Reducing Complexity in RE System Optimisations

Tom Brown

NEXT ENERGY 2. Themenfeldseminar Systemmodellierung und -analyse, 22nd February 2016

Frankfurt Institute of Advanced Studies (FIAS), Goethe-Universität Frankfurt
FIAS Renewable Energy System and Network Analysis (FRESNA)

FIAS Frankfurt Institute for Advanced Studies

Table of Contents

1. Introduction
2. Example: Electricity system optimisation
3. Network reduction
4. Time series reduction
5. Improving optimisation formulation
6. Conclusions

Introduction

Goal

Given a desired CO_2 reduction, what is the most cost-effective system?

$$\mathrm{Min}\begin{pmatrix}\text{Yearly system}\\ \text{costs}\end{pmatrix} = \sum_{n}\begin{pmatrix}\text{Annualised}\\ \text{capital costs}\end{pmatrix} + \sum_{n,t}(\text{Marginal costs})$$

subject to

- meeting energy demand at each node *n* and time *t*
- wind, solar, hydro (VRE) availability $\forall n, t$
- electricity/gas transmission constraints between nodes
- (installed capacity) $\leq$ (geographical potential)
- CO_2 constraint / RE share covering demand

Interesting Questions

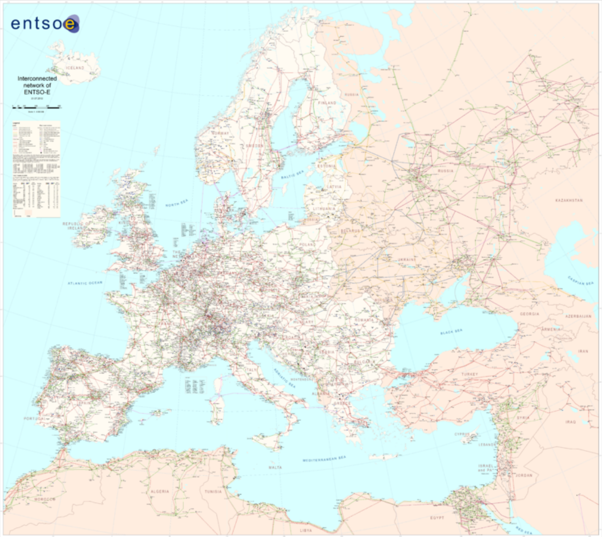
How do system characteristics and costs change as we ...

- ... restrict transmission expansion?
- ... impose import/export balance constraints?
- ... relax CO_2 constraint?
- ... include non-pumped-hydro storage?
- ... include meshed overlay direct current network and/or more phase-shifting transformers?
- ... couple electricity sector to heating/cooling and transport?

Some of these questions are driven by politics/social factors: compromises that take us away from economic optimum.

Problem 1: Spatial resolution

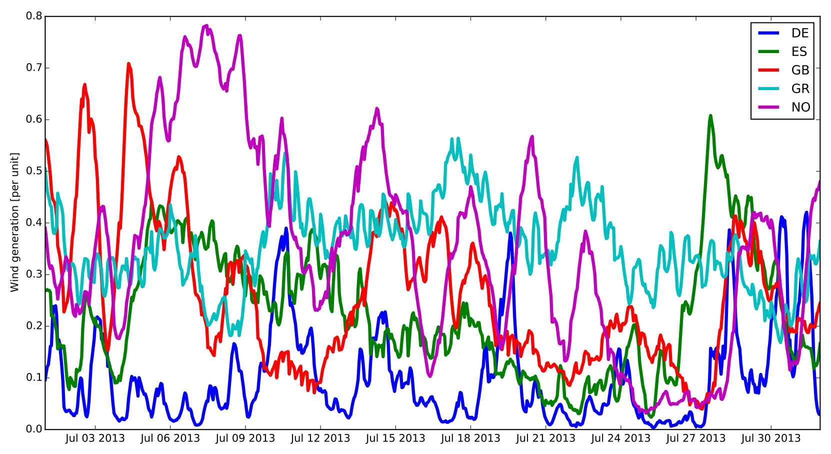
Need high spatial resolution to represent VRE resource variation and transmission constraints in electricity and gas networks.



Source: ENTSO-E

Problem 2: Temporal resolution

Need high temporal resolution to represent load and VRE resource variability and correlations. Wind generation in Europe in July 2013:



Problem 3: Model accuracy

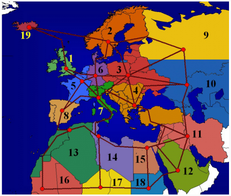
Modelling must respect physics.

- How much detail in the input data do we need?
- Optimise transmission simultaneously with generation capacity?
- How bad are linear approximations?
- Can we make the algorithms faster, to add detail in other areas?
- By looking at static situations, do we miss dynamic effects?

Examples from literature of electricity optimisation

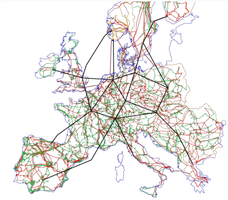
Study	Spatial resolution	Temporal resolution	What?	Flow physics
Czisch (2005)	low	high	g and t	transport
Hagspiel et al. (2014)	medium	low	g and t	linear
Egerer et al. (2014)	high	low	t only	linear

Czisch



Hagspiel et al.

Egerer et al.



Overarching goal

Find the "sweet spot" where:

- Computation time is finite (i.e. a week)
- Temporal resolution is "good enough"
- Spatial resolution is "good enough"
- Model detail is "good enough"

AND quantify the error we make by only being "good enough"

(e.g. are important metrics $\pm 10\%$ or $\pm 50\%$ correct?)

Example: Electricity system optimisation

Objective function

Objective function is total annual system cost, whose minimisation represents maximisation of "social welfare" (for fixed inelastic demand):

$$f(\bar{P}_l, \bar{g}_{n,s}, g_{n,s,t}) = \sum_l (c_l + o_l)\bar{P}_l + \sum_{n,s} c_{n,s}\bar{g}_{n,s} + \sum_{n,s,t} w_t o_{n,s} g_{n,s,t}$$

We optimise:

- the transmission capacity of all the lines l ($\bar{P}_l$)
- the generation capacities of all technologies (wind/solar/gas etc.) s at each node n
- the dispatch of each generator at each point in time t

Representative time points are weighted w_t such that $\sum_t w_t = 365 * 24$ and the capital costs c_* are annualised, so that the objective function represents the annual system cost.

Main constraints

- Demand $d_{n,t}$ is always met by generation or transmission

$$d_{n,t} = \sum_s g_{n,s,t} + \sum_{l \in n} f_{l,t}$$

- Dispatch $g_{n,s,t}$ cannot exceed availability $\bar{g}_{n,s,t}$

$$0 \leq g_{n,s,t} \leq \bar{g}_{n,s,t} \leq \bar{g}_{n,s}$$

- Rated capacity cannot exceed the installable potential $\hat{g}_{n,s}$

$$\bar{g}_{n,s} \leq \hat{g}_{n,s}$$

- Transmission flows cannot exceed capacities

$$|f_{l,t}| \leq \bar{P}_l$$

- CO_2 constraint is respected

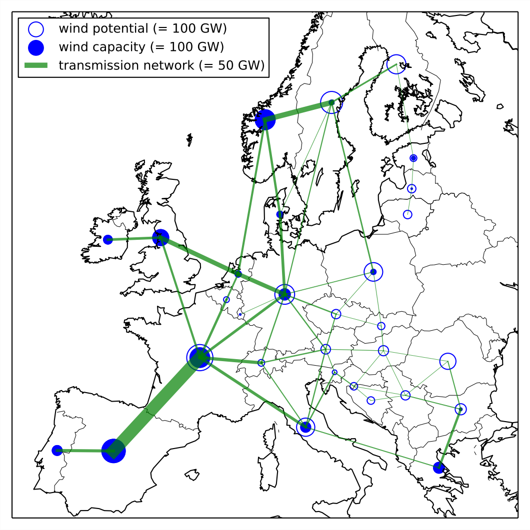
$$\sum_{n,s,t} \frac{1}{\eta_{n,s}} g_{n,s,t} e_{n,s} \leq CAP$$

Nota bene

- Renewables can be arbitrarily curtailed if that's cost-effective
- No assumptions made about which technology is preferred
- No assumptions made about spatial distribution of generation
- No assumptions made about transmission capacity
- Transmission capacity is optimized jointly with generation, e.g. sites far from loads requiring lots of transmission are balanced against sites nearer loads; also balancing made in time, e.g. transmission line not necessarily built out for extreme events

Example application: 30-node model of Europe

'Toy' model: 30 European countries optimised over all 8760 hours of a sample year, only wind, solar and gas generation (hydro is fiddly)



- Minimise electricity system costs (generation & transmission) assuming CO_2 reduction of 90%
- Energy consumed: 86% wind, 10% gas, 4% solar
- 16% curtailment of available wind energy
- Total system cost of € 131 billion / year ~ € 41 /MWh

Cost and other assumptions

Quantity	Cost	Unit
Wind onshore capital	1000	€/kW
Solar capital	1000	€/kW
Gas capital	700	€/kW
Gas marginal	50	€/MWh
Transmission line	400	€/MW/km
Gas CO_2 emissions	0.2	t/MWh$_{\mathrm{thermal}}$
EEA electricity CO_2 emissions 2015	1.5	Gt/a
Gas plant efficiency	40	%
Interest rate	5	%
Line lifetime	40	years
Generators lifetime	20	years

Projects at FIAS over next 18 months

- Look at coupling electricity with heating and transport, probably at country-scale aggregation (most studies on sector coupling have either no spatial resolution or no temporal resolution, missing Europe-wide smoothing benefits)
- Look at electricity sector optimisation with high spatial resolution to see how far it deviates from country-aggregated solutions

Network reduction

Reducing spatial resolution

We need spatial resolution to:

- capture the geographical variation of renewables resources (Niedersachsen versus Bayern) and the load
- capture spatio-temporal effects (e.g. size of wind correlations across the continent)
- represent important transmission constraints

BUT we do not want to have to model all 10,000 network nodes of the European system.

We want to aggregate nodes where load/RE resources are highly correlated and where there are no network bottlenecks inside the clusters (very similar to the concept of the electricity market bidding zone).

Many algorithms in the literature

There are lots of algorithms, particularly in the engineering literature:

- k-means clustering on (electrical) distance
- k-means on load distribution
- Community clustering (e.g. Louvain)
- Spectral analysis of Laplacian matrix
- Clustering of Locational Marginal Prices with nodal pricing (sees congestion and RE generation)
- PTDF clustering
- Cluster nodes with correlated RE time series

The algorithms all serve different purposes (e.g. reducing part of the network on the boundary, to focus on another part).

Not always tested on real network data.

Our metrics of interest

What we want from a network aggregation algorithm:

1. Preservation of major flows within original network
2. Preservation of overall volume of flows
3. For capacity optimisation: representative capacity extensions with aggregated network
4. Preservation of spatial distribution of generation capacity

Best method so far: load custering

Cluster nodes based on load using k-means.

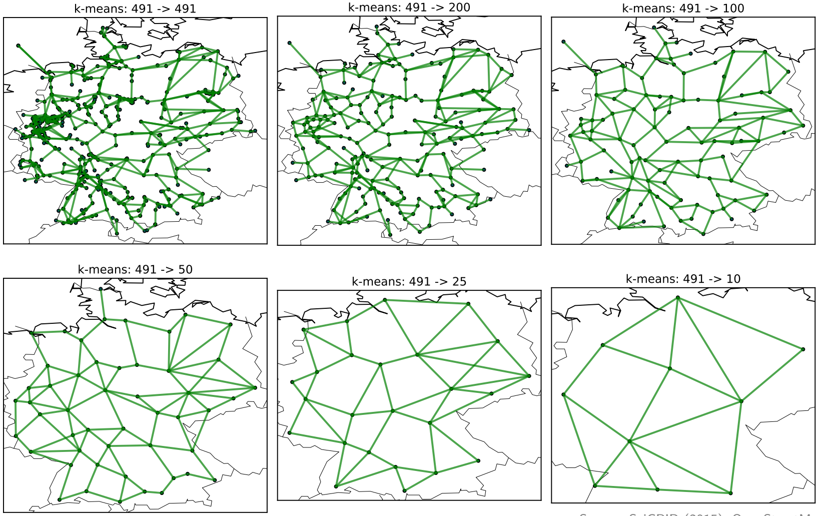
I.e. find k centroids and the corresponding k-partition of the original nodes that minimises the sum of squared distances from each centroid to its nodal members:

$$\min_{\{x_c\}} \sum_{c=1}^{k} \sum_{n \in N_c} w_n ||x_c - x_n||^2 \tag{1}$$

where each node is weighted w_n by the average load there.

NB: Totally ignores grid topology. It works because network is principally laid out to serve the load (with exception of large conventional power plants situated near e.g. mines/rivers).

k-means clustering



Source: SciGRID (2015), OpenStreetMap

Compare flows

Compare the aggregated 'microscopic' flows in the original network to the 'macroscopic' flows in the aggregated network.

We have two networks, our original one 1 and our aggregated one 2.

We have some $N_2 \times N_1$ matrix map B from the buses in 1 to 2 with ones wherever a bus in 1 is aggregated to a bus in 2.

Similarly we have an $L_2 \times L_1$ matrix map L from the lines in 1 to the lines in 2.

Aggregated microscopic flows:

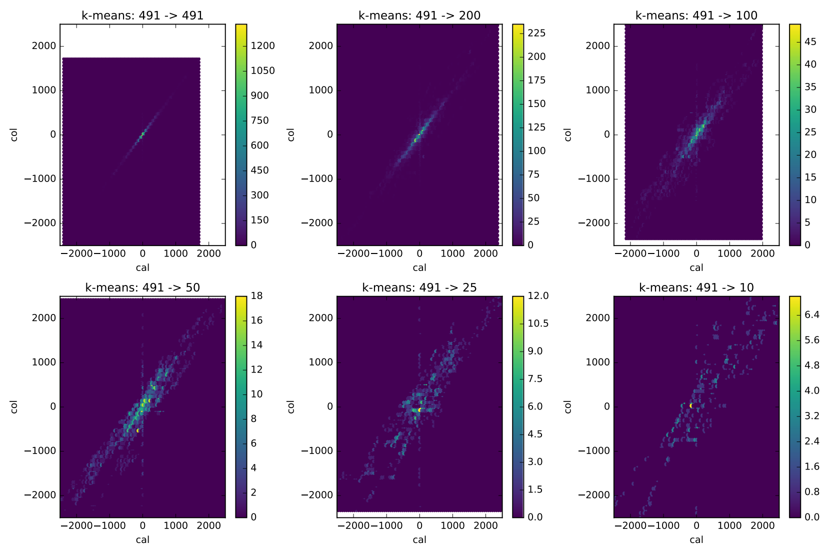
$$L \cdot PTDF_1 \cdot P \tag{2}$$

where P is an N_1 vector of power imbalances at the buses of 1 and $PTDF$ is the linear Power Transfer Distribution Factor.

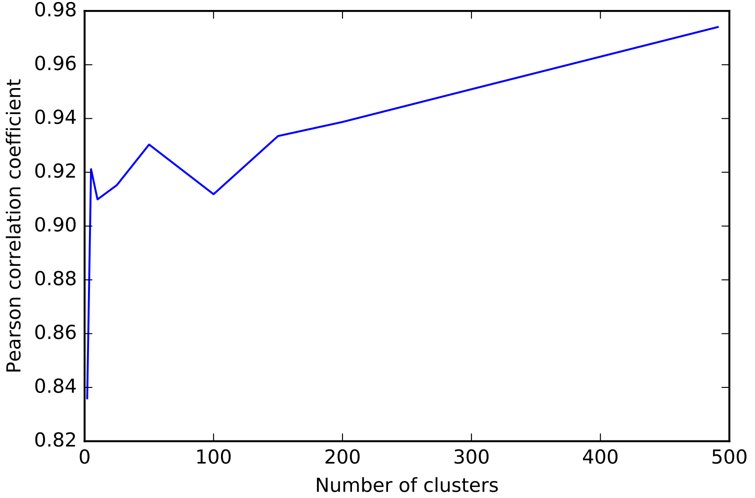
Macroscopic flows in aggregated network:

$$PTDF_2 \cdot B \cdot P \tag{3}$$

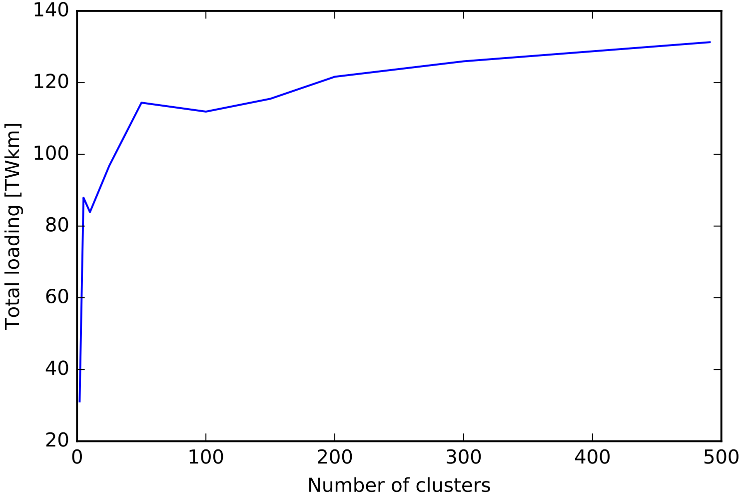
k-means clustering correlation



k-means clustering Pearson



k-means clustering total loading



Time series reduction

Basic idea

It is often desirable to take historical data, for example electrical load, wind and solar generation for many locations sampled hourly over several years, and then reduce the snapshots to a representative sample.

This is often done fairly manually, e.g. summer/winter + high/low wind/load days.

Can we do this more systematically and algorithmically?

The sample should be representative in the following senses:

1. It should reproduce statistics of the full historical data, such as the mean, (co-)variance and higher moments of the time series (so that, for example, capacity factors and spatial correlations for wind and solar are preserved).
2. It should preserve extreme events, such as high electrical demand and low VRE feed-in, which determine backup generation capacities.

More mathsy version

Suppose we have a set I of time series over times T, so that the value of each time series $i \in I$ at time $t \in T$ is given by $x_i(t)$. We treat this like an $|I|$-dimensional random vector $\mathbf{X} = \{X_i\}$ over the event space of time (with uniform probability measure).

The goal is to find a smaller sample $|S| \leq |T|$ and a weighting $w : S \to [0, 1]$ such that common statistics are preserved

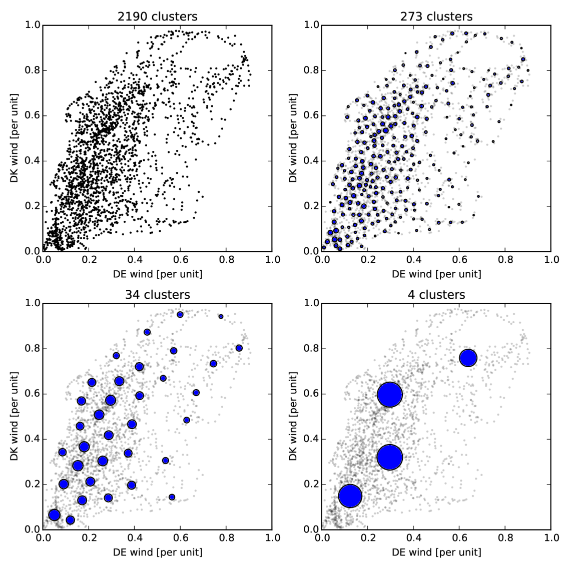
$$\mathbb{E}\left[f(\mathbf{X})\right] \sim \sum_{t \in T} \frac{1}{|T|} f(\mathbf{x}(t)) \sim \sum_{s \in S} w(s) f(\mathbf{x}(s))$$

(depending on the statistic, corrections can be made to get unbiased estimators) and that extremes are also preserved

$$\max_T / \min_T x_i(t) \sim \max_S / \min_S x_i(s)$$

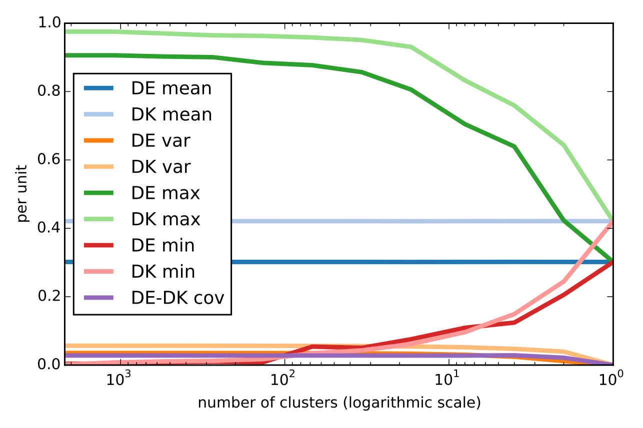
Note that the choice of S and w depends on all the samples $x_i(t)$. The weight imitates a probability measure and satisfies $\sum_{s \in S} w(s) = 1$.

2d example: wind production in Germany and Denmark



- *k*-means clustering creates *k* clusters
- Centroids chosen to minimise the sum of squared distances between the centroids and the original points
- $w(s) \propto$ number of points assigned to each cluster
- Envelope $\sim$ convex hull not so well captured

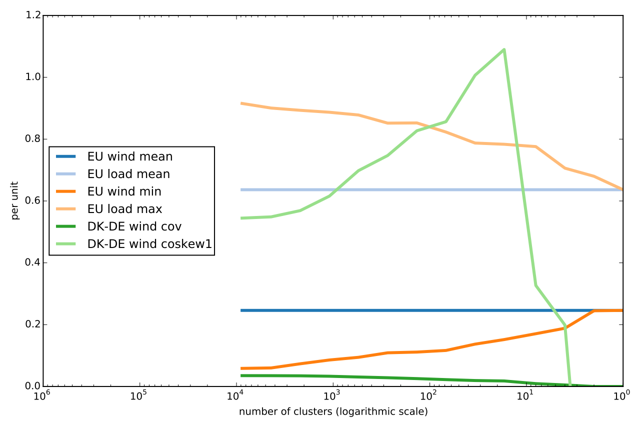
Statistics over time for clustering of $DE - DK$ wind



- Mean and (co-)variance well-preserved as number of clusters decreases
- Below 10 clusters, the extremes (min and max) are not so well captured

European example

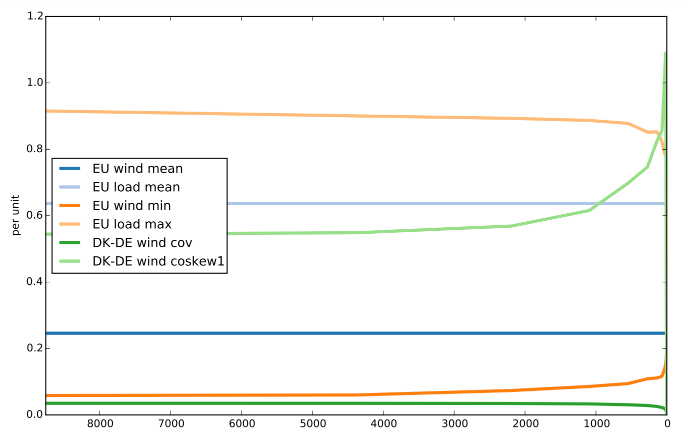
For a bigger European example, enlarge the dimensionality of the random vector from 2 to $90 = (30 \text{ countries}) \times (\text{wind, solar, load})$



- Means always preserved, thanks to algorithm
- With higher dimensionality, more clusters are required to preserve other statistics

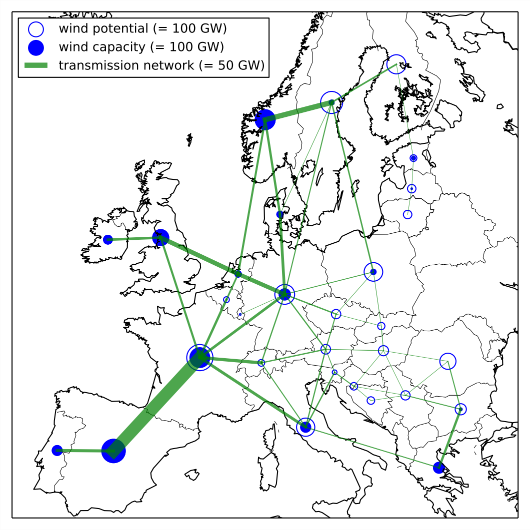
European example without logarithmic scale

Removing the logarithmic scale, it's not so bad...



Example application: 30-node model of Europe

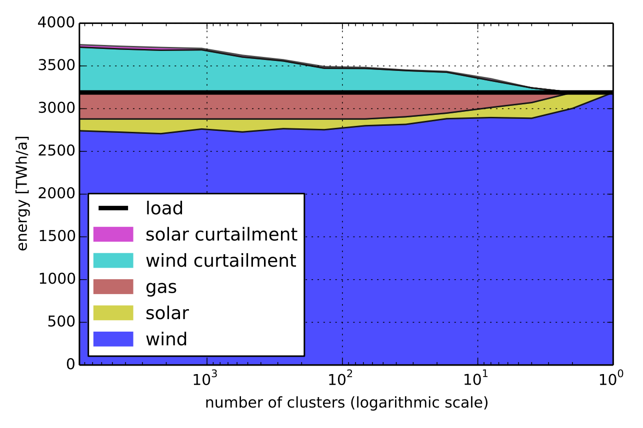
Demonstrate effects on optimisation with example we can solve for all 8760 hours of a sample year: the 30-node European model



- Minimise electricity system costs (generation & transmission) assuming CO_2 reduction of 90%
- Energy consumed: 86% wind, 10% gas, 4% solar
- 16% curtailment of available wind energy
- Total system cost of € 131 billion / year ~ € 41 /MWh

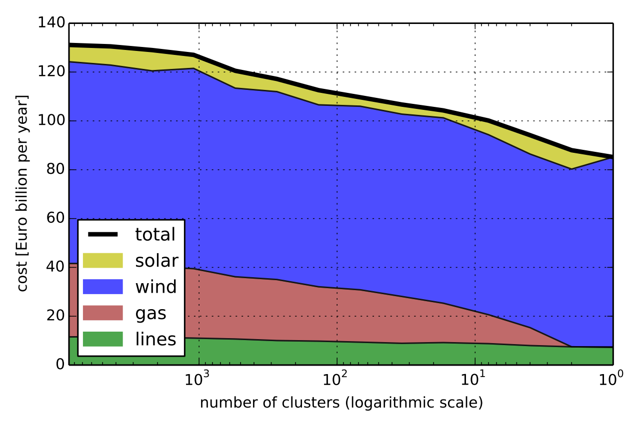
Energy progression

The European electrical load of ~ 3200 TWh/a is mostly met by wind.



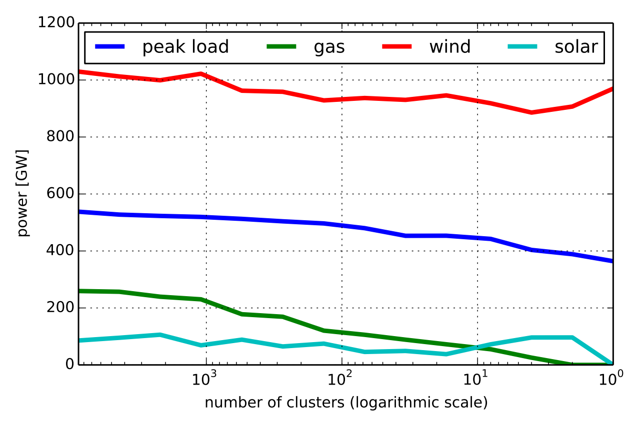
- Main feature is preserved: domination of wind
- For low number of clusters, optimisation doesn't see hours with high load and low VRE, so removes backup energy

Cost progression



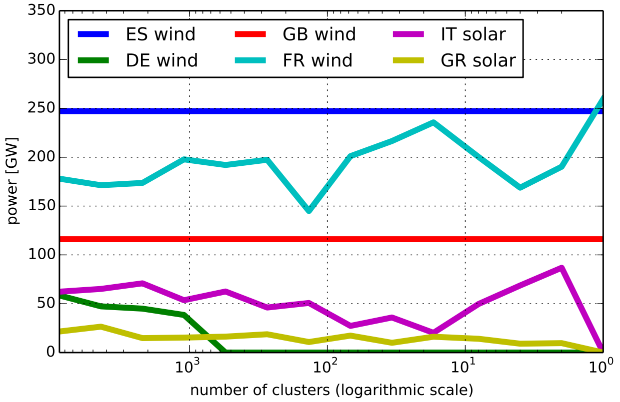
- This story is repeated in the cost summation
- Costs with just one cluster are 34% lower than with all snapshots
- Balance between generation and transmission is well preserved

Installed capacity / peak load progression



- Total wind and solar capacities remain relatively consistent
- Backup gas capacity determined by maximum residual load (load - VRE), which drops from around 260 GW to 0 GW as number of clusters decreases

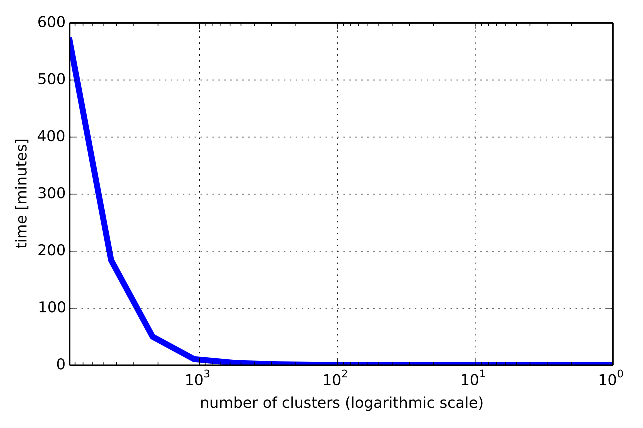
Installed capacity distribution



- Distribution of capacity around Europe shows some variation in countries which are not at the upper capacity potential limit (France, Germany for wind and both Italy and Greece for solar)

Computation time progression

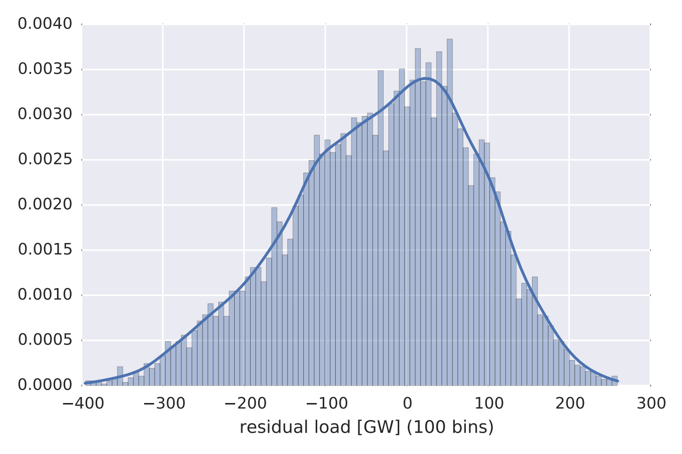
The pay-off is big in terms of computational time:



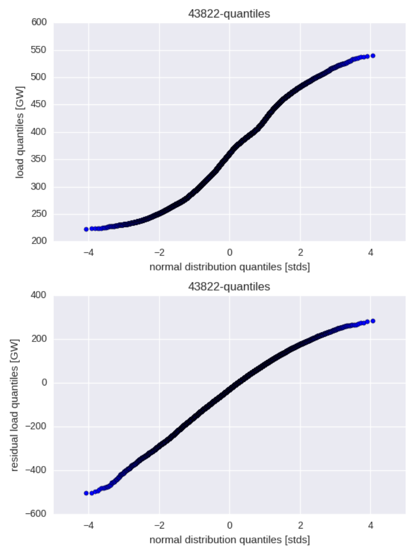
- Reducing number of snapshots by factor 8 reduces computation time from 9.4 hours to 11 minutes.
- This gain comes with little loss of modelling accuracy.

Forcing extreme events

By analysing tail of high residual load (load - VRE), can force in extreme events with appropriate weighting/measure (i.e. include once-in-ten-years event) to capture backup capacity correctly.

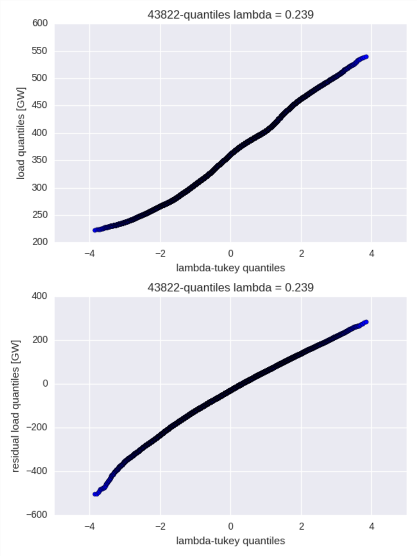


Distribution of extreme events: 5 year sample QQ



- Both load and residual exhibit shorter tails than the normal distribution
- This is GOOD - it means extreme events are less likely than normal

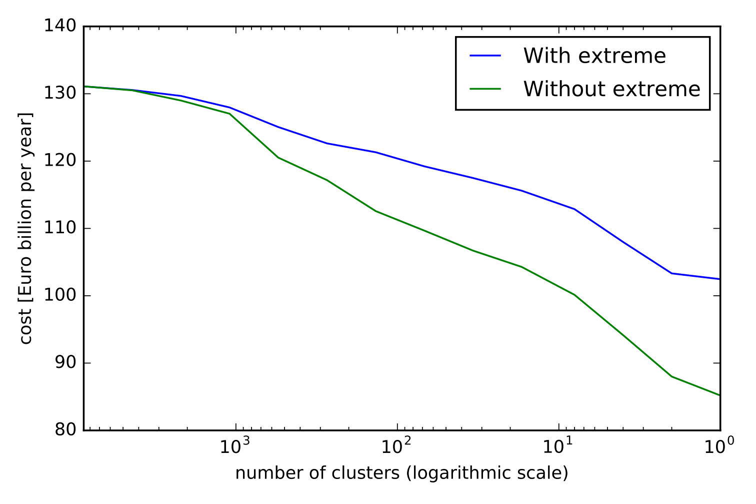
Fitting the tails



- Can fit them to Tukey-Lambda with value $\lambda = 0.239$ (normal distribution is $\lambda = 0.14$; fatter tails are $\lambda < 0.14$, like logistic or t-distribution; $\lambda = 1$ is the uniform distribution)

Results: extreme event maintain total cost better

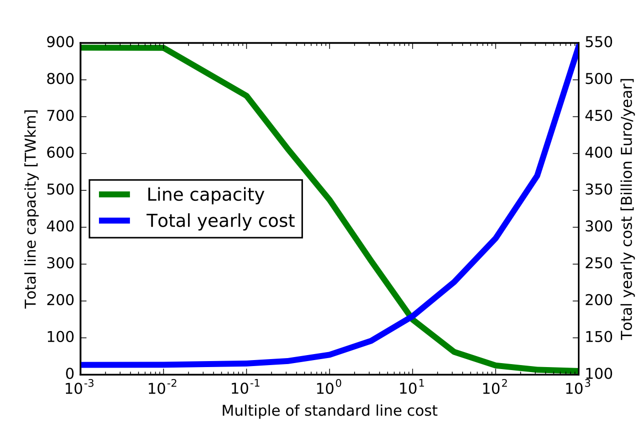
Inserting the once-in-five-years extreme event of high residual load into the data with the correct weight means the total cost is better kept as you reduce the number of nodes



Line cost experiment 1/2

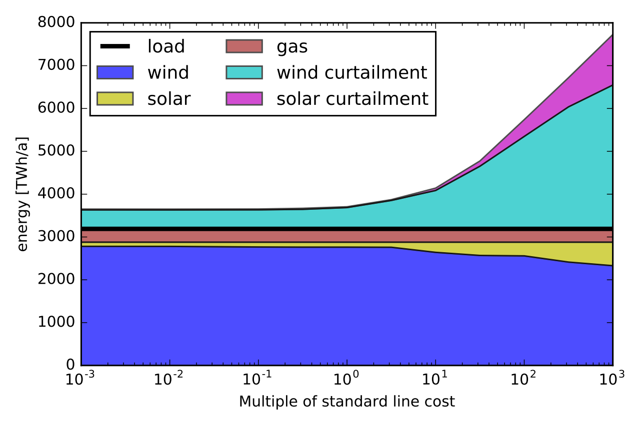
Can do simple sensitivity analysis very quickly, e.g. on changing line cost.

See effect of restricting transmission: Far left is Europe as copper-plate (as much transmission as necessary), far right is no transmission.



- Model results **very** sensitive to change in line costs
- If line costs increase by factor 10 ($\sim$ using underground cables instead of overhead lines), only 32% of lines get built and system costs increase by over 40%

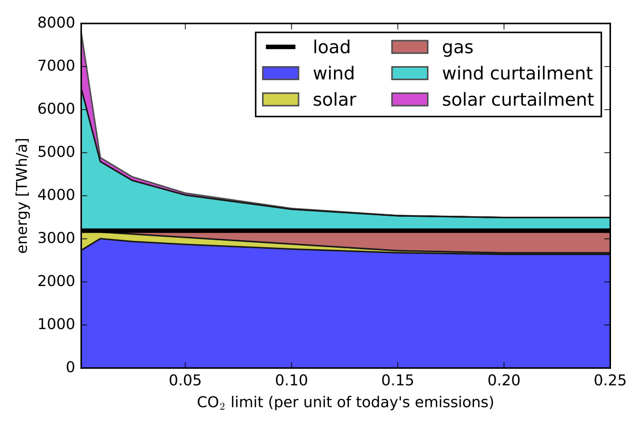
Line cost experiment 2/2



- Curtailment of wind jumps from 14% to 59% as transmission restricted
- Increasing shares of solar as transmission is restricted, with even higher curtailment than for wind (68% curtailment)

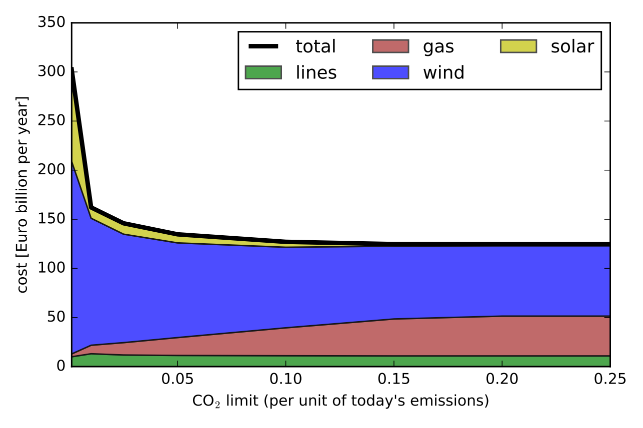
CO_2 restriction experiment 1/3

See effect of reducing backup energy, which in this model is $\propto CO_2$ (although in reality we have carbon-free hydroelectricity)



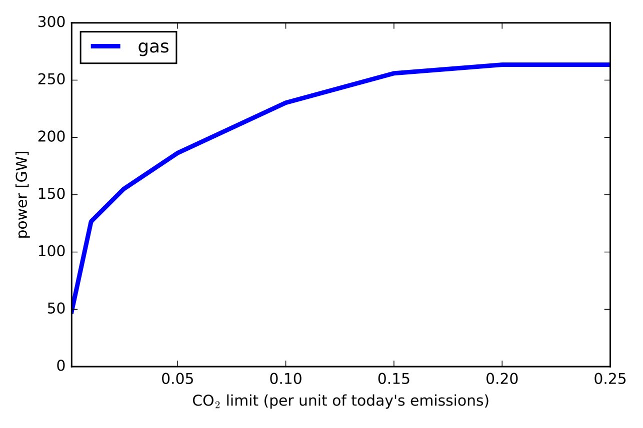
- From 20% and higher, CO_2 restriction is no longer binding $\Longrightarrow$ low-CO_2 is cheaper!
- Cost-optimal backup around 15-20% by energy, which is exactly level of hydroelectricity
- Lower than 1%, curtailment sky-rockets

CO_2 restriction experiment 2/3



- Cost sky-rockets below 1% but quite flat about 5%, so cost-optimum does not require much backup energy

CO_2 restriction experiment 3/3



- But backup capacity still remains high
- Ideally still require 50% of peak load
- Political/security considerations might still require 100-110% backup capacity to peak load ratio

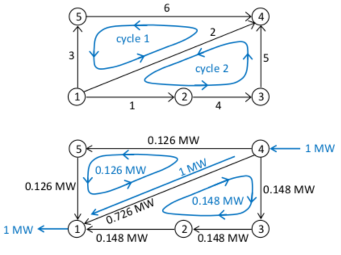
Improving optimisation formulation

Attacking model formulation

Reformulate and benchmark linear optimal power flow formulation:

- Power Transfer Distribution Factor (PTDF) (dense but no auxilliary variables)

$$F_\ell = \sum_i PTDF_{\ell i} P_i$$



- Voltage angles (sparse but N auxilliary variables)

$$F_\ell = \sum_{k,i} B_{\ell k} K^T_{ki} \theta_i$$

- Kirchoff's voltage law around closed cycles (cf. *Ronellenfitsch, Timme, Witthaut (2015)* - semi-sparse and only $L - N + 1$ auxilliary variables)

$$F_\ell = T_\ell + \sum_c C_{\ell c} L_c$$

Conclusions

Conclusions

- There is lots of room for improvement in overall system optimisation.
- Data reduction in time and space can improve optimisation speed, allowing more interesting modelling complexity.
- Moreoever data reduction is essential as the modelling becomes more detailed (integrating other sectors like transport and heating)
- Learning where data reduction does not affect the results tells us important information about the system sensitivities.
- Extreme events must be balanced carefully against “typical” system behaviour, depending on their effect on the results.
- New algorithms can also speed up the optimisation.

Open Modelling Adverts

Small advert 1: We'd like our modelling and results to be as transparent and repeatable as possible. Therefore we aim to make all our code and data available. All the software for the optimisation is available here:

https://github.com/FRESNA/PyPSA

"Python for Power System Analysis" also does non-linear power flow for those interested in voltage and reactive power.

Small advert 2: Open Energy Modelling Initiative `http://openmod-initiative.org/` - join the mailing list! Let's put an end to pointlessly-repeated data munging and increase transparency. Workshop at KTH, Stockholm, 28-29 April, 2016.