RESGen
–
Renewable Energy Scenario Generation Platform

Emil B. Iversen[1,2] **Pierre Pinson**[1,*]**, Igor Arduin**[1,3]

[1] Technical University of Denmark, Center for Electric Power and Energy
[2] Maersk, Analytics Group (DK)
[3] Vattenfall, Wind Power Division (DK)
[*] email: ppin@dtu.dk - webpage: www.pierrepinson.com

IEEE PESGM 2016, Boston, MA (US) - 19 July 2016

with acknowledgements to EPRI and Danish Research Council for funding (e.g., 5s - Future Electricity Markets and CITIES), data providers (e.g., ECMWF)

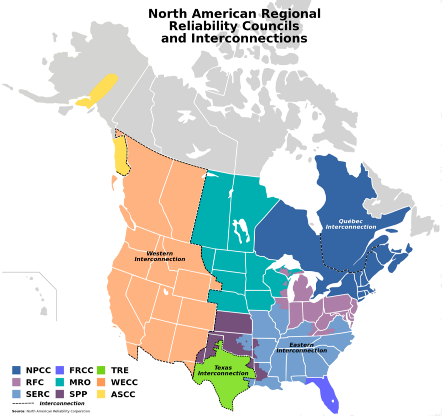
Outline

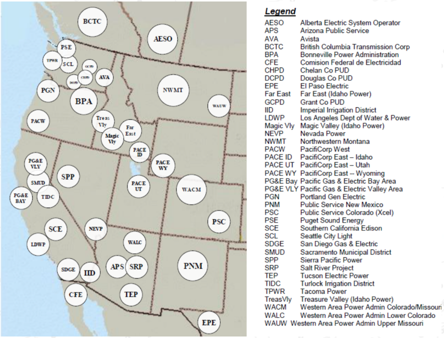
- **Context**
  - Decision-making framework
  - Open code and data
- **Methodological aspects**
  - Input data
  - Predictive marginals
  - (Gaussian) Copulas and covariance models
- **Sample results**
- **Final remarks**

1 **Context**

RESGen: Renewable Energy Scenario Generation platform

- EPRI, US, sponsored us to produce open-source code to generate space-time trajectories (for wind and solar power):
  - the application area is the whole US Western Interconnection!
  - it will be used to feed system studies at various US institutions (e.g., DoE)

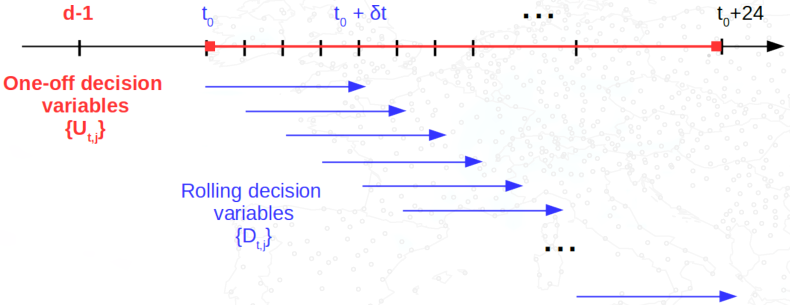




- RESGen is **open-source** (python code and data): any one can use it for operational/planning studies, for any test case of interest, or further develop it...

Originality stemming from the decision framework

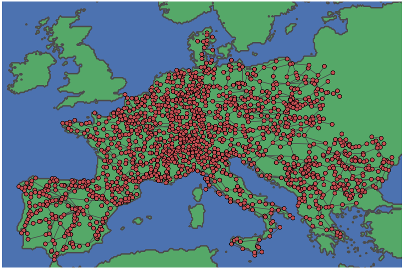
- The first setup for the platform originated from discussion with EPRI
- More than an interest in the stochastic processes themselves, a more complete package was envisaged:
  - simulation of *space-time renewable power generation* (wind and solar),
  - simulation of *single-valued forecasts*, possibly of various quality,
  - simulation of *probabilistic forecasts* and *multivariate space-time trajectories* (somewhat simulating forecast errors),

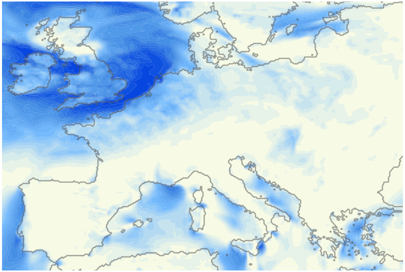


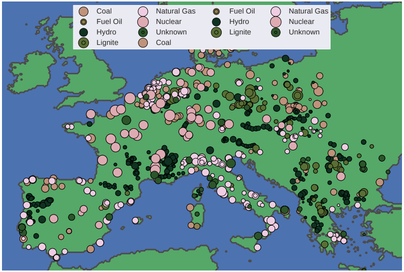
... and, eventually, *diagnostics* and *verification tools*!

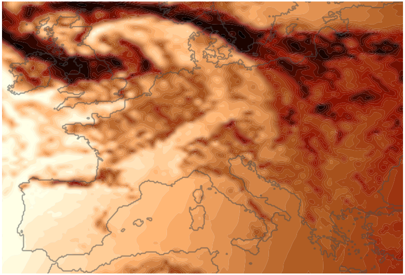
Later on, a bit more ambitious

Let us build a (very) large-scale dataset for the whole European system (RE-Europe)





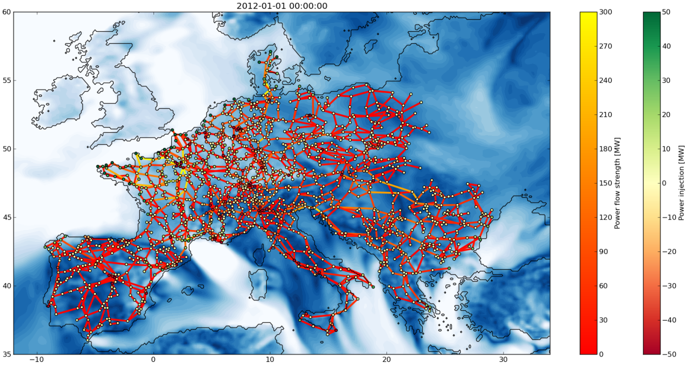




Also **open-access** (python code and data) for research purposes at `zenodo.org` (ackn. due to Tue V. Jensen for hard work on the RE-Europe dataset)

The big picture...

- The **"grand forecasting challenge"**: predict *renewable power generation*, *dynamic uncertainties* and *space-time dependencies* at once for the whole Europe...!



- **Linkage with future electricity markets:**
  - Monitoring, simulating and forecasting of the complete **"Energy Weather"** over Europe
  - Provides all necessary information for coupling of various existing markets (e.g., day-ahead, balancing), as well as simulation for planning decisions

2. **Methodological aspects**

Methodological framework - Inputs and dimensions

- Since the decision framework is similar to the case of operational forecasts, the core of our approach is similar...

- **Inputs available:**
  - renewable power (wind or solar) observations $y_{s_j,t}$ for a number of discrete locations s_j and times t
  - single valued forecasts $\hat{y}_{s_j,t+k|t}$ corresponding to the different problems of interest, i.e., day-ahead and rolling intra-day

- **3 interesting problems:**
  - improve input single valued forecasts $\hat{y}_{s_j,t+k|t}$ (not discussed further)
  - generate predictive marginals $\hat{F}_{s_j,t+k|t}$, somewhat describing the conditional densities of the observations given the forecasts
  - generate sets of high-dimensional space-time trajectories $\hat{\mathbf{z}}_t^{(b)}$, at each and every time t of interest, jointly informing on J locations and K lead times

- **For our 2 setups:** US - $J = 25$, $K = 6$ or 24 (dim(z)=150 or 528)
  EU - $J = 1500$, $K \leq 96$ (dim(z)$\sim$150.000)

- All **predictive marginals**

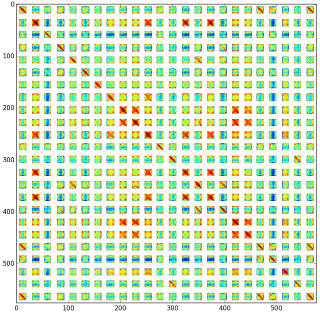
$$\hat{F}_{s_j,t+k|t}(y), \quad j = 1, \ldots, J, \quad k = 1, \ldots, K$$

are linked together using a Gaussian copula

- This is equivalent to working with a latent Gaussian random field in space and in time, then transformed to the power space on a point-wise basis

- **A key result** is that, if predictive marginals $\hat{F}_{s_j,t+k|t}$ are *probabilistically calibrated*, then

$$Z_{s_j,t+k|t} = \Phi^{-1}\left(\hat{F}_{s_j,t+k|t}(Y_{s_j,t+k})\right) \sim \mathcal{N}(0,1)$$

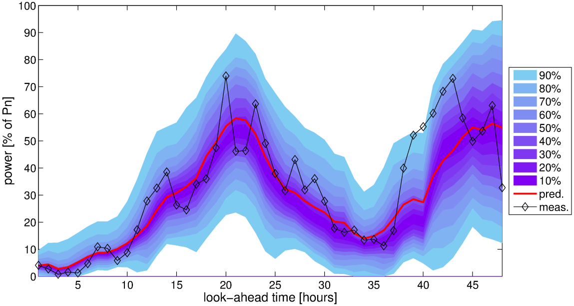


- Consequently the **full dependence structure** simplifies to
  - An empirical covariance matrix $\mathbf{\Sigma}$, to be estimated based on (lots of) data - *US dataset*
  - A covariance model $C(\delta k, \delta s)$, separable or not - *EU dataset*

Obtaining predictive marginals

- **Probabilistic forecasting** of renewable power generation has been widely studied over the last 10 years:
  - in both parametric and nonparametric framework
  - to issue *quantile*, *interval* and *density* forecasts
  - for lead times between a few minutes and several months (or years!) ahead

- Alternative approaches:
  - based on quantile regression
  - using conditional kernel density estimation
  - with a GL-Normal assumption
  - etc.



- In the present case, we use *spline-based linear and nonlinear quantile regression* for a number of nominal levels α_i, $i = 1, \ldots, m$

Dependence structures - Modelling and estimation

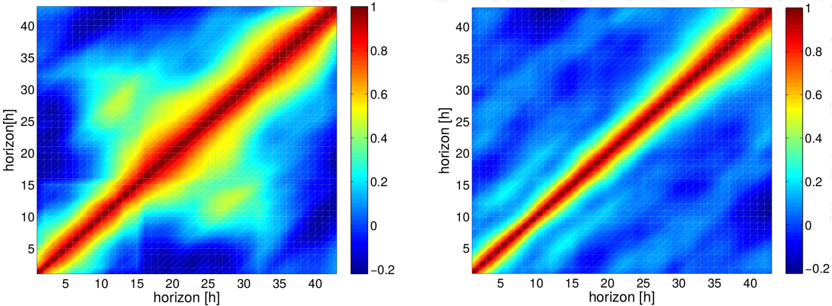
- **Nonparametric:** Simply estimating $\boldsymbol{\Sigma}$ based on all observations transformed through predictive marginals,

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{T-1} \sum_t \mathbf{z}_t \mathbf{z}_t^\top$$

where $\mathbf{z}_t = [z_{s_j,t+k|t}]_{j,k}, \qquad z_{s_j,t+k|t} = \Phi^{-1}\left(\hat{F}_{s_j,t+k|t}(y_{s_j,t+k})\right)$

(possibly needing some scaling since $\mathbf{z}_{s_j,t+k|t}$ may not be perfectly $\mathcal{N}(0,1)$ in practice)

- More advanced approaches may involve *time-varying* covariance tracking, *regime-switching*, etc.



- **Parametric:** By first assuming separability, we fit and analyse simple covariance models for time and space, e.g.
  - exponential, $C(\delta k) = \exp\left(-\delta k/\tau_k\right)$ (similar for δs)
  - powered exponential, $C(\delta k) = \exp\left(-(\delta k/\tau_k)^{\gamma_k}\right)$ (similar for δs)
  - Cauchy, $C(\delta k) = \left(1 + \tau_k \delta k^{2\gamma_k}\right)^{-1}$ (similar for δs)

- We also generalize to non-separable models following the approach of Gneiting *et al.* (2007), based on a coupling parameter β to be estimated in a second stage

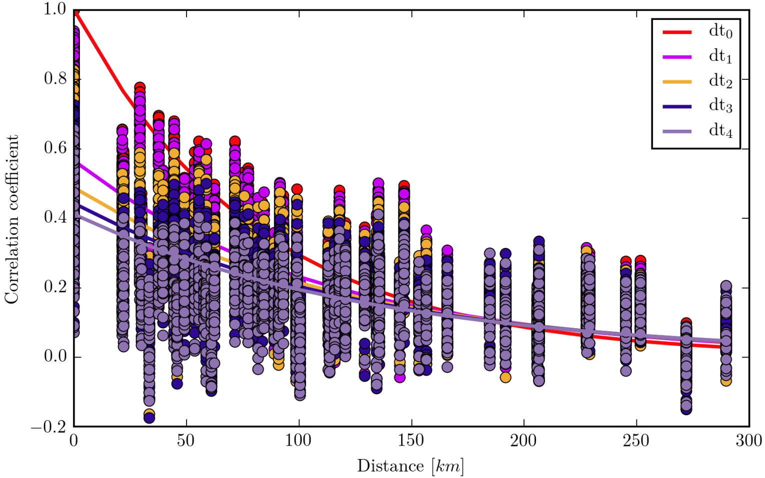
- That we used, eventually, reads

$$C(\delta_k, \delta_s) = \left(1 + \tau_k \delta_k^{2\gamma_k}\right)^{-1} \exp\left\{ - \frac{\delta_s}{\tau_s (1 + \tau_k \delta_k^{2\gamma_k})^{\frac{\beta}{2}}} \right\}$$

For the case $\beta = 0$, one retrieves a separable case.

- Estimation is performed with either direct or 2-stage weighted least squares

Dependence structures - Example non-separable fit

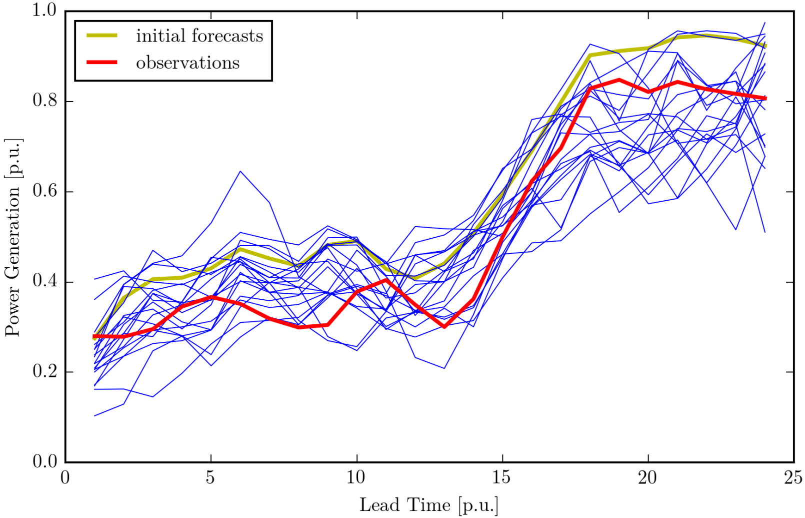


Example for a Denmark subset of data

For our various test cases, the improvement in RMSE from fitting a non-separable model is in the order of 40%...

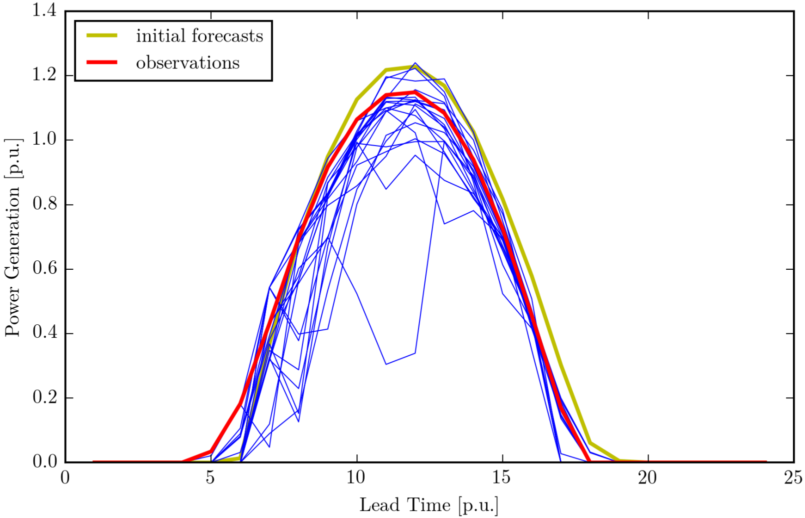
3. **Sample results**

Example day-ahead wind scenarios



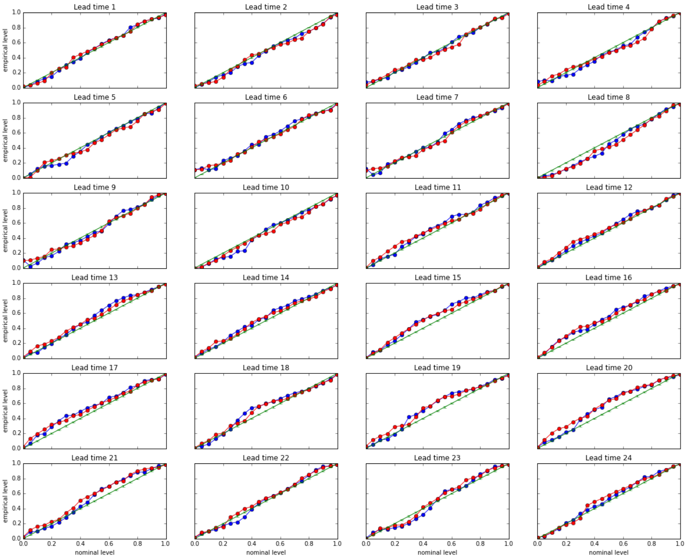
Node 1038 in Denmark (RE-Europe dataset), 5 May 2014, forecasts issued at 12:00

Example day-ahead solar scenarios



Node 205 in Spain (RE-Europe dataset), 2 June 2014, forecasts issued at 00:00

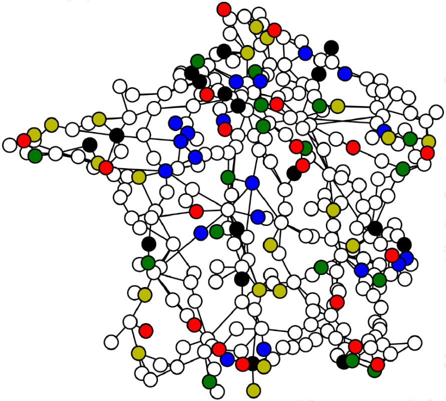
Calibration of predictive marginals



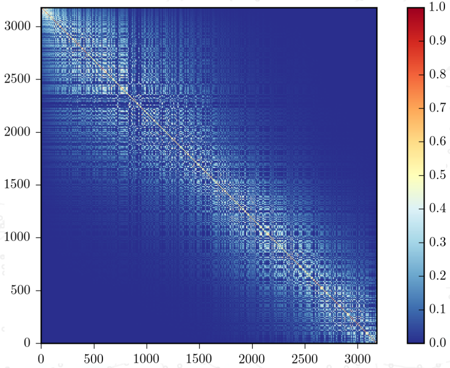
Example calibration results for day-ahead wind trajectories for Western US (given area) - Other properties of the trajectories (e.g., skill, dependence structure) were analysed

An example setup in France

- We considered more focused setups for Western US, France, Portugal, Denmark



318 nodes - various experiments fitting on 5% of data, cross-validating, etc.



... then building the full dependence structure for all nodes (and 10 lead times in that example)

- It was naturally found that space-time dependencies varies in space and in time!

Final remarks

- There is a crucial need for methods permitting to model and simulation high-dimensional space-time stochatic processes for renewable energy applications

- Here the approach presented is simple and flexible...

- Next steps will focus on:
  - nonstationary covariance structures (in space and in time)
  - anisotropic and conditional covariance models
  - modelling of precision matrices
  - possibly using SPDEs

- But also, in terms of verification
  - better understanding of the characteristics of existing toolbox
  - new scores and diagnostic tools

  - Finally, it is a **collaborative venture!** - Download it from Github and improve it...

Relevant references

I. Arduin (2016). Development of an open-source platform for wind and solar probabilistic forecasting. M.Sc. thesis, Technical University of Denmark (pdf)

I. Arduin, E.B. Iversen, P. Pinson (2016). RESGen - Europe set-up. `https://github.com/igorarduin/RESGen_EU`

P. Pinson, R. Girard (2012). Evaluating the quality of scenarios of short-term wind power generation. *Applied Energy* **96**: 12-20 (pdf)

P. Pinson (2013). Wind energy: Forecasting challenges for its operational management. *Statistical Science*, **28**(4): 564-585 (pdf)

P. Pinson, J. Tastu (2013). Discrimination ability of the Energy score. Tech. Rep., Technical University of Denmark (pdf)

J. Tastu, P. Pinson, H. Madsen (2015). Space-time trajectories of wind power generation: Parameterized precision matrices under a Gaussian copula approach. *Lecture Notes in Statistics: Modeling and Stochastic Learning for Forecasting in High Dimension*, Springer: 267-296 (pdf)

Thanks for your attention!