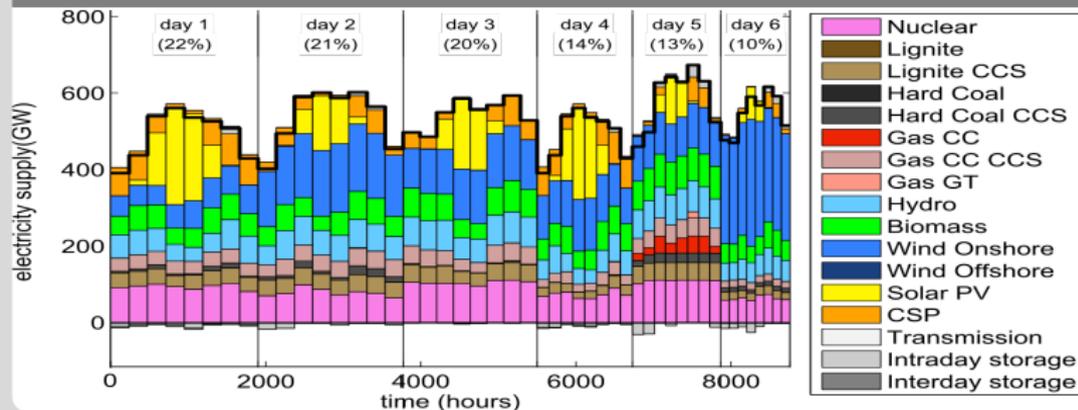


# Verkürzung von Erzeugungs- und Verbrauchszeitreihen unter Beibehaltung der energiebezogenen Merkmale

Sinan Birbalta

IIP & ITI



# Table of contents

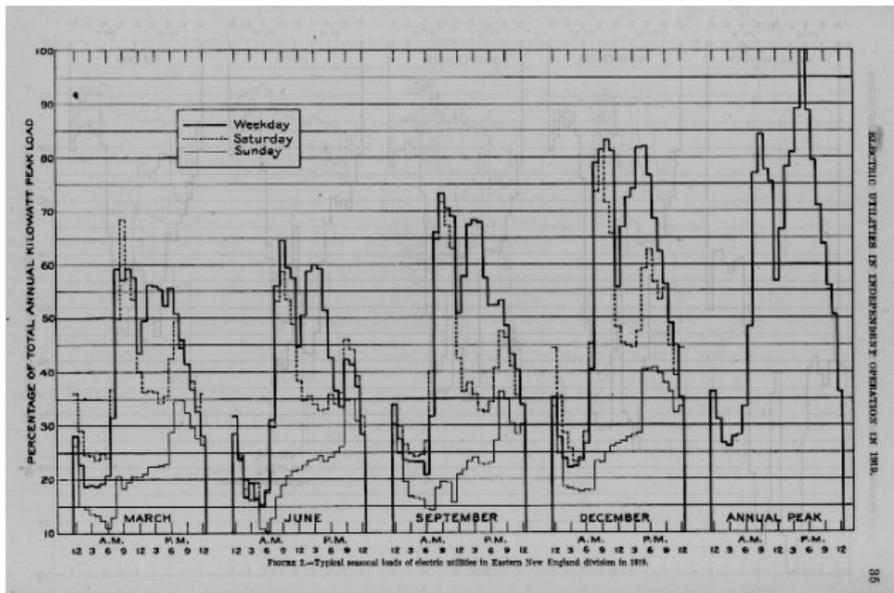
1. Einführung
2. Übersicht
3. Downsampling
4. Heuristik
5. Cluster
6. Mixed integer linear programming
7. Bewertung und Fazit

- Modell des Energiesystems einzelner Länder oder Regionen (EU)
- Optimierung des Ausbaus von Kraftwerken und der Netze
- zu jedem Zeitpunkt muss gelten:  
nachgefragte Strommenge  $\stackrel{!}{=} \text{eingspeiste Menge}$
- Berücksichtigung von Wetterdaten, technischen Kennwerten, Nachfrage, etc.
- Zielfunktion: Minimierung der Gesamtkosten,  $CO_2$ -Kosten, etc.

- Ausgangslage entweder
  - bestehendes System (energy-system optimization models) oder
  - von Grund auf (generation expansion planning models)
- Problem der Abregelung (curtailment) und des Redispatch
  - falls zu viel Energie vorhanden oder
  - falls Leitungskapazität nicht ausreichend

# Was sind repräsentative Zeitreihen?

- Frühere Ansätze waren Schätzung des typischen Verbrauchers
- 9 Typtage für einen Wochentag, Samstag, Sonntag im Winter, Sommer und Übergangszeit



**Abbildung:** Quelle: United States Geological Survey, 1921

# Warum repräsentative Zeitreihen?

- Ziel: Reduktion der Rechenzeit für Energie-(Optimierungs-)Modelle
- Früher lagen Daten nur in 15-minütiger Auflösung für große Verbraucher vor
- Außerdem teilweise Wahl der Zeitreihen "nach Gefühl" oder Standardlastprofile der BDEW
- Heute immer mehr, unflexible Produzenten
- Demand Site Management

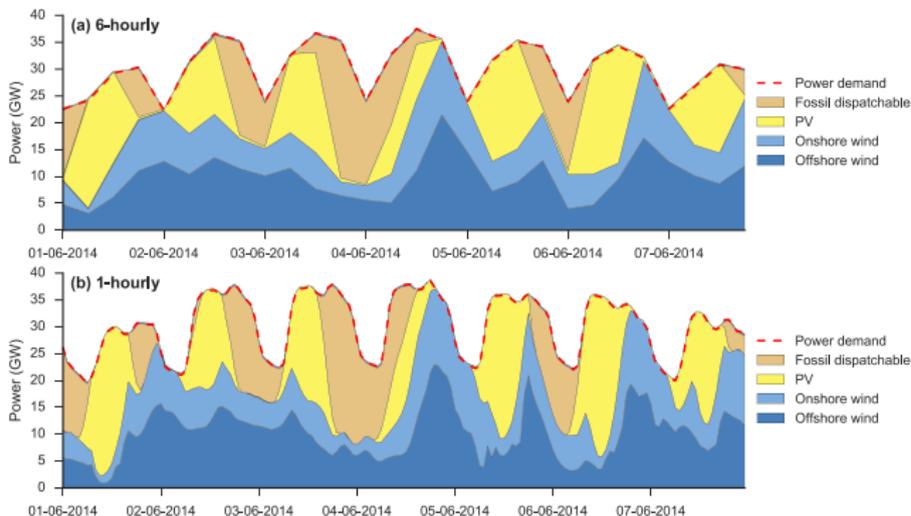
Methode	Wichtige Parameter
Randomisiert	Anzahl zufälliger Ziehungen
Downsampling	Auflösung/Fenstergröße
Heuristik	Auswahlmethode; Resampling oder auslassen
K-means Cluster [NSHK16]	Distanz definieren; Anzahl Cluster;
Hierarchische Cluster [NSHK16]	Cluster-Zentren oder nächstgelegene Tage
Linear Programming [PHD <sup>+</sup> 17]	Anzahl repräsentativer Tage und Auflösung
Hybrid [PHD <sup>+</sup> 17]	Anzahl repräsentativer Tage und Auflösung

- Gleitender Mittelwert mit unterschiedlicher Fenstergröße

$$m(t) = \frac{1}{n} \sum_{i=0}^{n-1} x(t-i)$$

- führt zu Glättung der Zeitreihe

## ■ Spitzen werden "abgeschnitten"



**Abbildung:** Downsampling bei Daten im Zeitraum einer Woche [Pfe17]

## ■ starke Abweichungen für PV, Wind weniger aber auch betroffen

- Tage mit scheinbar wichtiger Charakteristik werden gewählt
  - unterschiedliche Wetter- oder Lastbedingungen
  - Tage mit minimaler und maximaler Nachfrage
  - aus jedem Monat oder jeder Jahreszeit
- kombinierbar mit Downsampling

- seltene Muster werden vernachlässigt
  - weil übersehen oder
  - weil Auswahl zu Überrepräsentation führen würde
- Repräsentiert entweder Wind oder PV gut, nicht beides

- Vergleich von Tagesdauerkurven nach
  - Varianz
  - Korrelation untereinander
  - Approximation der Jahresdauerlinie (RMSE)
- Wahl von repräsentativen Zeitreihen pro Cluster
  - Clusterzentrum oder
  - Zeitreihe, die am dichtesten am Zentrum liegt
- "Muster" werden gewichtet nach Größe des Clusters
- seltene Lastverhalten werden nicht vernachlässigt, wie bei Heuristik

- Zeitreihen einzelner Tage werden betrachtet
- Betrachtung von k-means und hierarchisch nach Ward
- Vorgabe gewünschter
  - Cluster-Anzahl (k-Means)
  - Cluster-Größe (hierarchisch)
- Vorteil hierarchischer Algorithmus: deterministisch
- Ziel: Minimierung des Abstands der Clusterzentren zu den Mitgliedern
- Zurodnung nach Distanz von Datenpunkten, hier Tage oder Stunden
- Metriken: Korrelation, Error oder andere Werte

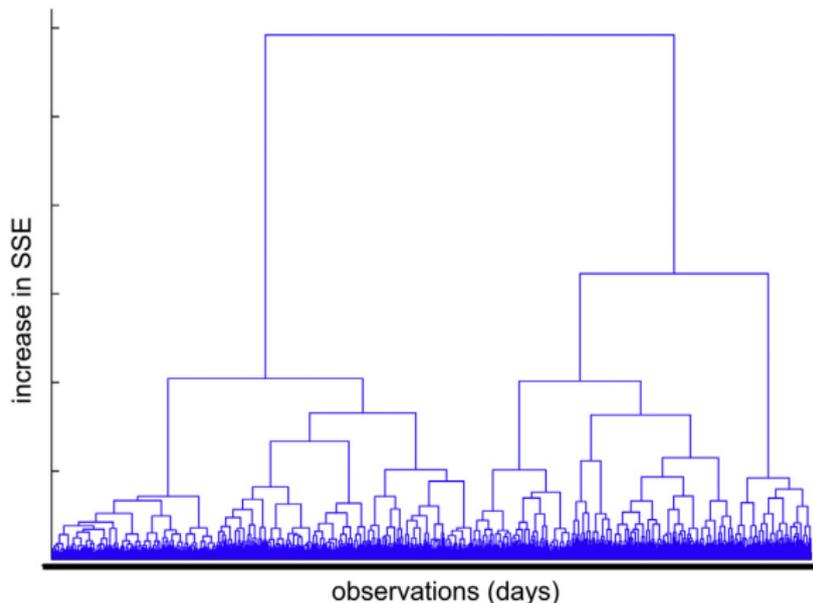
0.  $n$ -dimensionalen Vektor für jeden Tag erstellen,  
 $n = \text{Anzahl Zeitschritte pro Tag} \cdot \text{Anzahl Zeitreihen}$

## 1 Normalisierung der Zeitreihen

- Erzeugung nach maximalem Output pro Technologie (global)
- Nachfrage nach Maximum in jeweiliger geographischer Lage

## 2 Cluster-Algorithmus anwenden

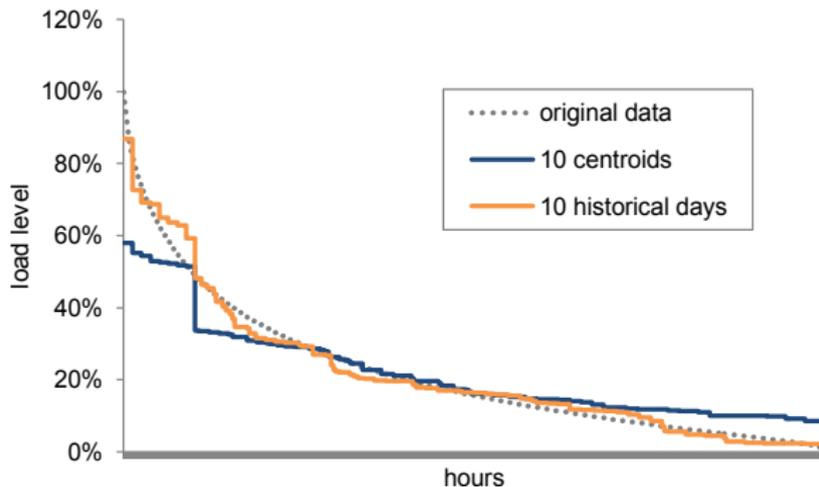
- k-means:  $k$  Anfangswerte jeweils einem Cluster zuordnen, dann iterativ die Elemente hinzufügen, die die Varianz am wenigsten erhöht (Animation)
- Ward: Verbindet iterativ in jedem Schritt zwei Cluster, sodass sich die kumulierte Varianz innerhalb der Cluster minimal erhöht



**Abbildung:** Zunahme des Fehlers bei zunehmendem Clustering [NSHK16]

$$SSE = \sum_c \sum_{d \in D_c} \|V_d - \bar{V}_c\|^2, V_d: \text{beobachteter Tag}, \bar{V}_c: \text{Cluster-Zentrum}$$

- 3 Kandidaten für Cluster auswählen
  - Optimale Anzahl bzw. Größe der Cluster ist vorher nicht bekannt
  - Folgende Schritte mit jedem Kandidatenset durchführen
- 4 Repräsentativen Tag pro Cluster wählen
  - arithmetisches Mittel oder besser:
  - beobachteter Tag, der den kleinsten Abstand zum arithmetischen Mittel hat
- 5 Jeden Tag nach Clustergröße gewichten
  - arithmetisches Mittel oder besser:
  - beobachteter Tag, der den kleinsten Abstand zum arithmetischen Mittel hat
- 6 Summe der gewichteten Tage auf 1 nivellieren

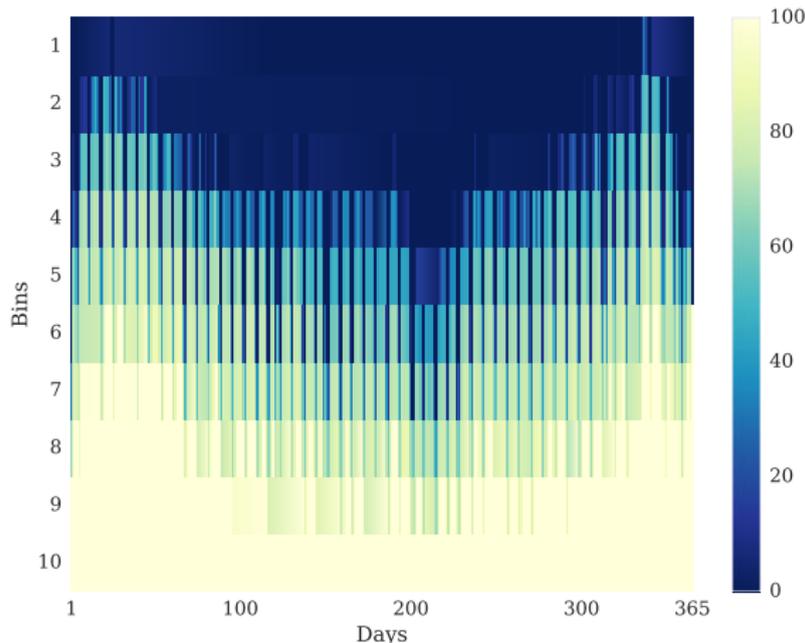


**Abbildung:** Vergleich einer Lastkurve von Onshore Wind zwischen Wahl des arithmetischen Mittels und beobachtetem Tag [NSHK16]

$$SSE = \sum_c \sum_{d \in D_c} \|V_d - \bar{V}_c\|^2, \quad V_d: \text{beobachteter Tag}, \quad \bar{V}_c: \text{Cluster-Zentrum}$$

- 5 Jeden Tag nach Clustergröße gewichten
  - arithmetisches Mittel oder besser:
  - beobachteter Tag, der den kleinsten Abstand zum arithmetischen Mittel hat
- 6 Summe der gewichteten Tage auf 1 nivellieren

- Auswahl der Tage mit Gewichtung für jeden Tag, ähnlich Clustering
- Abstand zur Dauerkurve wird minimiert (NRMSE)
- Parameter
  - Anzahl Tage  $n$
  - Anzahl Slots pro Tag  $s$
- $b = n \cdot s$  bins werden für jede Dauerkurve approximiert
- Jeder bin entspricht einem Zeitschritt
- Durch die Minimierung erhält jeder bin am Ende eine Gewichtung

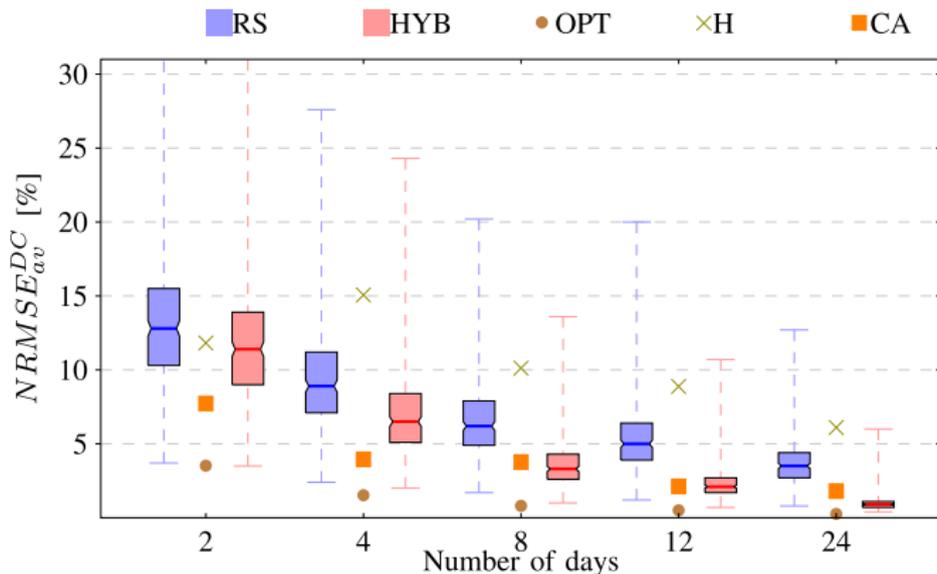


**Abbildung:** 10 Bins für eine Lastkurve in Belgien 2014, [PHD<sup>+</sup>17]

- mehrmalige Zufallswahl von Untermengen von Tagen
- dabei immer die Tage mit höchstem Fehler bei der Approximation zur Dauerkurve wegwerfen
- "feasible" für bis zu 50000 Untermengen [PHD<sup>+</sup>17]

- Kombiniert Zufallsauswahl und LP
- Gewichtungen werden optimiert, wie MILP

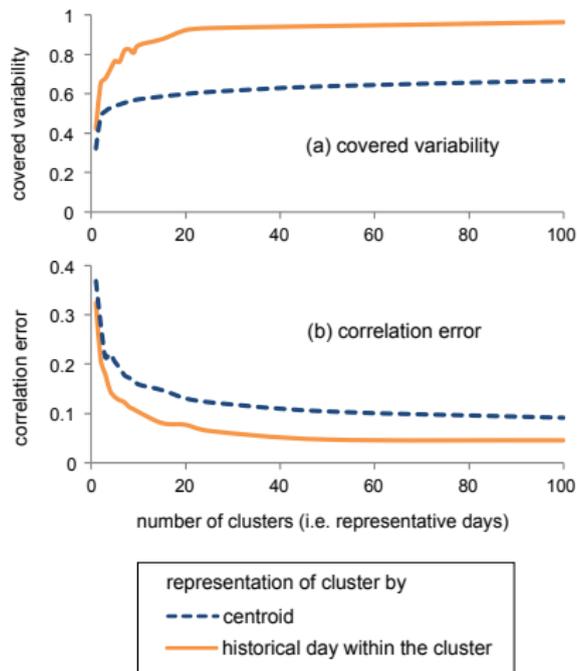
- Methoden wurden auf Energiesystemmodelle angewendet
- Verschiedene Maße wurden berechnet
  - Fehler in Approximation der Jahresdauerlinie (Normalized Root Mean Square Error)
  - Vergleich der Mittelwerte jeder Zeitreihe (Relative Energy Error)
  - Abweichung der Korrelation zwischen den Zeitreihen (Correlation Error)
- neueres Paper betrachtet nicht die *verbesserte* Cluster-Methode



**Abbildung:** Normalized Root Mean Square Error bei verschiedenen Verfahren [PHD<sup>+</sup> 17]

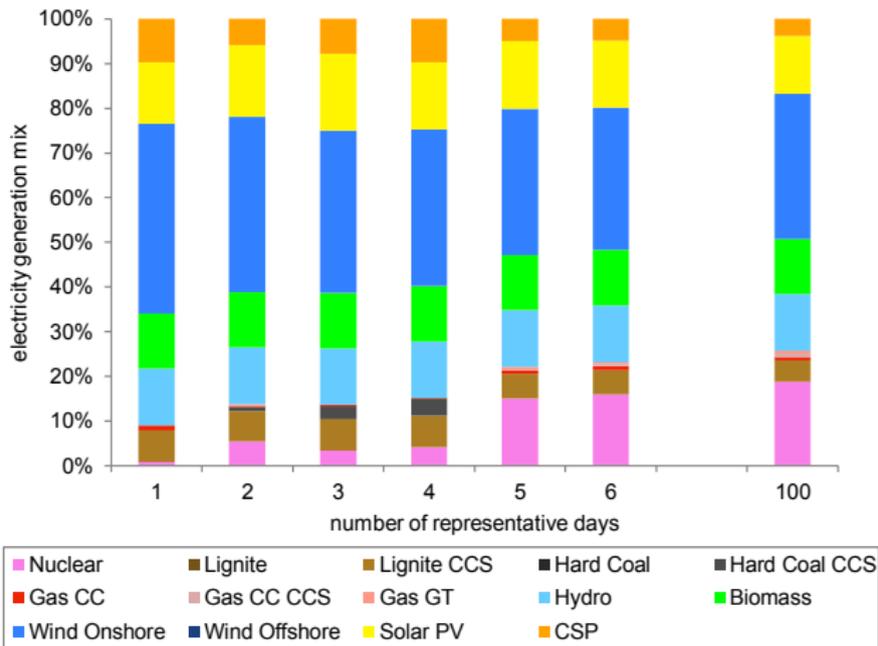
**RS:** Random Selection, **HYB:** Hybrid, **OPT:** LP, **H:** Heuristik, **CA:** Cluster Algorithmus

# Approximation der Lastkurve



**Abbildung:** Abnahme des Varianz- und Korrelations-Fehlers [NSHK16]

# Approximation der Lastkurve



**Abbildung:** Erzeugungsmix als Ergebnis eines generation expansion planning models [NSHK16]

Methoden	Aufwand	Genauigkeit
Randomisiert	+ +	-
Downsampling	+ +	-
Heuristik	+ +	- -
K-means Cluster	-	+
Hierarchische Cluster	-	+
Linear Programming	- -	+ +
Hybrid	-	+ +

- Übergreifender Vergleich der Methoden schwierig
- Neue Methoden wurden nicht an Optimierungsmodell getestet
- Mehr Zeitschritte führen zu höherer Genauigkeit
- vor allem wichtig für fluktuierende Erneuerbare
- Rechenzeit und Genauigkeit korrelieren nicht linear
- Letztendlich kommt man eher nicht um eine eigene Evaluierung herum

-  Paul Nahmmacher, Eva Schmid, Lion Hirth, and Brigitte Knopf, *Carpe diem: A novel approach to select representative days for long-term power system modeling*, Energy **112** (2016), 430–442.
-  Stefan Pfenninger, *Dealing with multiple decades of hourly wind and PV time series in energy models: A comparison of methods to reduce time resolution and the planning implications of inter-annual variability*, Applied Energy **197** (2017), 1–13.
-  Kris Poncelet, Hanspeter Hoschle, Erik Delarue, Ana Virag, and William Drhaeseleer, *Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems*, IEEE Transactions on Power Systems **32** (2017), no. 3, 1936–1948.