Innovation Fund Denmark

DTU

Probabilistic Forecasting for System Operators in a Low-Carbon Society

Henrik Madsen (DTU)

http://www.smart-cities-centre.org

http://www.henrikmadsen.org

CITIES
Centre for IT Intelligent Energy Systems

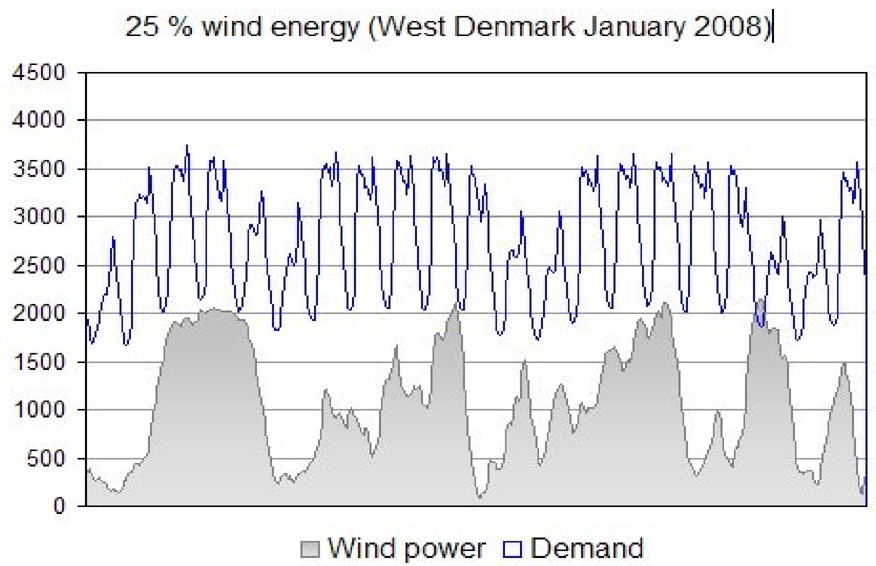
PMAPS 2020

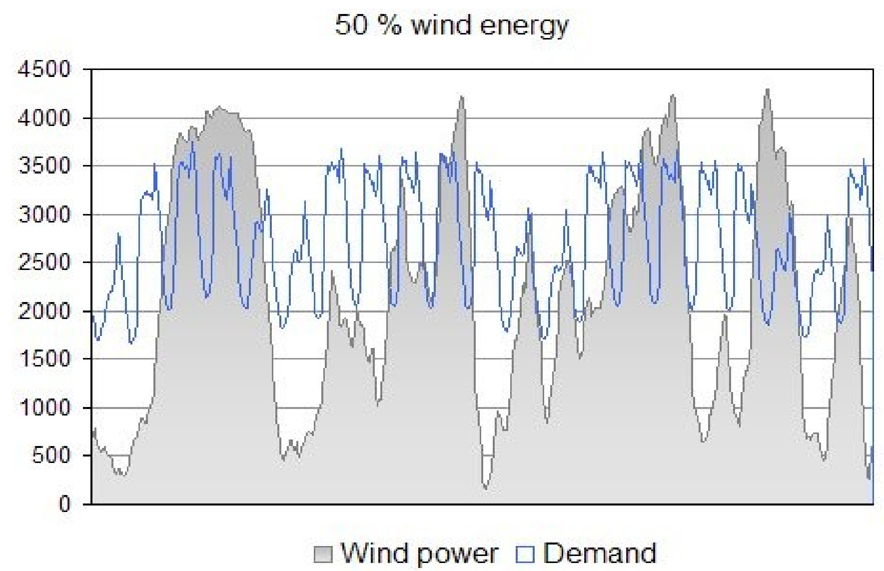
FED FLEXIBLE ENERGY DENMARK

Challenges

The Danish Wind Power Case

…. balancing of the power system





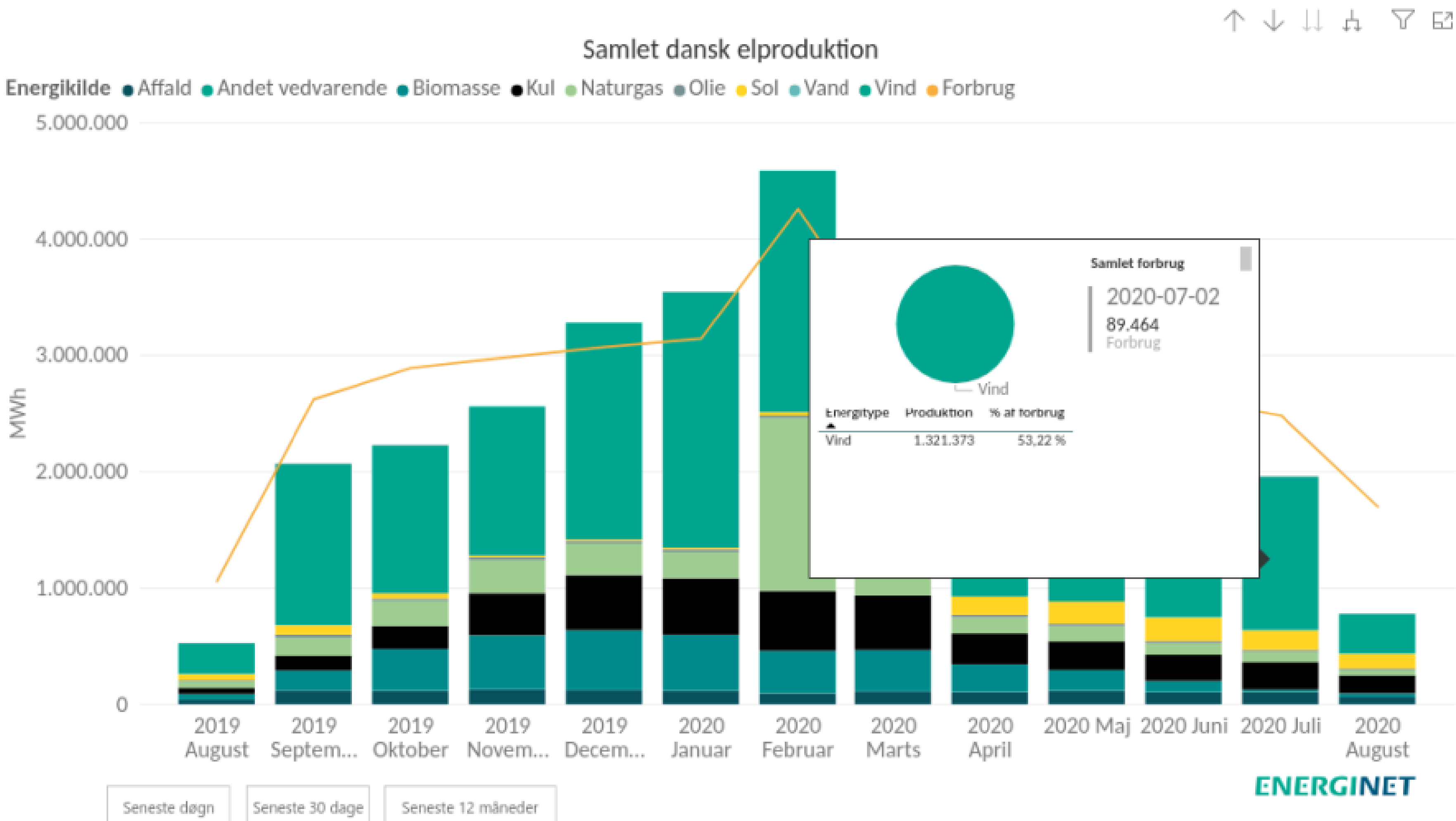
In 2008 wind power did cover the entire demand of electricity in 200 hours (West DK)

In 2020 Forecasting and Flexibility are essential

That's the topic of 'Flexible Energy Denmark'

(For several days the wind power production is more than 100 pct of the power load)

The Danish Wind Power Case



Multivariate Point Forecasting

Uncertainty and adaptivity

Errors in MET forecasts will end up in errors in wind power forecasts, but other factors lead to a need for adaptation which however leads to some uncertainties.

The total system consisting of wind farms measured online, wind turbines not measured online and meteorological forecasts will inevitably change over time as:

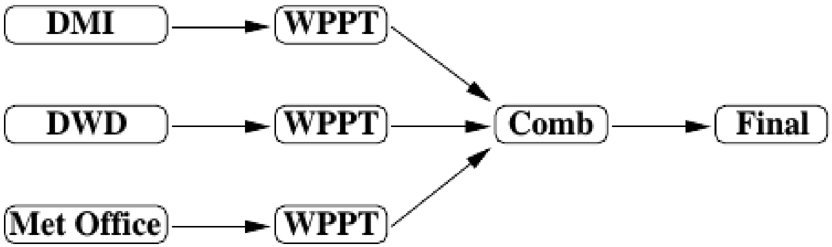
- the population of wind turbines changes,
- changes in unmodelled or insufficiently modelled characteristics (important examples: roughness and dirty blades),
- changes in the NWP models.

A wind power prediction system must be able to handle these time-variations in model and system. An adequate forecasting system may use **adaptive and recursive model estimation** to handle these issues.

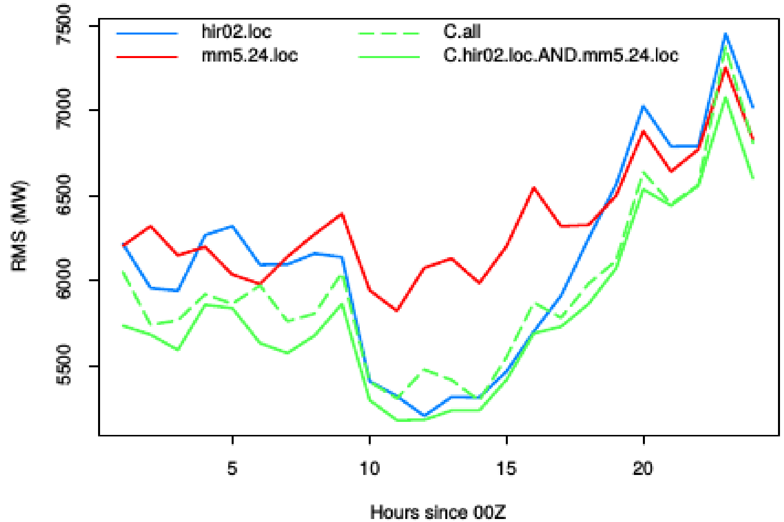
Any reasonable wind and solar power forecasting tool will automatically calibrate the model to the actual situation.

Combined forecasting

- A number of power forecasts are weighted together to form a new improved power forecast.
- These could come from parallel configurations of WPPT using NWP inputs from **different MET providers** or they could come from other power prediction providers.
- In addition to the improved performance also the robustness of the system is increased.



The example show results achieved for the Tunø Knob wind farms using combinations of up to 3 power forecasts.



Typically an improvement on 10-15 pct in accuracy of the point prediction is seen by including more than one MET provider. Two or more MET providers imply information about uncertainty

Innovation Fund Denmark

DTU

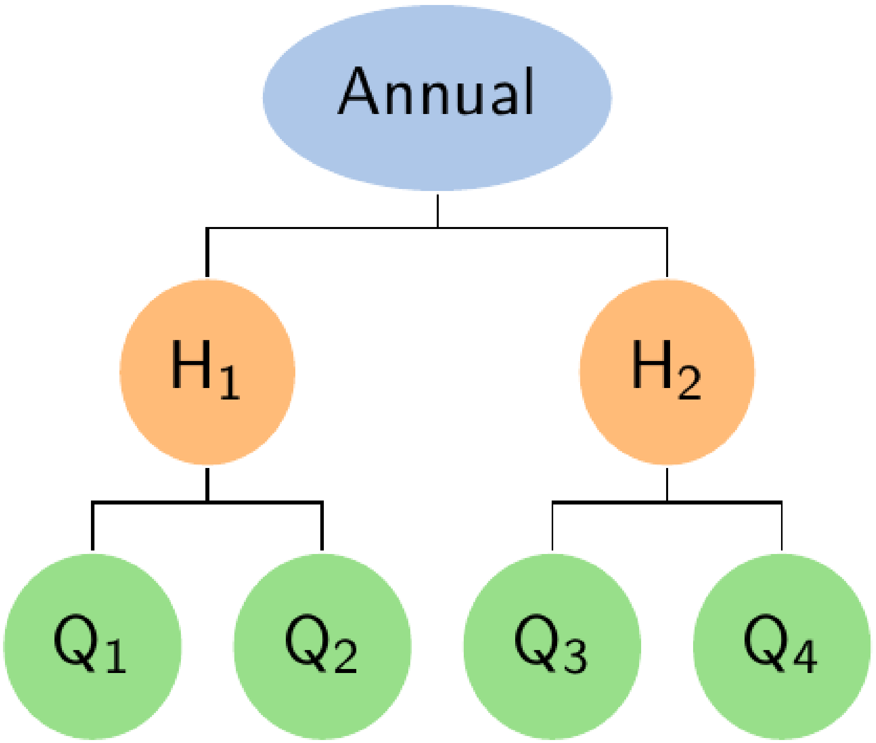
Multivariate Forecasting using Temporal Hierarchies

CITIES
Centre for IT Intelligent Energy Systems

PMAPS 2020

FED FLEXIBLE ENERGY DENMARK

Temporal hierarchy for quarterly series



$$S = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Optimal forecast reconciliation

- Given *base* forecasts $\hat{y} \in \mathbb{R}^n$, we want to find *reconciled* forecasts $\tilde{y} \in \mathbb{R}^n$, which are *coherent*:

$$\begin{aligned} \text{minimize} \quad & (\tilde{y} - \hat{y})^T \Sigma^{-1} (\tilde{y} - \hat{y}) \\ \text{subject to} \quad & \tilde{y} = SP\tilde{y} \end{aligned}$$

- If $\Sigma \in \mathbb{R}^{n \times n}_{++}$ were known, the solution would be given by the *generalized* least-squares estimator

$$\tilde{y} = S \left(S^T \Sigma^{-1} S \right)^{-1} S^T \Sigma^{-1} \hat{y}$$

Weighted least-squares estimation

Athanasopoulos et al. (2017) proposed three estimators that approximate Σ:

$$\Lambda_{\mathsf{struc}} = \mathrm{diag}\,(4, 2, 2, 1, 1, 1, 1)$$
$$\Lambda_{\mathsf{svar}} = \mathrm{diag}\left(\sigma^2_{\mathsf{A}}, \sigma^2_{\mathsf{H}}, \sigma^2_{\mathsf{H}}, \sigma^2_{\mathsf{Q}}, \sigma^2_{\mathsf{Q}}, \sigma^2_{\mathsf{Q}}, \sigma^2_{\mathsf{Q}}\right)$$
$$\Lambda_{\mathsf{hvar}} = \mathrm{diag}\left(\sigma^2_{\mathsf{A}}, \sigma^2_{\mathsf{H}_1}, \sigma^2_{\mathsf{H}_2}, \sigma^2_{\mathsf{Q}_1}, \sigma^2_{\mathsf{Q}_2}, \sigma^2_{\mathsf{Q}_3}, \sigma^2_{\mathsf{Q}_4}\right)$$

Autocovariance scaling
(from P. Nystrup et. al. DTU and LU, 2020)

Estimate the full autocovariance matrix within each aggregation level:

$$\Sigma_{\text{acov}} = \begin{bmatrix} \sigma^2_A & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma^2_{H_1} & \sigma^2_{H_1,H_2} & 0 & 0 & 0 & 0 \\ 0 & \sigma^2_{H_1,H_2} & \sigma^2_{H_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma^2_{Q_1} & \sigma^2_{Q_1,Q_2} & \sigma^2_{Q_1,Q_3} & \sigma^2_{Q_1,Q_4} \\ 0 & 0 & 0 & \sigma^2_{Q_1,Q_2} & \sigma^2_{Q_2} & \sigma^2_{Q_2,Q_3} & \sigma^2_{Q_2,Q_4} \\ 0 & 0 & 0 & \sigma^2_{Q_1,Q_3} & \sigma^2_{Q_2,Q_3} & \sigma^2_{Q_3} & \sigma^2_{Q_3,Q_4} \\ 0 & 0 & 0 & \sigma^2_{Q_1,Q_4} & \sigma^2_{Q_2,Q_4} & \sigma^2_{Q_3,Q_4} & \sigma^2_{Q_4} \end{bmatrix}$$

Innovation Fund Denmark

DTU

Markov scaling

(from P. Nystrup et. al. DTU and LU, 2020)

Only requires estimation of first-order autocorrelation coefficients:

$$\Sigma_{\mathrm{Markov}} = \Lambda^{1/2} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & \rho_{\mathrm{H}} & 0 & 0 & 0 & 0 \\ 0 & \rho_{\mathrm{H}} & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \rho_{\mathrm{Q}} & \rho_{\mathrm{Q}}^2 & \rho_{\mathrm{Q}}^3 \\ 0 & 0 & 0 & \rho_{\mathrm{Q}} & 1 & \rho_{\mathrm{Q}} & \rho_{\mathrm{Q}}^2 \\ 0 & 0 & 0 & \rho_{\mathrm{Q}}^2 & \rho_{\mathrm{Q}} & 1 & \rho_{\mathrm{Q}} \\ 0 & 0 & 0 & \rho_{\mathrm{Q}}^3 & \rho_{\mathrm{Q}}^2 & \rho_{\mathrm{Q}} & 1 \end{bmatrix} \Lambda^{1/2}$$

Example: Load Forecasting (SE1-SE4)

Base forecasts from the additive double-seasonal Holt–Winters method (Taylor 2012):

$$
\begin{aligned}
y_t &= l_{t-1} + s^{(1)}_{t-p_1} + s^{(2)}_{t-p_2} + \phi e_{t-1} + \varepsilon_t \\
e_t &= y_t - \left(l_{t-1} + s^{(1)}_{t-p_1} + s^{(2)}_{t-p_2} \right) \\
l_t &= l_{t-1} + \alpha e_t \\
s^{(1)}_t &= s^{(1)}_{t-p_1} + \gamma_1 e_t \\
s^{(2)}_t &= s^{(2)}_{t-p_2} + \gamma_2 e_t
\end{aligned}
$$

Reconciled hourly forecasts (SE1-SE4)

	In sample (2016)				Out of sample (2017)			
	SE1	SE2	SE3	SE4	SE1	SE2	SE3	SE4
Identity	-8^{**}	-2	-8^{**}	-4^{*}	-1	2	-1	0
Structural	-10^{**}	-4^{**}	-9^{**}	-5^{**}	-4^{**}	-1	-2	-1
Series variance	-10^{**}	-4^{**}	-7^{**}	-4^{**}	-4^{**}	-1	-2	-1
Hierarchy variance	-10^{**}	-4^{**}	-10^{**}	-8^{**}	-4^{**}	-2	-6^{**}	-6^{*}
Structural Markov	-10^{**}	-5^{*}	-13^{**}	-9^{**}	-4^{**}	-1	-6^{**}	-4^{*}
Series Markov	-11^{**}	-5^{**}	-12^{**}	-8^{**}	-5^{**}	-2	-6^{**}	-4
Hierarchy Markov	-11^{**}	-7^{**}	-23^{**}	-17^{**}	-5^{**}	-5^{*}	-17^{**}	-13^{**}
Autocovariance	-14^{**}	-10^{**}	-26^{**}	-20^{**}	-7^{**}	-7^{**}	-20^{**}	-16^{**}
Series GLASSO	-27^{**}	-21^{**}	-27^{**}	-27^{**}	-19^{**}	-18^{**}	-21^{**}	-23^{**}
Series shrinkage	-26^{**}	-22^{**}	-29^{**}	-29^{**}	-18^{**}	-19^{**}	-23^{**}	-25^{**}
Cross-covariance	-35^{**}	-33^{**}	-55^{**}	-48^{**}	-23^{**}	-30^{**}	-47^{**}	-41^{**}

Probabilistic forecasting

Some methods

- Quantile regression
- Stochastic differential equations
- Again – adaptivity, combined forecasting, multivariate, ...

Quantile regression

A (additive) model for each quantile:

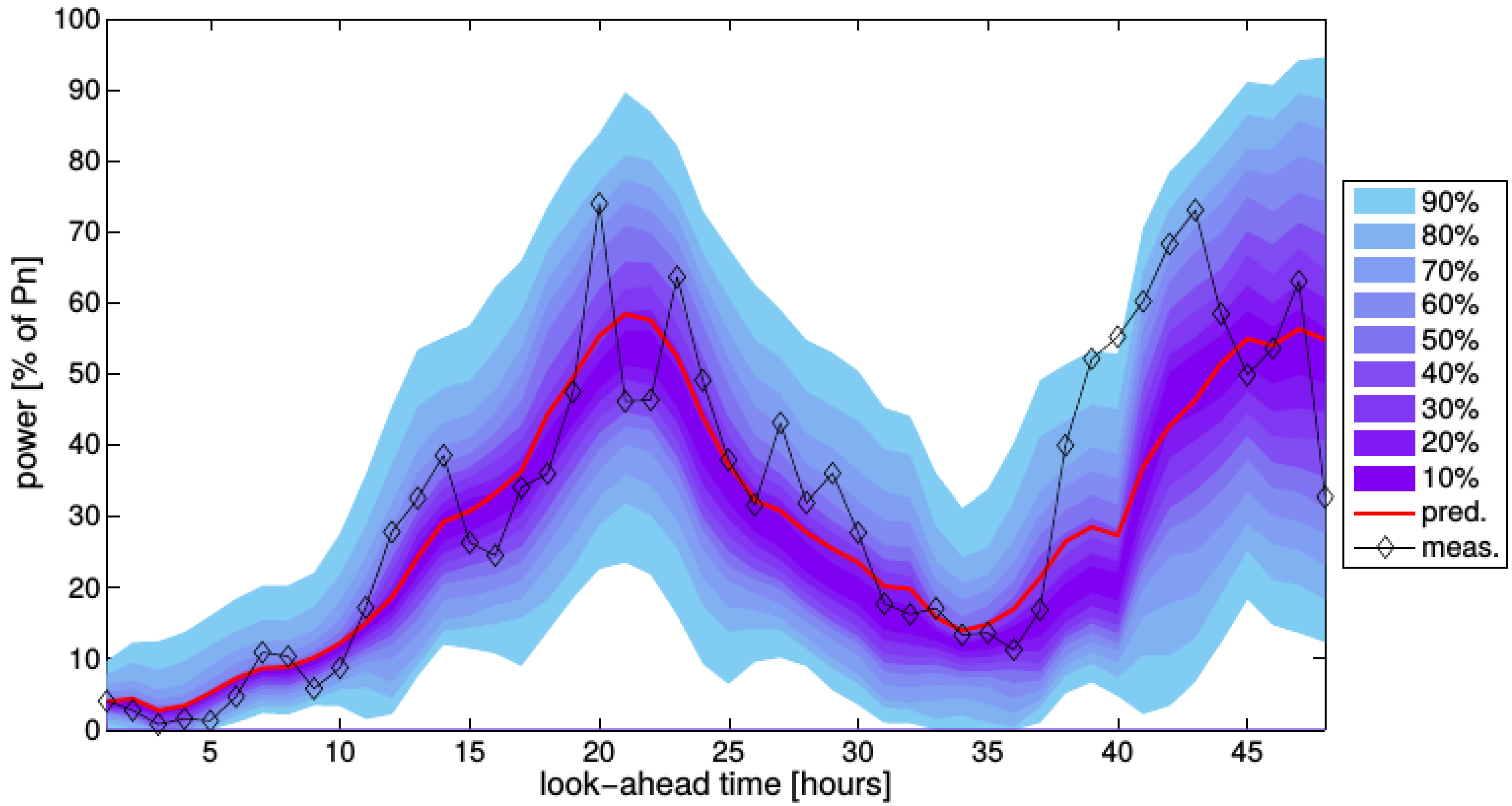
$$Q(\tau) = \alpha(\tau) + f_1(x_1;\tau) + f_2(x_2;\tau) + \ldots + f_p(x_p;\tau)$$

$Q(\tau)$	Quantile of **forecast error** from an **existing system**.
x_j	Variables which influence the quantiles, e.g. the wind direction.
$\alpha(\tau)$	Intercept to be estimated from data.
$f_j(\cdot;\tau)$	Functions to be estimated from data.

Notes on quantile regression:

- Parameter estimates found by minimizing a dedicated function of the prediction errors.
- The variation of the uncertainty is (partly) explained by the independent variables.

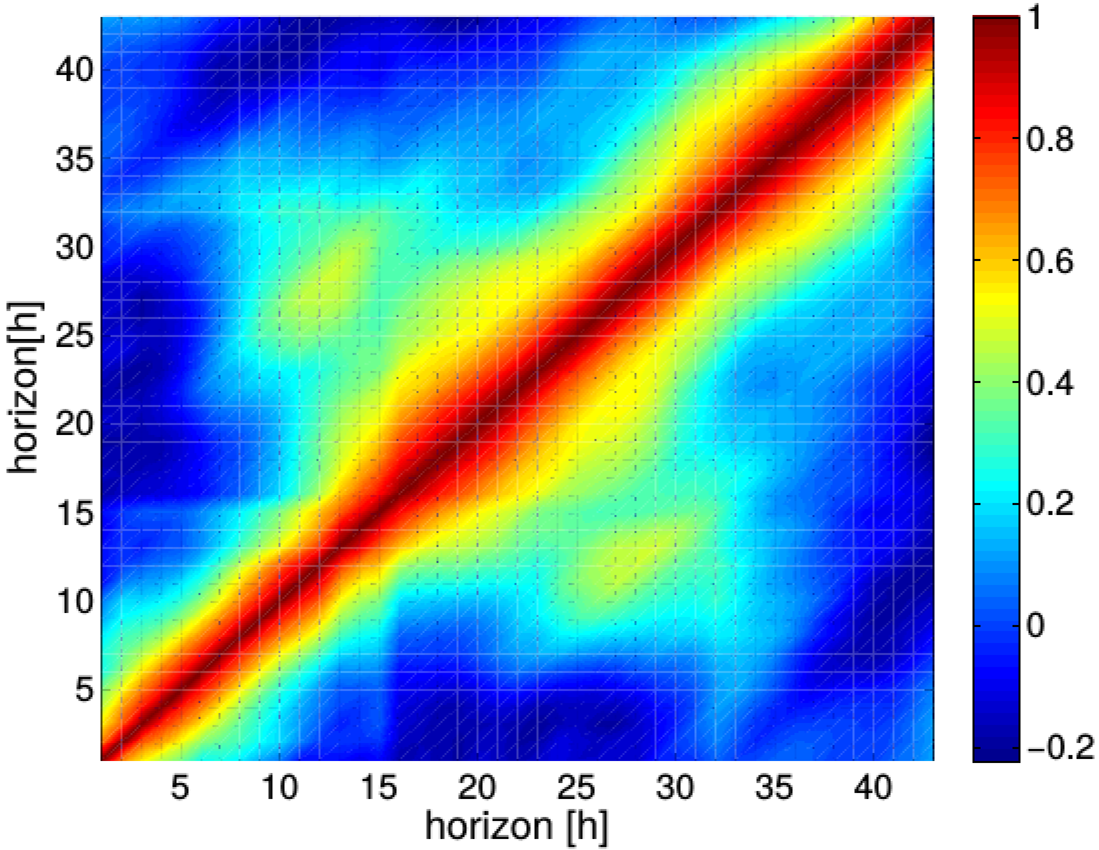
Example: Probabilistic forecast



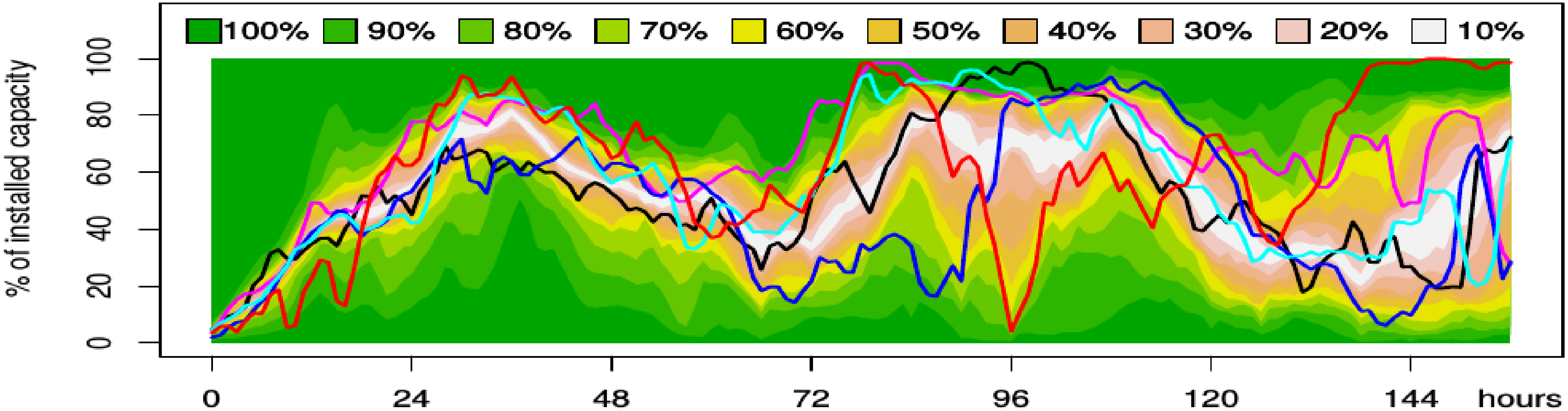
- Notice how the confidence intervals varies ...
- But the correlation in forecasts errors is not described so far.

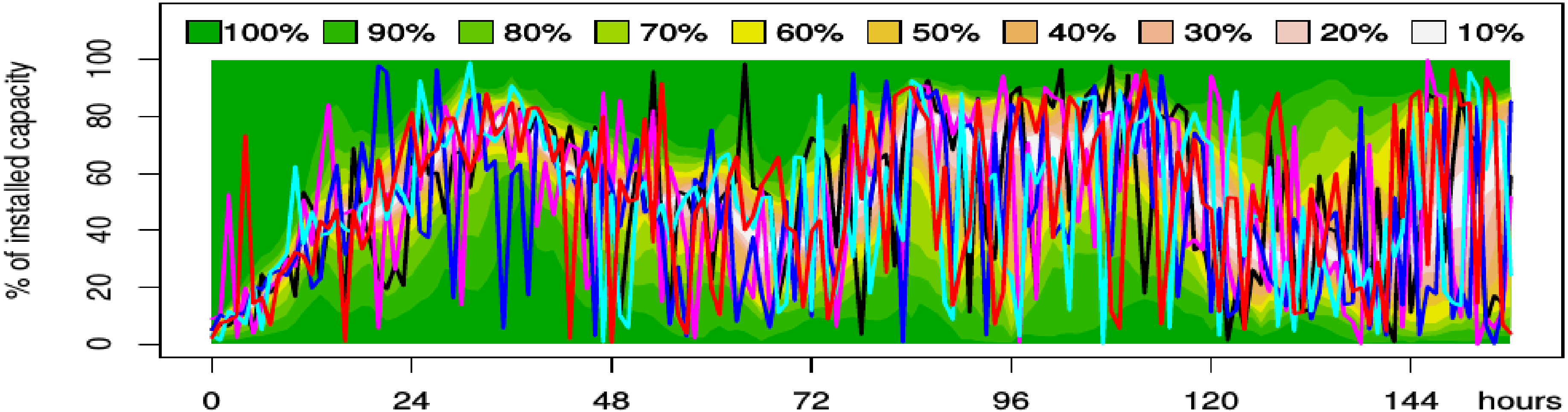
Correlation based structure

- It is important to model the **interdependence structure** of the prediction errors.
- An example of interdependence covariance matrix:

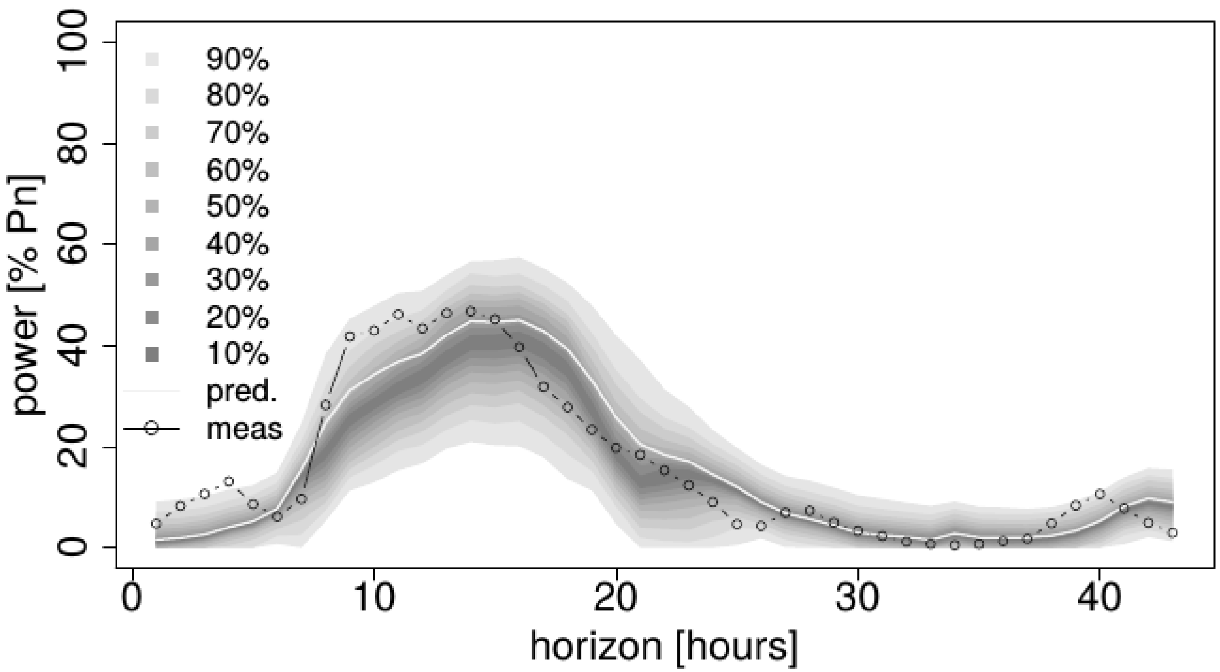


Correct and Naive Scenarios



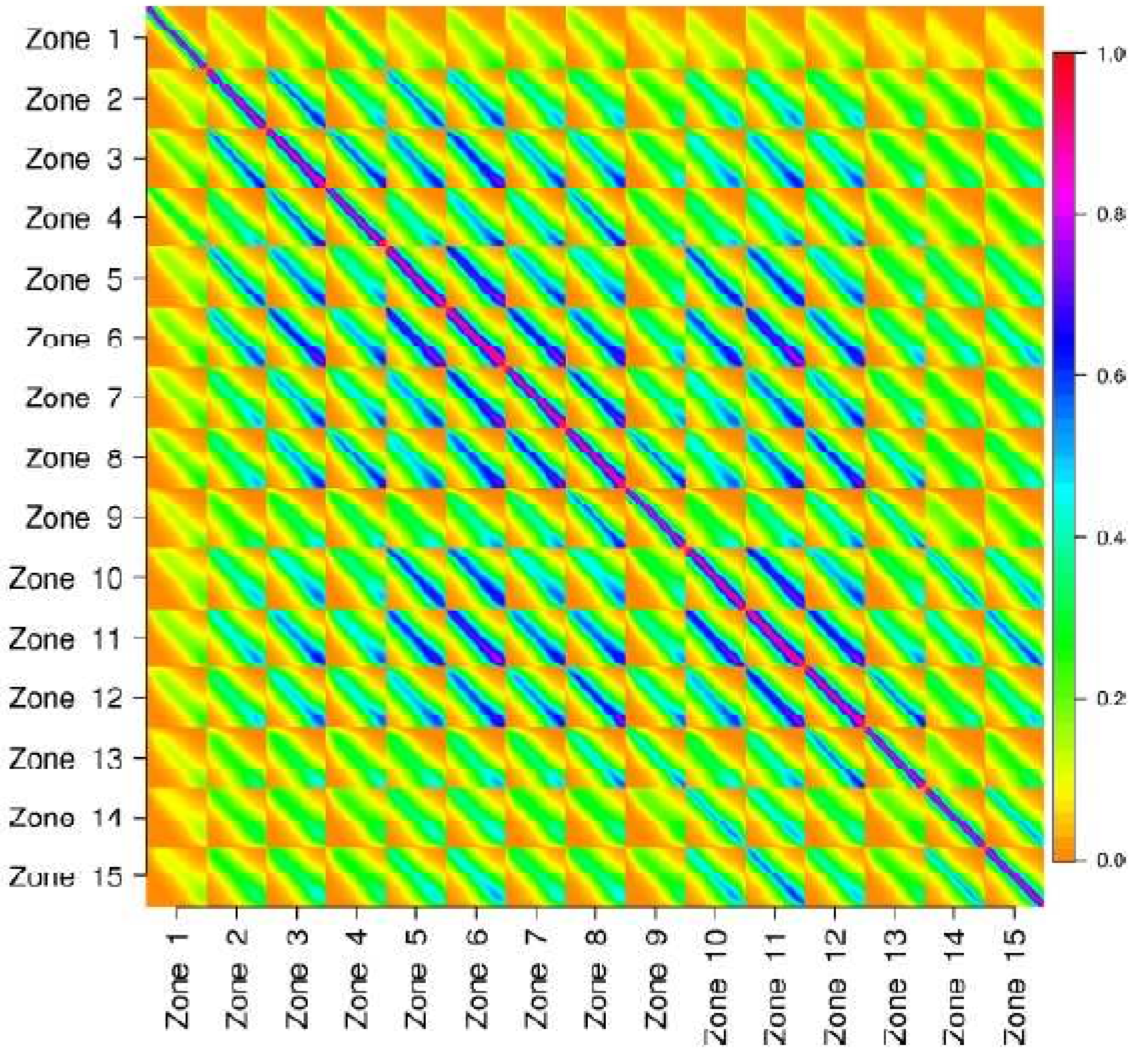


Space-Time dependencies ...



This is not enough...

Space-Time Correlations



State-of-the-art
Probabilistic Forecasting

SDEs for Forecasting

The basic stochastic differential equation formulation:

$$X_t \quad = \quad X_0 + \int_0^t f(X_s, s)ds + \int_0^t g(X_s, s,)dW_s,$$

We use the short-hand interpretation of this integral equation:

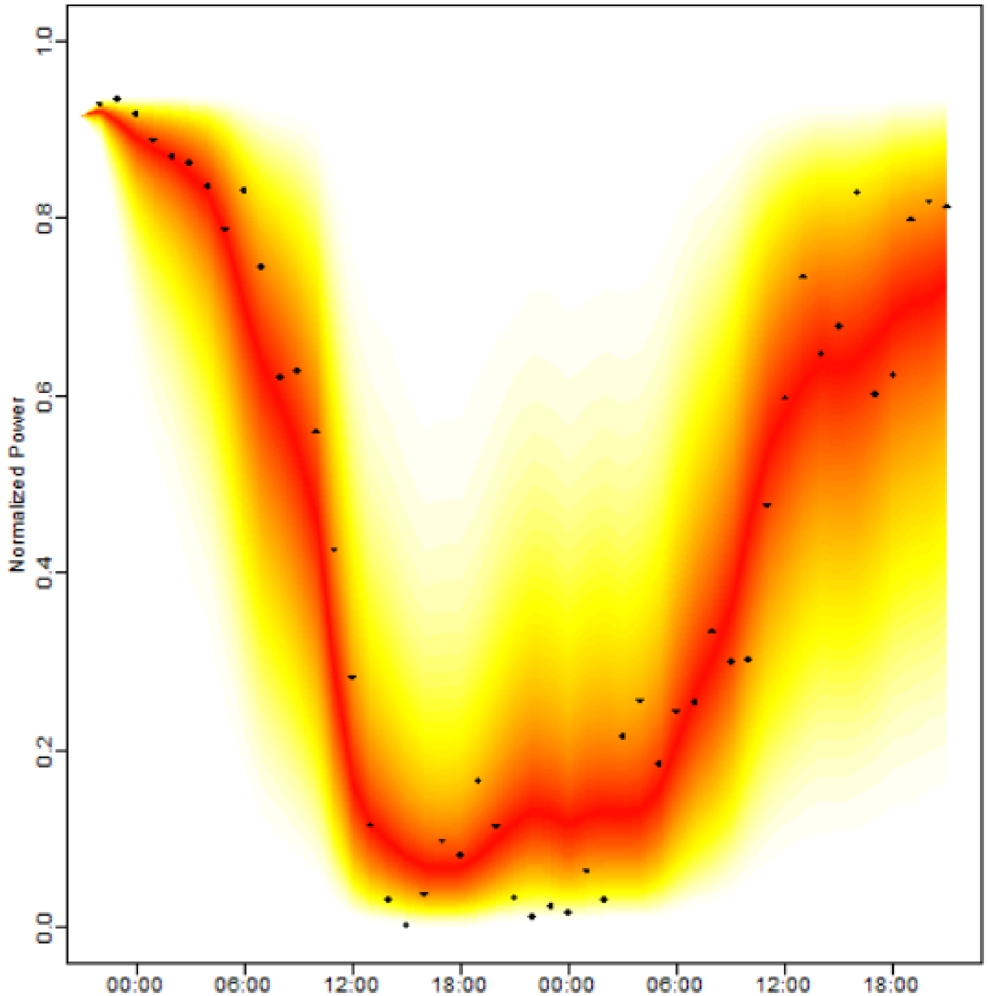
$$\begin{aligned} dX_t &= f(X_t, t)dt + g(X_t, t)dW_t \\ Y_k &= h(X_{t_k}, t_k, e_k). \end{aligned}$$

The predictive density, $j(x, t)$, can be found by solving (with $g(X_t, t) = \sqrt{2D(X_t, t)}$):

$$\frac{\partial}{\partial t} j(x, t) = -\frac{\partial}{\partial x}\left[f(x, t)j(x, t)\right] + \frac{\partial^2}{\partial x^2}\left[D(x, t)j(x, t)\right]. \qquad (1)$$

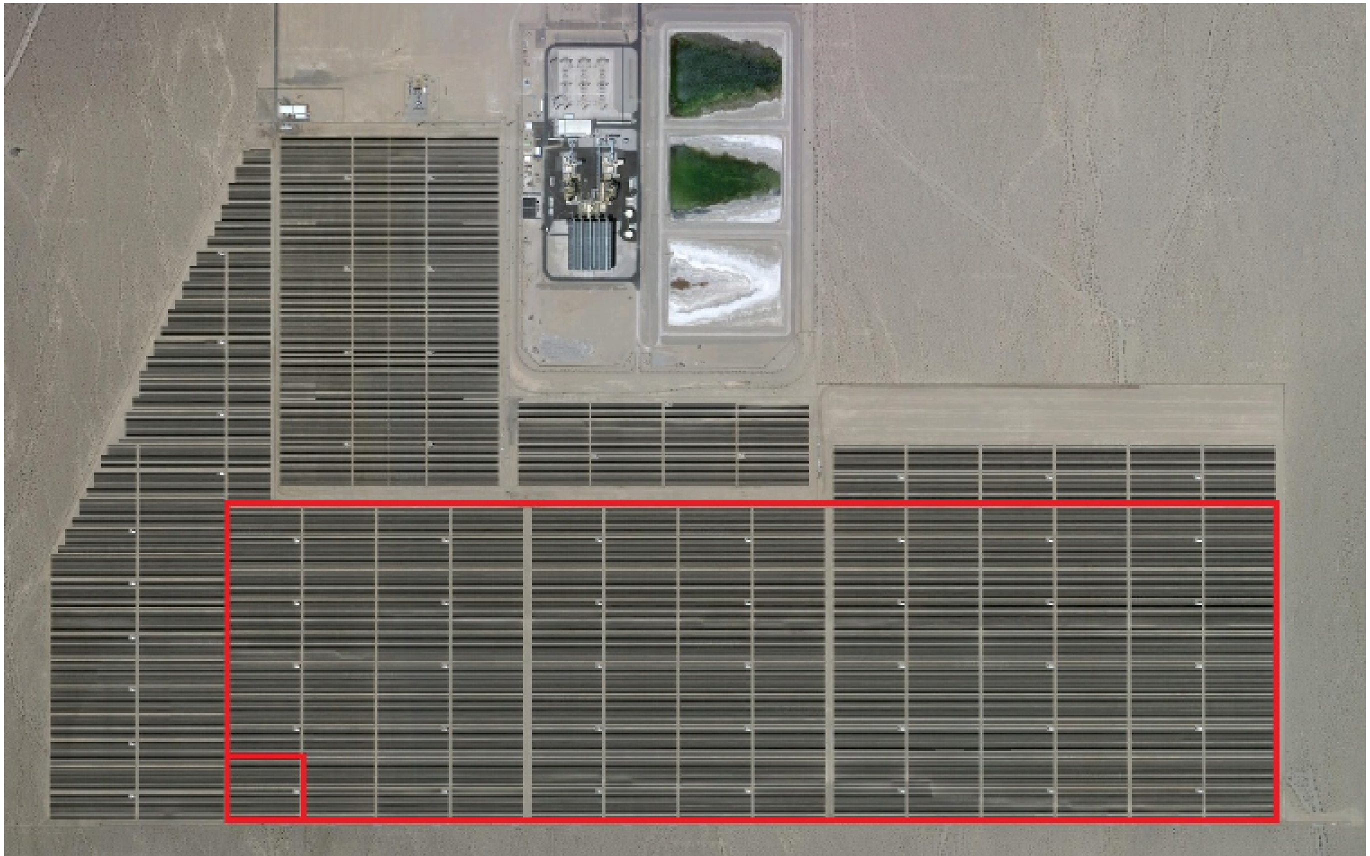
Multi-Horizon Prob. Forecasting

Predictive density of production in percent out of rated power for the Klim wind farm:



Spatio-Temp. Forecasting
Solar Power Plant

- A solar power plant with a nominal output of 151 MW.
- Measurements of 91 inverters every second for one year.
- We consider a cutout of 5 by 14 inverters for modeling.



Innovation Fund Denmark

DTU

SPDE Model Performance

	Auto-Regressive	Model
CRPS_5	0.00262	**0.00131**
CRPS_{20}	0.00982	**0.00666**
CRPS_{60}	0.02886	**0.02455**
CRPS_{120}	0.04883	**0.04675**

Center Denmark

Green transition paved by green innovation


CENTER
DENMARK
Connect networks and data
for a green world
Danmarks
nationale Center
Fremme den grønne omstilling.
Samle og bygge bro, mellem
forskning, teknologi, natur og formidling,
på tværs af interesseorganisationer,
virksomheder, skoler og
universiteter.

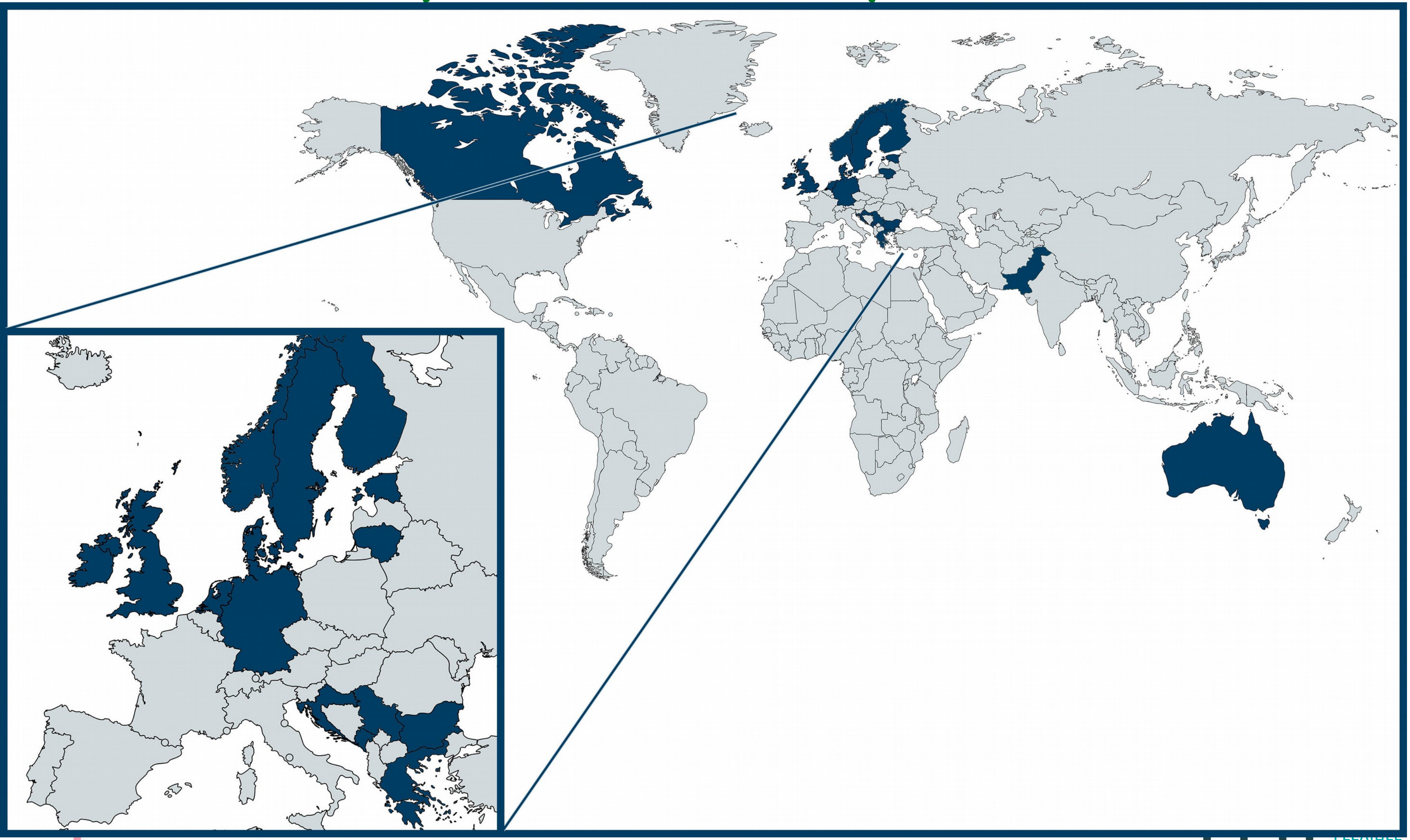
Innovation Fund Denmark

DTU

CITIES
Centre for IT Intelligent Energy Systems

FED FLEXIBLE ENERGY DENMARK

Wind Power Forecasting Installations (WPPT/WindFOR)



Innovation Fund Denmark

DTU

CITIES
Centre for IT Intelligent Energy Systems

PMAPS 2020

FED FLEXIBLE ENERGY DENMARK

Lessons Learned

- The forecasting models must be **adaptive** (in order to taken changes of dust on blades, changes roughness, etc., into account).
- Reliable estimates of the **forecast accuracy** is very important (check the reliability by eg. reliability diagrams).
- Reliable probabilistic forecasts are important to gain the **full economical value**.
- Use **more than a single MET provider** for delivering the input to the prediction tool – this improves the accuracy of wind power forecasts with 10-15 pct.
- Estimates of the **correlation in forecasts errors** important.
- Forecasts of '**cross dependencies**' between load, prices, wind and solar power are important.
- **Probabilistic forecasts are very important for asymmetric cost functions**.
- Probabilistic forecasts can provide **answers** for questions like
  - What is the probability that a given storage is large enough for the next 5 hours?
  - What is the probability of an increase in wind power production of more that 50 pct of installed power over the next two hours?
  - What is the probability of a down-regulation due to wind power on more than x GW within the next 4 hours.

For more information ...

See for instance

www.smart-cities-centre.org

...or contact

– Henrik Madsen (DTU Compute)

hmad@dtu.dk

Acknowledgements:

Innovation Fund Denmark projects: CITIES and Energy Flexible Denmark (FED).

ENFOR a/s (www.enfor.dk) for providing a number of plots etc.

Some 'randomly picked' books on modeling and renewable integration ....