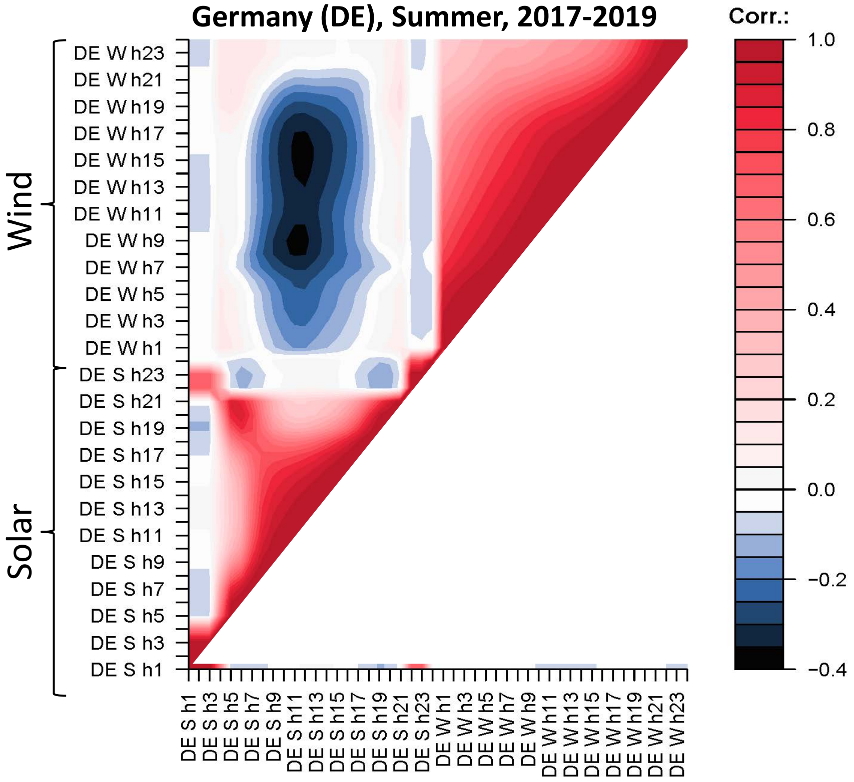
Correlation: Hourly wind & solar PV availability

Correlation wind vs. solar over all hours a year

Region	2017	2018	2019
Austria	-0.12	-0.14	-0.17
Switzerland	-0.16	-0.05	-0.12
Germany (on- and offshore)	-0.17	-0.24	-0.22
Germany offshore	-0.15	-0.21	-0.16
France	-0.15	-0.21	-0.16
Italy	-0.07	-0.09	-0.10

- Negative correlation can be higher in certain hours, up to: -0.4
- Positive correlation at
  - late-evening solar
  - late-evening wind

Correlation for each hour

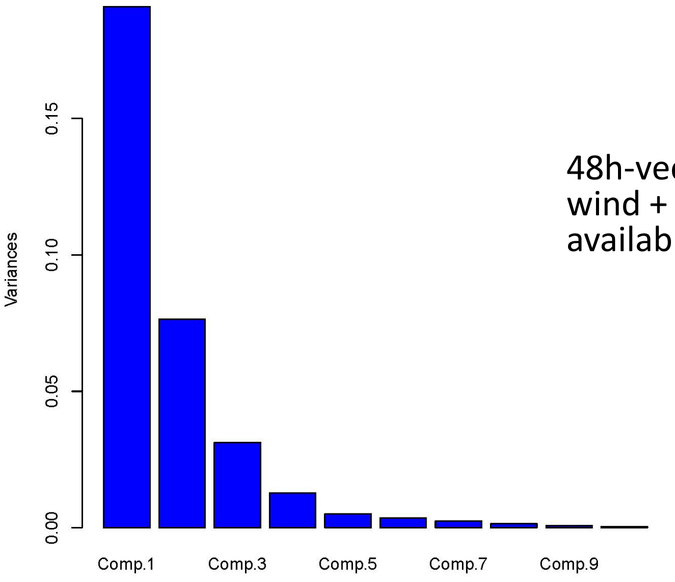


(wind: on + offshore)

Principal Component Analysis (PCA)

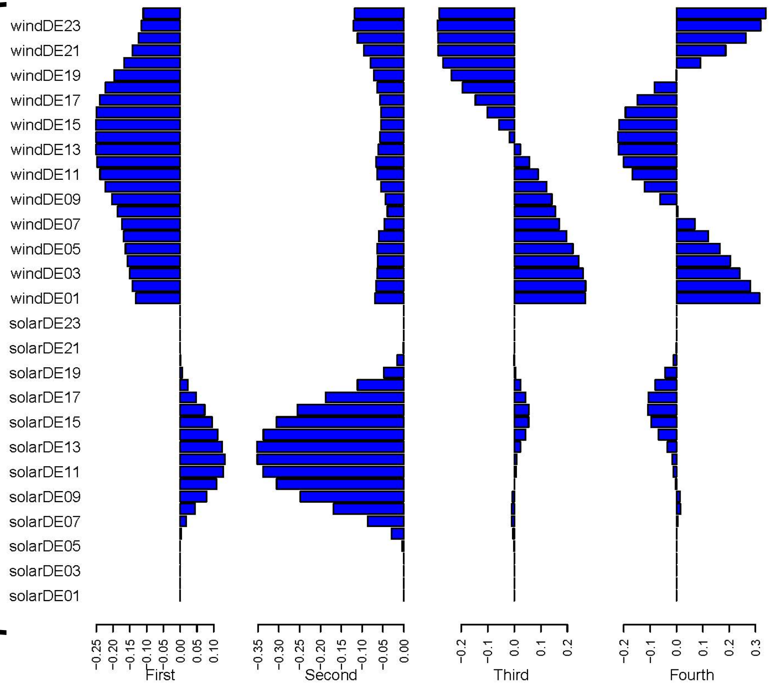
- PCA decomposes the covariances, PCA yields uncorrelated loadings

Example: Germany, summer, 2017-19



Variance of PCs ordered by variance

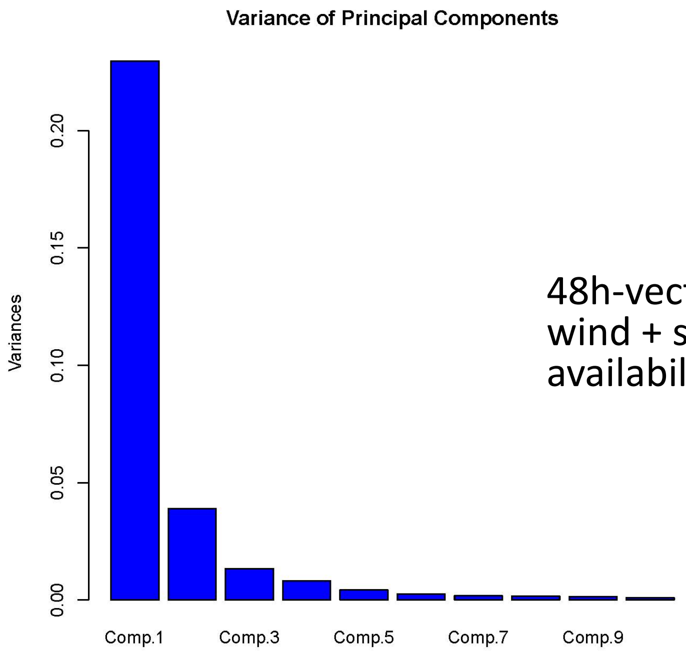
48h-vector wind + solar availability



Loadings of the ordered PCs

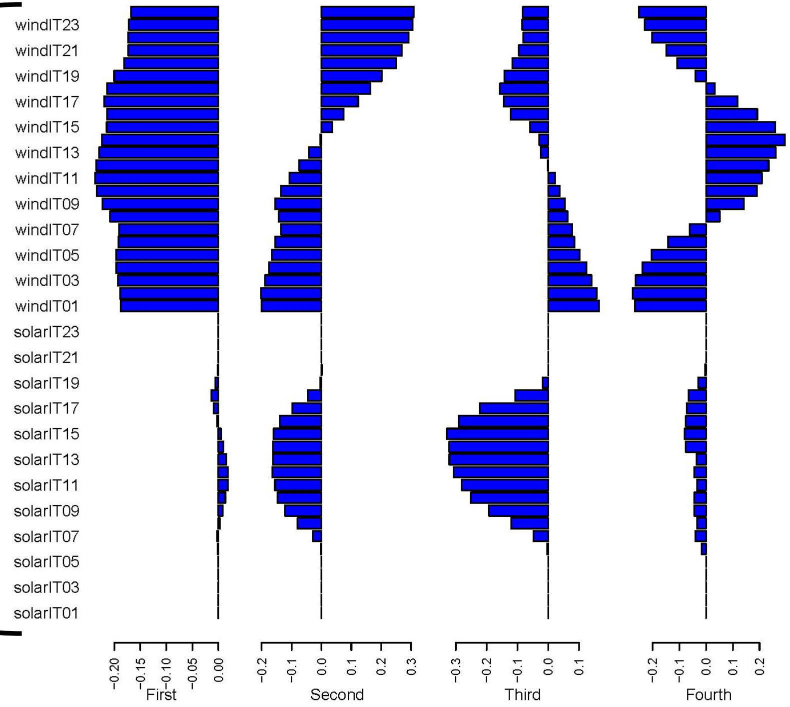
Italy: PCA of hourly wind & solar

Again: Summer, 2017-19



Variance of PCs ordered by variance

48h-vector wind + solar availability



Loadings of the ordered PCs

Scenario generation: Factor model given by PCA

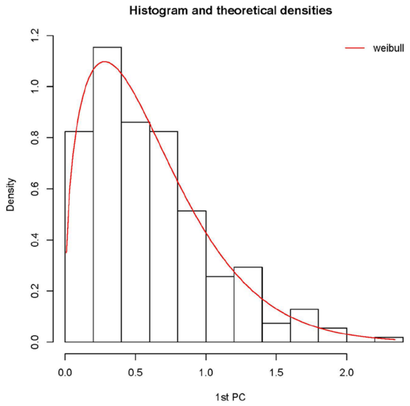
- PCA approximates covariance matrix of X by sum of uncorrelated loadings:

$$X \approx \sum_{i=1}^{k} P_i u_i, \qquad k < n$$

  - $X \in \mathbb{R}^n$: original random vector with values in n-dimensional space (n = 48),
  - $P_i \in \mathbb{R}$: random variable, i^{th} PC,
  - $u_i \in \mathbb{R}^n$: loadings of PC (deterministic vector)

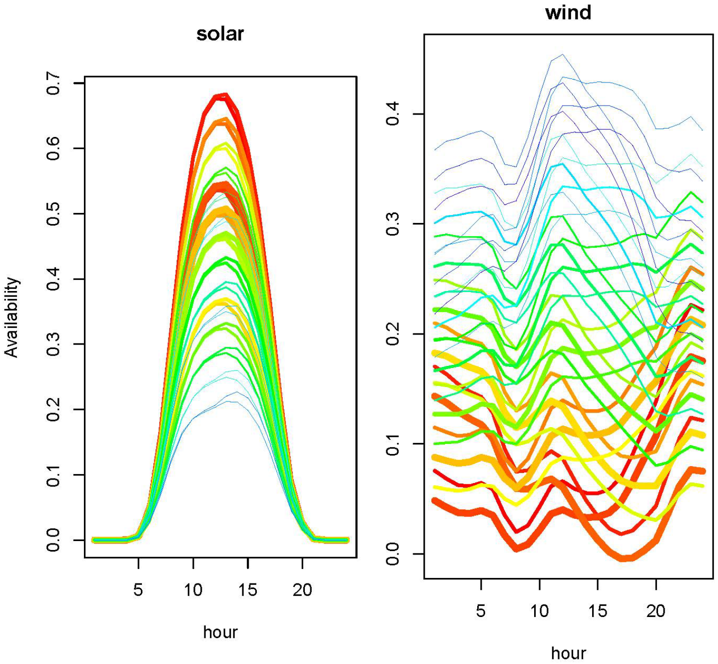
- **Factor model:** $X = UF + \varepsilon$
  - $F = (P_1, \dots, P_k)^T \in \mathbb{R}^k$: lower-dimensional factor,
  - $U = (u_1, \dots, u_n)$
  - Distribution of factors P_i are fitted by continuous distributions and then discretized:

Example: Germany, summer, P_1



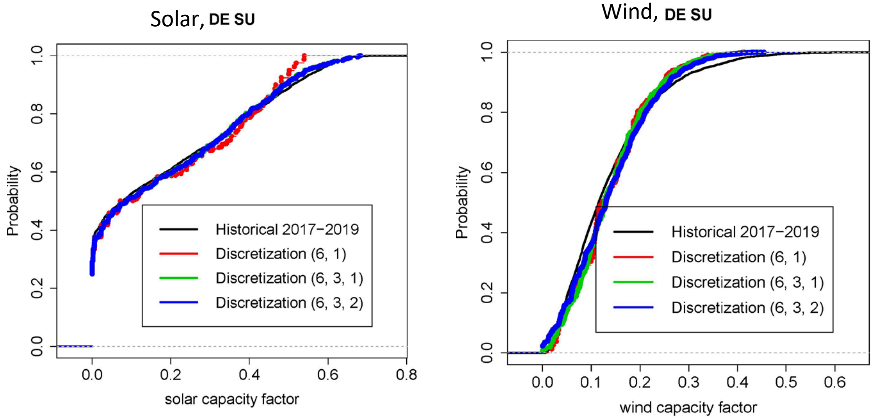
Scenarios; Quality

Example: Germany, summer; 36 scenarios;
line width = probability weight



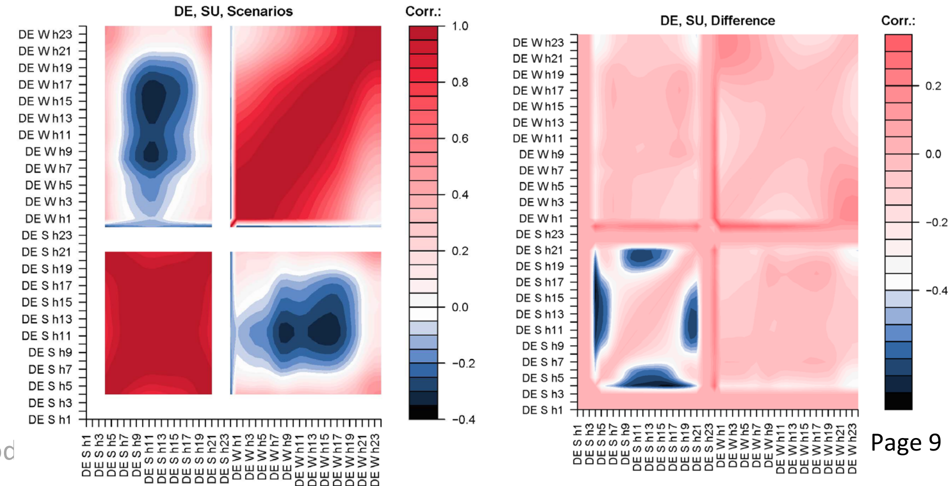
- number of components and discretizations:
  1st, 2nd, 3rd PC = 6, 3, 2

Duration curves



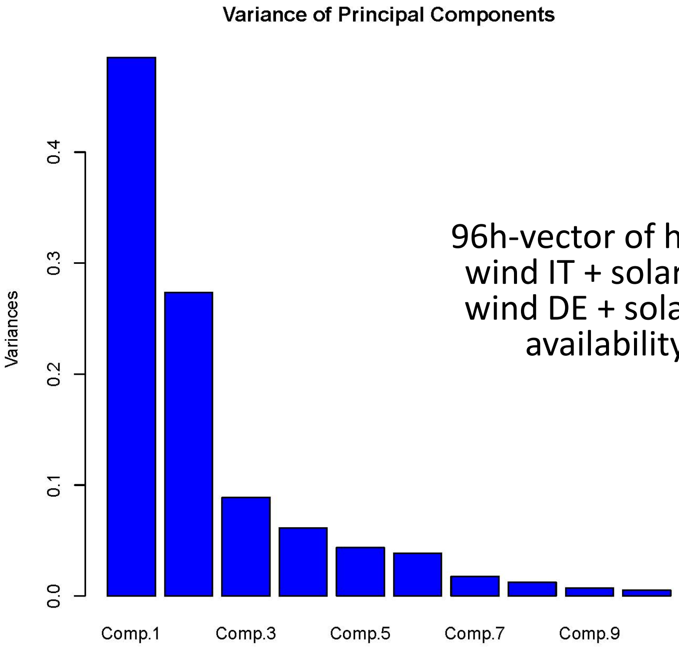
Correlation of scenarios

Correlation: error



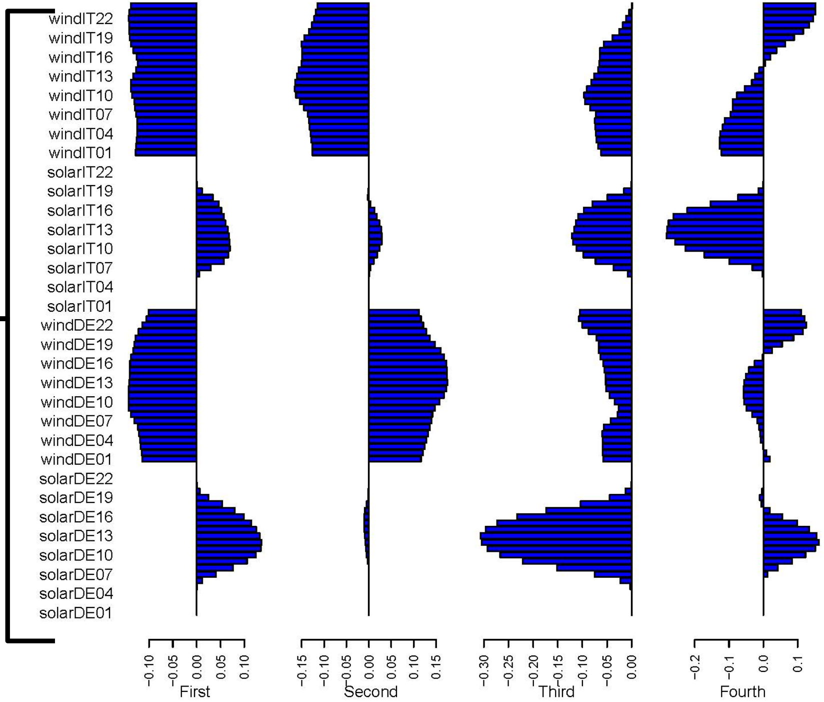
Several countries: PCA over two countries?

Example: Italy + Germany



Variance of the PCs having highest variance

96h-vector of hourly wind IT + solar IT + wind DE + solar DE availability

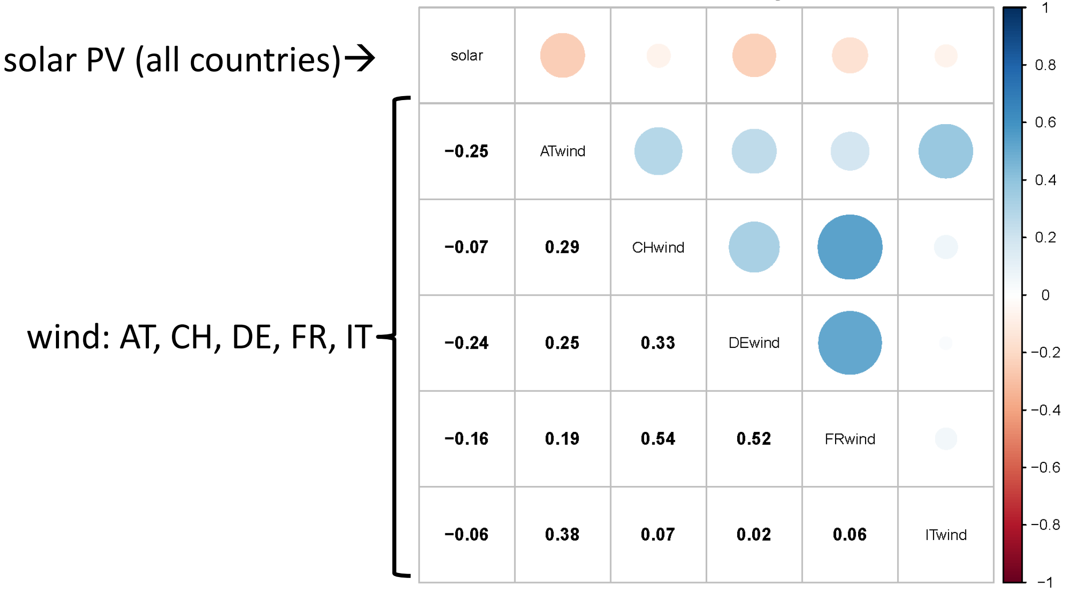


Loadings of the ordered PCs

Across regions: Daily wind & solar availability

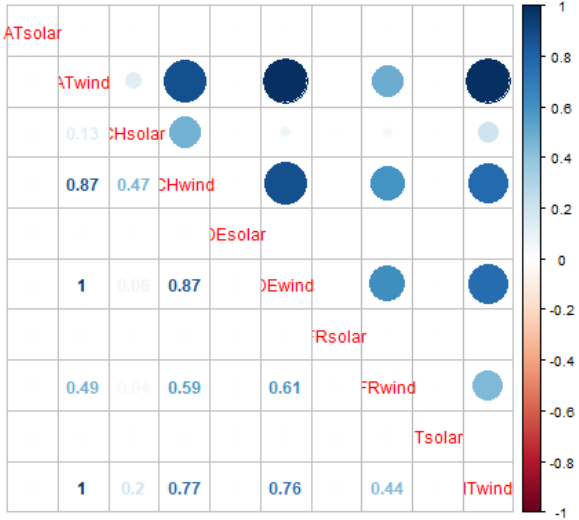
- Regions: Switzerland and surrounding countries: CH, AT, DE, FR, IT
- Keep dimension low (i): Cross-regional correlation between daily availability (avg. of hourly)
- Keep dimension low (ii): By statistical analysis: If sun is shining, then usually in all countries

Correlation matrix daily wind and solar



Tail-dependence of wind & solar across regions

- Tail dependence := Probability of joint, extremely-high values (or extremely low values)
- Daily wind & solar availability across regions: High tail-dependence = 0; Low tail-dependence =



Within a day:

- Likely: Joint calms across regions
- Unlikely: Joint storms, or dark- & calmness

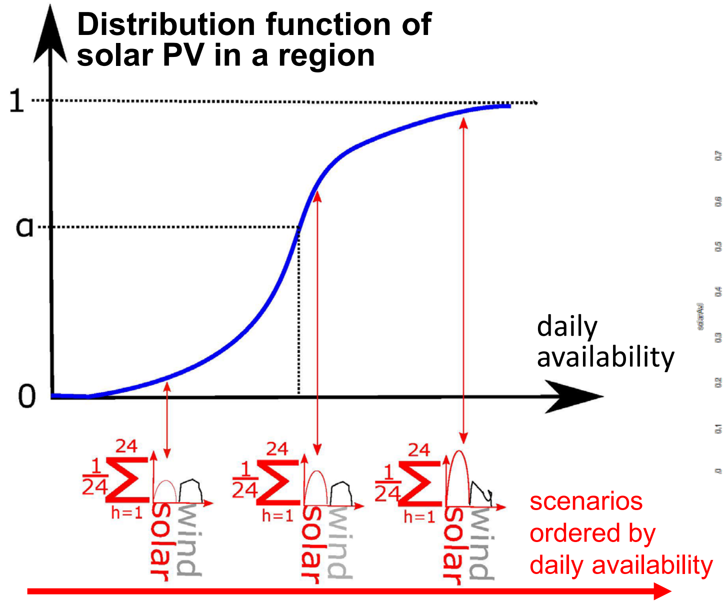
- Scenario generation: Random sampling from multivariate distribution of the variables
- Estimation of distribution? Gaussian has tail-dependences = 0. We use: t-distribution

Random sampling of copula of t-distribution

- Copulas in spatial energy time series: see e.g. Zhan et al. (2019), Camal et al. (2019)
- A random sample of a copula are quantiles of its marginal distributions. Two steps:
  (i) Sample quantile α for daily wind, solar, for each region (-> daily values across regions)
  (ii) Identify with hourly scenario having closest quantile α (ordered by daily values)

(i) Random sample of copula for solar in region

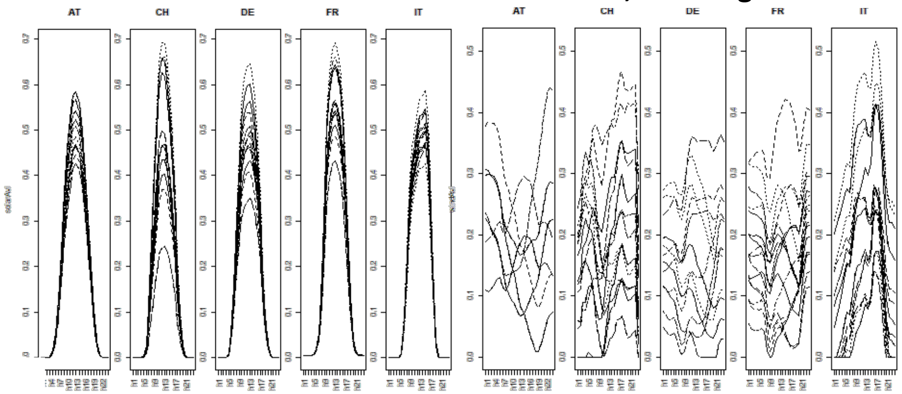
(ii) Identification with hourly scenario in region



20 random samples from t-copula → 20 scenarios

solar, each region

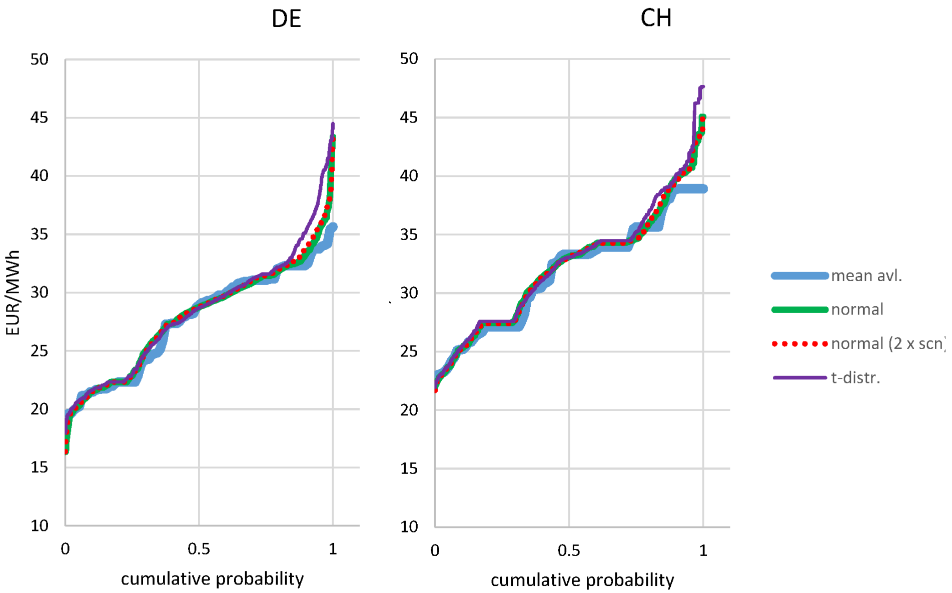
wind, each region



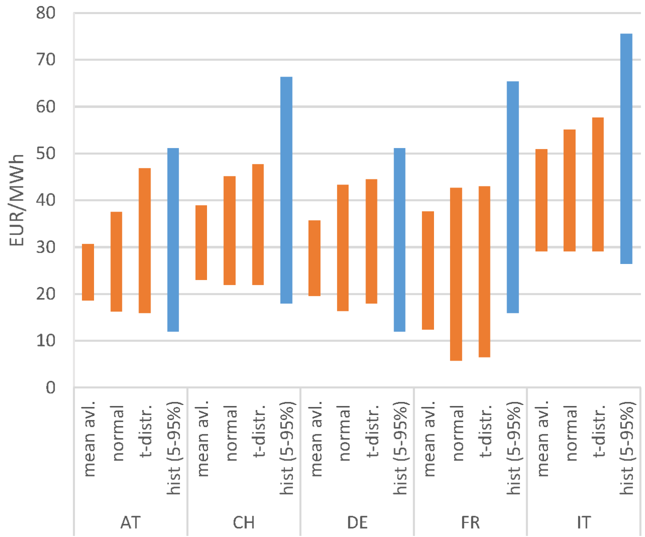
Results in an electricity market model

- BEM: Cross-border electricity market model: Switzerland and surrounding countries (Panos & Densing, 2019)
- BEM is run for this work in “basic” marginal-cost mode (price-peaks in model too low)

Price duration curves



Price ranges



PAUL SCHERRER INSTITUT PSI

Conclusions

- To capture dependences between time series of renewable supply is difficult:
  - Correlation and extreme events can be captured with daily inter-regional resolution and hourly intra-regional resolution: Fat-tail copulas and PCA

- **Limitation:** To match correlations, we need several (statistical) representative days per season: Suitable for daily or seasonal storage, but not (yet) for consecutive days; dimension goes up!

- Why **not** use the original 8760h model?
  - Numerical intractability

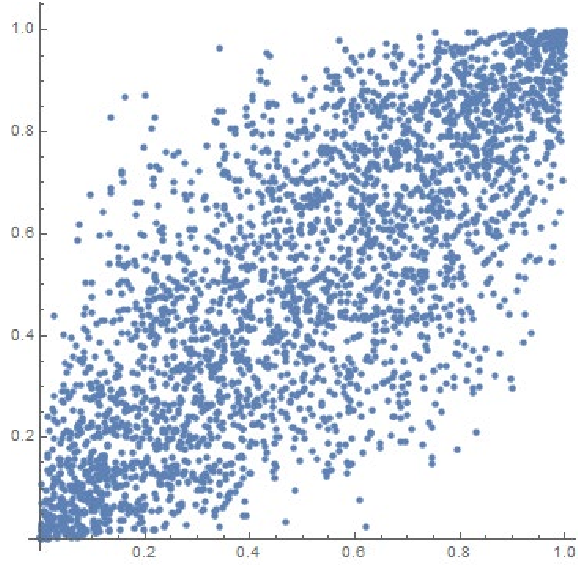
- Why use the original 8760h model?
  - Dependencies are trivially captured
  - Energy modelers are not meteorologists

- Densing & Wan, 2022. Low-dimensional scenario... *Applied Energy*. 10.1016/j.apenergy.2021.118075
- R-package: https://gitlab.psi.ch/energy-economics-group/representative-days.

Correlation is not enough

Copula: Multivariate random variable, values in [0,1], to capture only interdependencies

Random samples of
Bivariate **Gaussian (normal) copula**,
corr = 80%



Random samples of
Bivariate **t-copula**
corr=80%