Regime-switching models for prediction of wind power fluctuations

P. Pinson, P.-J. Trombe and Henrik Madsen

Technical University of Denmark, Informatics and Mathematical Modelling

* pp@imm.dtu.dk, http://www.imm.dtu.dk/∼pp

IEA Task 23 - Power Fluctuations Workshop, February 2009

Fluctuations of offshore wind power

- Fluctuations at **large offshore wind farms** have a significant impact on the control and management strategies of their power output
- Focus is given to the **minute scale**. Thus, the effects related to the turbulent nature of the wind are smoothed out
- When looking at time-series of power production at **Horns Rev** (160MW) and **Nysted** (165 MW), one oberves successive periods with fluctuations of larger and smaller magnitude

- We aim at building models
  - based on historical wind power measures only...
  - ... but able to reproduce this observed behavior
  - this calls for **regime-switching models**

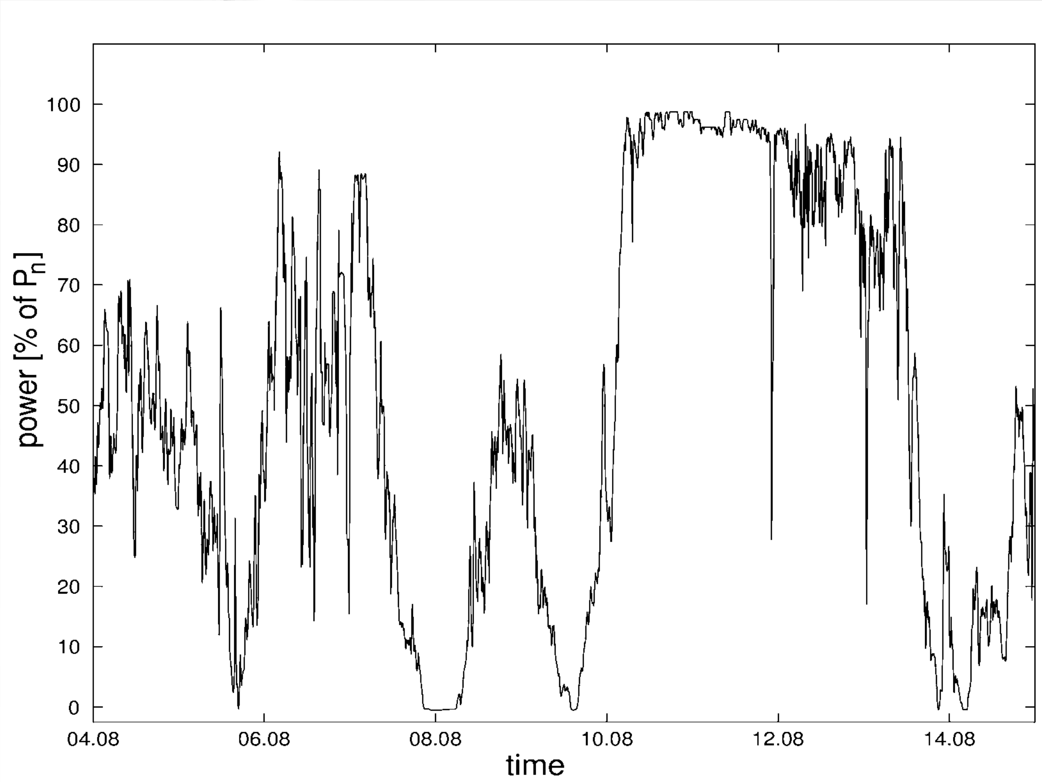
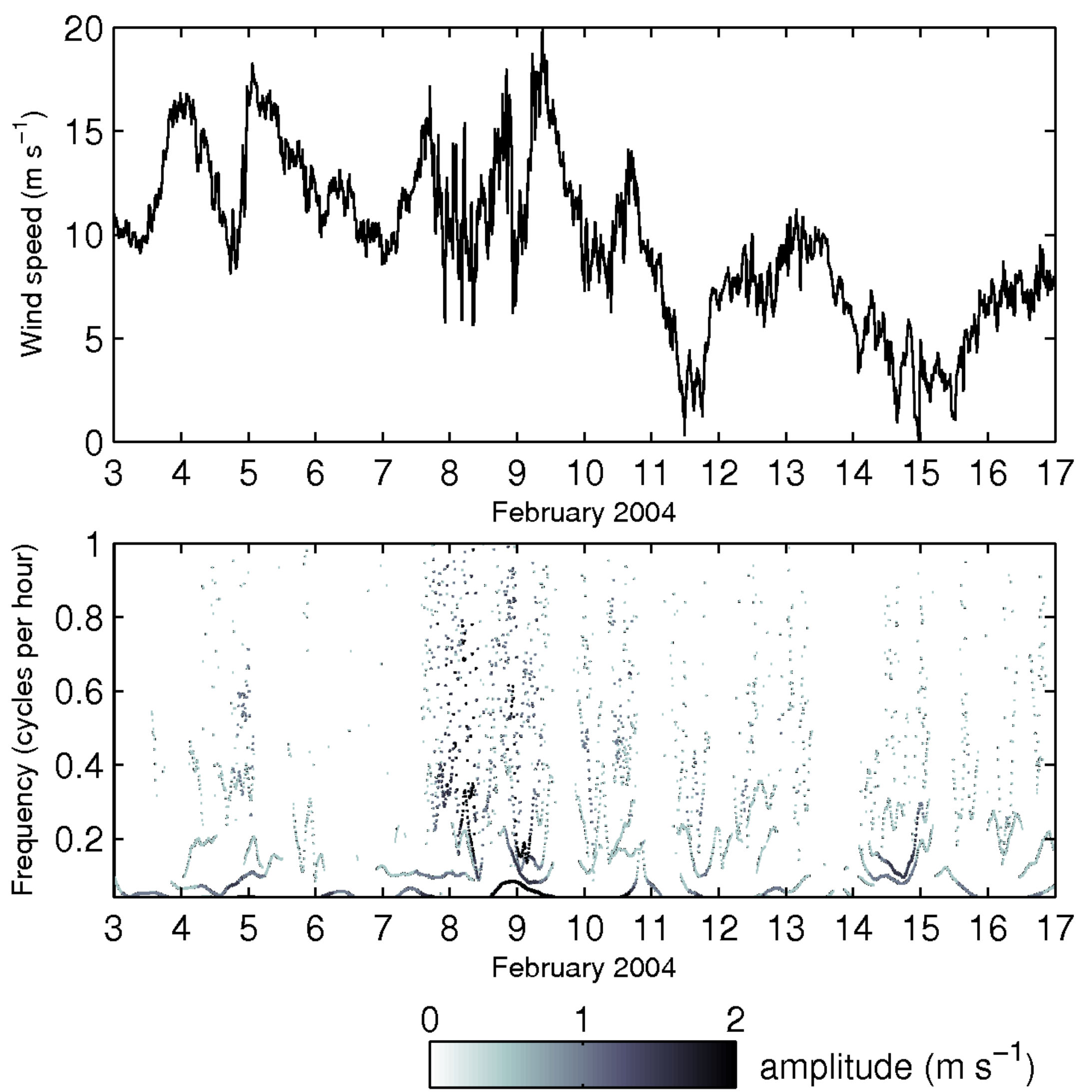


Illustration of nonstationarity

- Episode with one 2 weeks of 10-minute wind speed measurements
- Spectral analysis of the time-series with the Hilbert-Huang transform
- Evidence of successive periods with significantly different spectral behaviour of wind speed time-series



Current results, and how to proceed ?

- Previous research[*] led us to the following conclusions:
  - Regime-switching approaches can indeed permit to capture the statistical characteristics of offshore wind power fluctuations
  - A favourable approach is that involving **Markov-switching models**, i.e. for which the regime sequence is not observable but can be estimated
  - Improvements in terms of 1-step ahead forecasting accuracy with respect to state-of-the-art methods have been shown to be highly significant (for averaging rates 1, 5, 10mns)
  - The estimated regime sequence can be studied along with the evolution of some met. variables in order to determine ‘what’ governs the regime-switching

- Specific developments required for the wind power application:
  - **time-varying** parameters
  - **nonparametric uncertainty information** provided along with forecasts

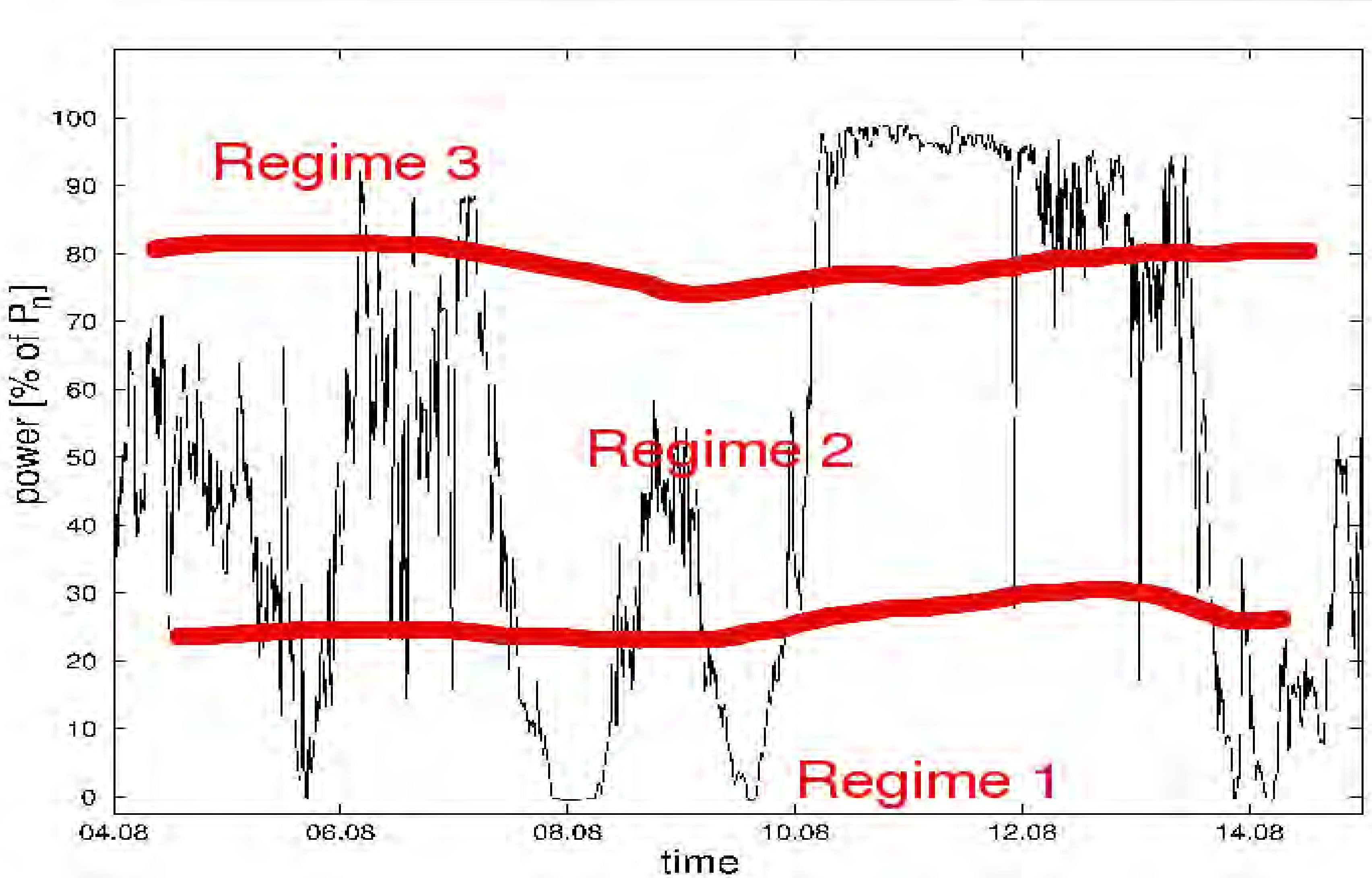
* P. Pinson et al, “Regime-switching modeling of fluctuations of offshore wind generation”, *J. Wind Eng. Ind. Aerodyn.* 2008

* P. Pinson and H. Madsen, “Adaptive modeling and forecasting of wind power fluctuations with MSAR models”, *Int. J. Forecasting* 2008 (submitted)

Outline

- **Markov-switching modeling: some basics**
  - model itself
  - point and probabilistic forecasting
- **Adaptive estimation of model parameters**
  - methodology
- **Analysis of tracking and convergence abilities (artificial dataset)**
- **Simulations on offshore wind power data**
  - test-case: forecasting at Horns Rev and Nysted (10mn averaged data)
  - results

The ‘classical’ regime-switching approach



Markov-switching model

- Write $\{y_t\}$ the time-series of power measures
- The basic idea of MSAR (Markov Switching AutoRegressive) models is that the regime sequence $\{s_t\}$ is governed by an **unobservable process**
- An MSAR model for $\{y_t\}$ then writes

$$y_t = \theta_0^{(z_t)} + \sum_{i=1}^{p} \theta_i^{(z_t)} y_{t-i} + \sigma_{z_t} \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0,1)$$

- $\{z_t\}$ is assumed to follow a first-order **Markov chain** (with R possible states):

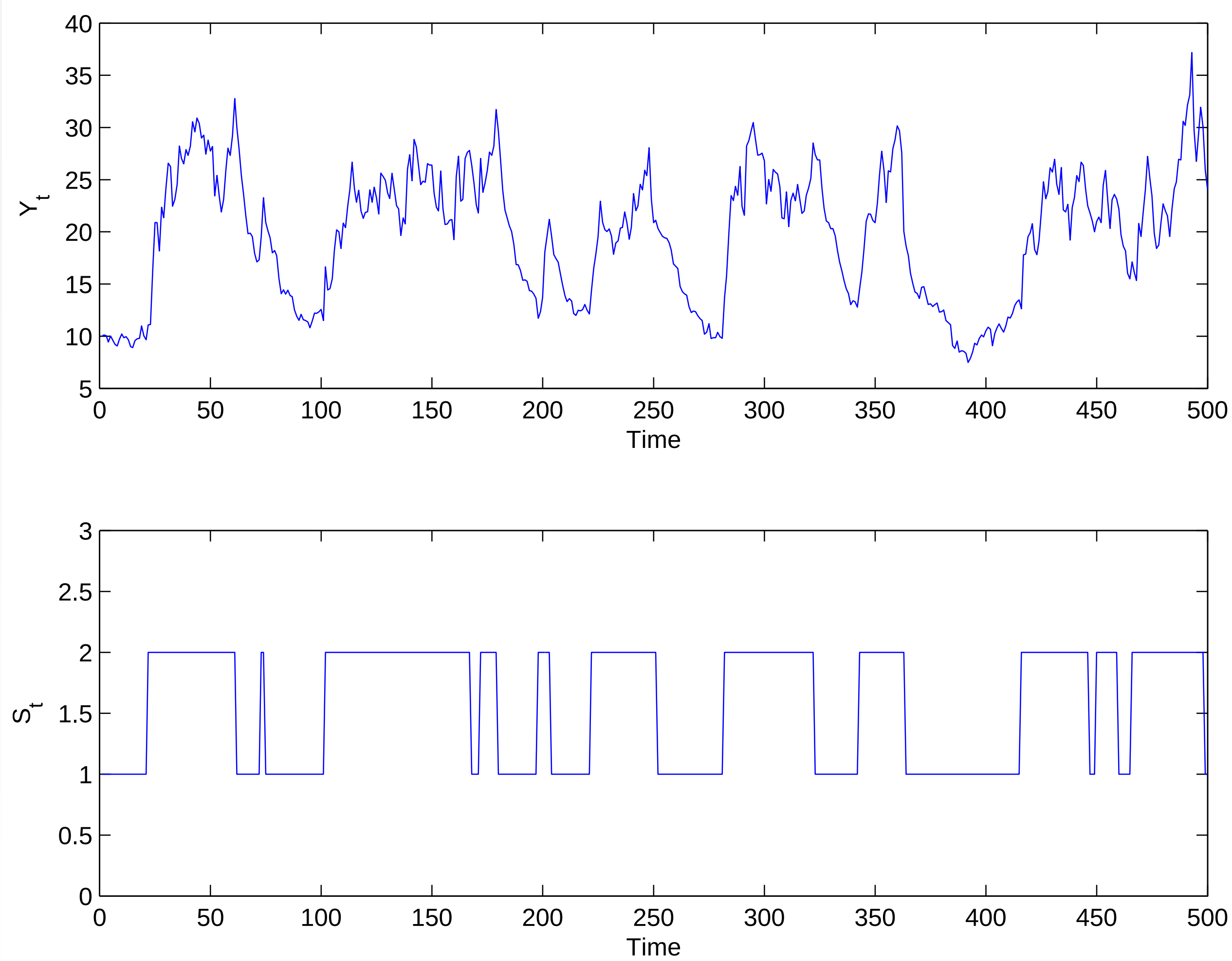
$$P(z_t = j | z_{t-1} = i, z_{t-2}, \ldots, z_0) = P(z_t = j | z_{t-1} = i), \quad \forall i, j, t$$

such that all probabilities governing the switches are summarized by the so-called **transition matrix** $\mathbf{P} = \{p_{ij}\}_{i,j=1,\ldots,R}$

- The set of model parameters is thus $\boldsymbol{\Theta} = (\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(R)}, \boldsymbol{\sigma}, \mathbf{P})^\top$

MSAR process: illustration

- MSAR with 2 regimes
- $\mathbf{P} = \begin{pmatrix} 0.95 & 0.05 \\ 0.05 & 0.95 \end{pmatrix}$
- $\boldsymbol{\theta}^{(1)\top} = (1, 0.9)$
  $\boldsymbol{\theta}^{(2)\top} = (5, 0.8)$
- $\sigma^{(1)} = 0.8$
  $\sigma^{(2)} = 1.4$



The interest of MSAR models is that they may permit to capture some effects that cannot be modelled with past values of the process only

Point forecasting with MSAR model

- Consider 3 regimes with 3 different models
- Denote by $\hat{y}^{(1)}_{t+1|t}$, $\hat{y}^{(2)}_{t+1|t}$, $\hat{y}^{(3)}_{t+1|t}$, the forecasts from the model in each regime
- Write $\hat{\xi}^{(1)}_{t+1|t}$, $\hat{\xi}^{(2)}_{t+1|t}$, $\hat{\xi}^{(3)}_{t+1|t}$, the forecast probabilities of being in each of the 3 regimes,
- The final forecast is given by

$$\hat{y}_{t+1|t} = \sum_{j=1}^{3} \hat{\xi}^{(j)}_{t+1|t} \hat{y}^{(j)}_{t+1|t}$$

- Example:

$$\begin{aligned} \hat{y}^{(1)}_{t+1|t} &= 0.15, & \hat{\xi}^{(1)}_{t+1|t} &= 0.2 \\ \hat{y}^{(2)}_{t+1|t} &= 0.35, & \hat{\xi}^{(2)}_{t+1|t} &= 0.5 \\ \hat{y}^{(3)}_{t+1|t} &= 0.40, & \hat{\xi}^{(3)}_{t+1|t} &= 0.3 \end{aligned}$$

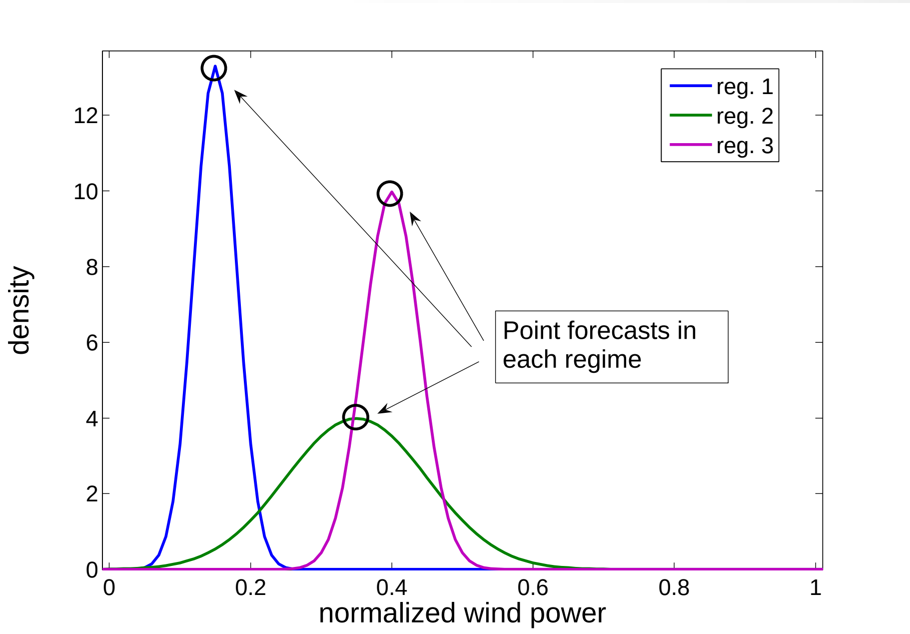
so... $\hat{y}_{t+1|t} = 0.325$

DTU

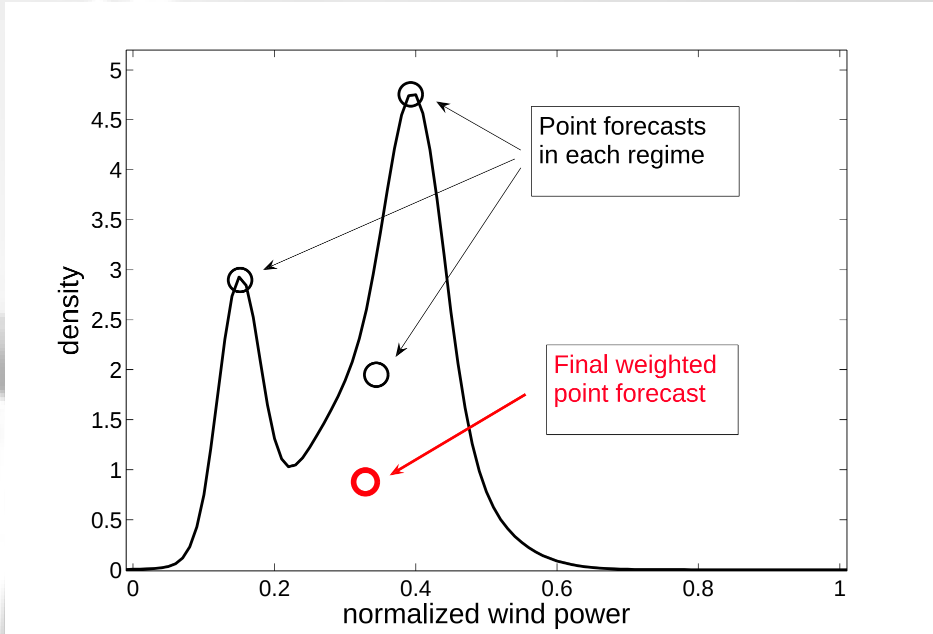
Probabilistic forecasting with MSAR model

- Instead of considering single values, one may considers the distributions of potential power production $\hat{f}^{(j)}_{t+1|t}(y)$, $j = 1, 2, 3$, in each regime...
- The averaging method is similar, i.e.

$$\hat{f}_{t+1|t}(y) = \sum_{j=1}^{3} \hat{\xi}^{(j)}_{t+1|t} \hat{f}^{(j)}_{t+1|t}(y)$$



densities in each regime



weighted density

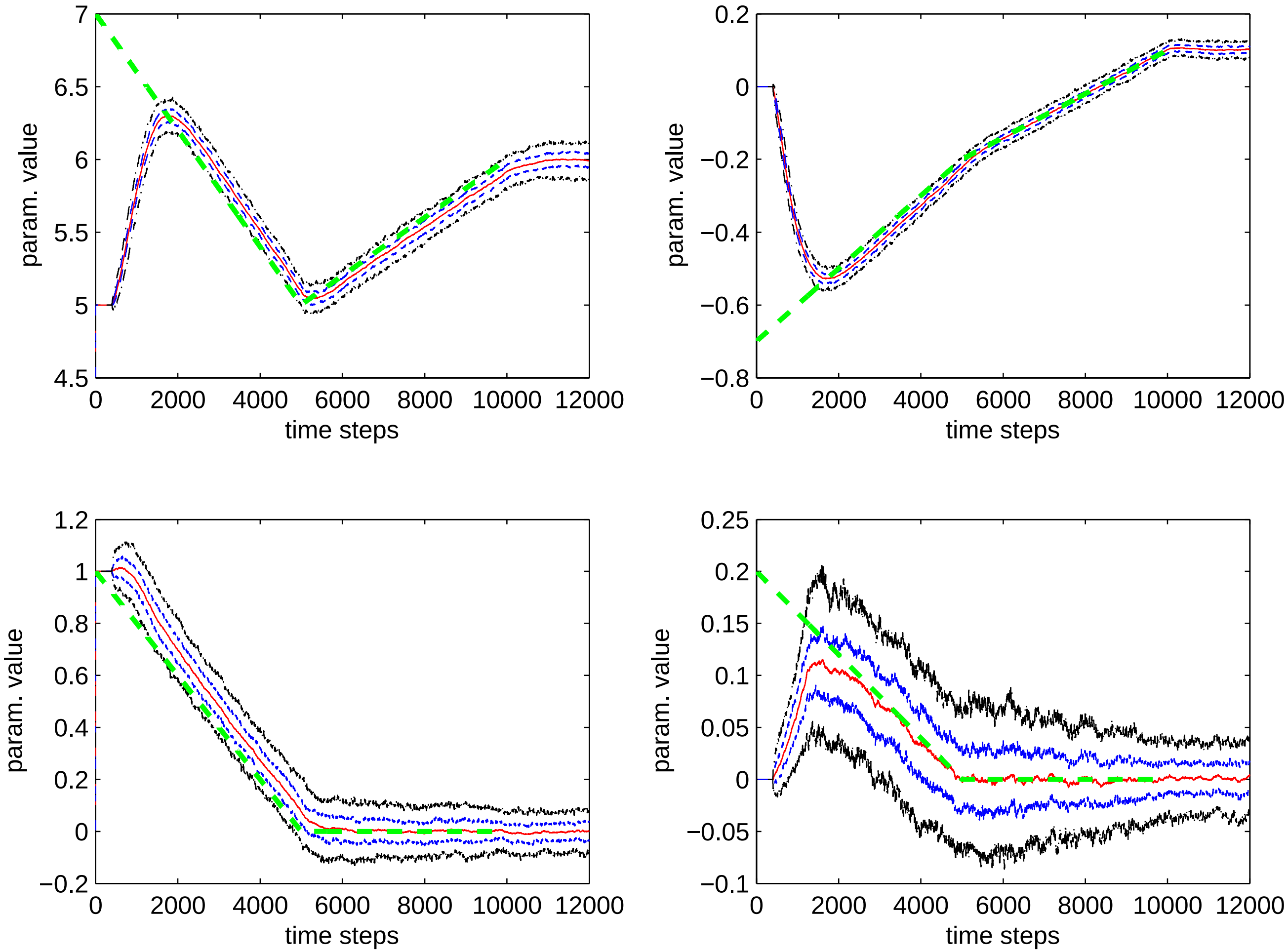
Adaptive estimation of model parameters

- The main objectives of employing an adaptive estimation procedure are:
  - to **accomodate the long term variations** of the process characteristics (owing to e.g. seasonality, ageing of turbines, climate change)
  - to **lower computational costs** of model parameter estimation
- At a given time t, what we want to update is:
  - the probability $\xi_{t|t}^{(j)}$ of currently being in regime j ($j = 1, \ldots, R$)
  - the autoregressive parameters $\boldsymbol{\theta}_t^{(j)}$ in each regime
  - the conditional densities $\eta_t^{(j)}$ in each regime. A gaussian assumption makes it more simple: only the variance is to be tracked...
  - the matrix $\mathbf{P}_t$ or transition probabilities
- For adaptive estimation, we have employed:
  - a **robust parametrization** of the model parameters
  - **probabilistic inference** of the regime sequence
  - **recursive maximum likelihood** estimation
  - **regularization** for having smoother variations of the model parameter estimates

Simulation analysis

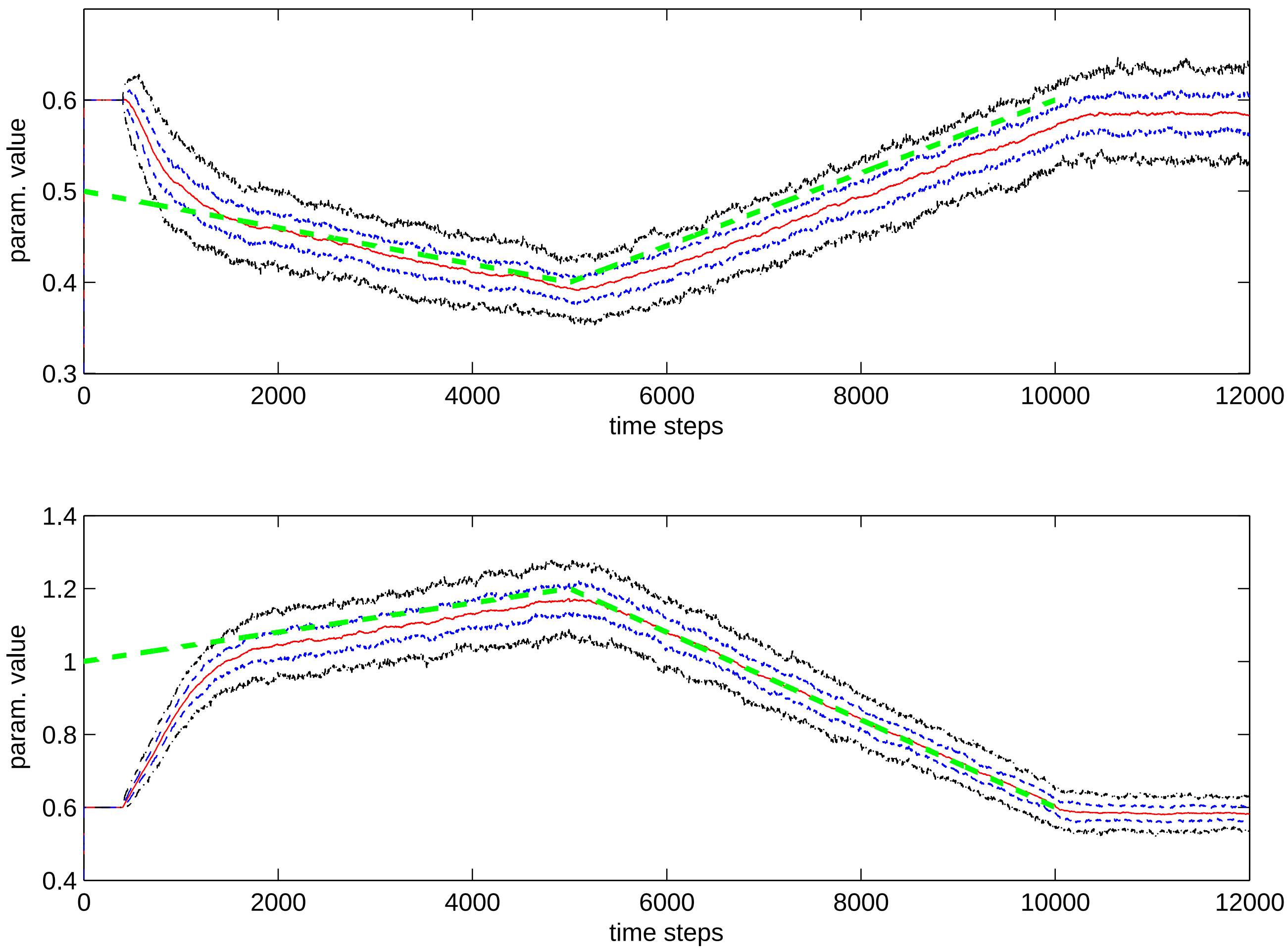
- Markov-switching process with 2 regimes,
- MSAR processes with parameters linearly evolving in time:
  - *Time step 0:*
    Regime 1: $y_t = 7 - 0.7y_{t-1} + \varepsilon_1,\ \varepsilon_1 \sim \mathcal{N}(0, 0.5)$
    Regime 2: $y_t = 1 + 0.2y_{t-1} + \varepsilon_2,\ \varepsilon_2 \sim \mathcal{N}(0, 1)$
    and a switching probability of 0.15
  - *Time step 5000:*
    Regime 1: $y_t = 5 - 0.2y_{t-1} + \varepsilon_1,\ \varepsilon_1 \sim \mathcal{N}(0, 0.4)$
    Regime 2: $y_t = \varepsilon_2,\ \varepsilon_2 \sim \mathcal{N}(0, 1.2)$
    and a switching probability of 0.2
  - *Time step 10000:*
    Regime 1: $y_t = 6 - 0.1y_{t-1} + \varepsilon_1,\ \varepsilon_1 \sim \mathcal{N}(0, 0.6)$
    Regime 2: $y_t = \varepsilon_2,\ \varepsilon_2 \sim \mathcal{N}(0, 0.6)$
    and a switching probability of 0.15
- 200 hundreds simulations with the same initialization of the model parameters — We then study distributions of errors in model parameter estimates...

Simulation analysis - results



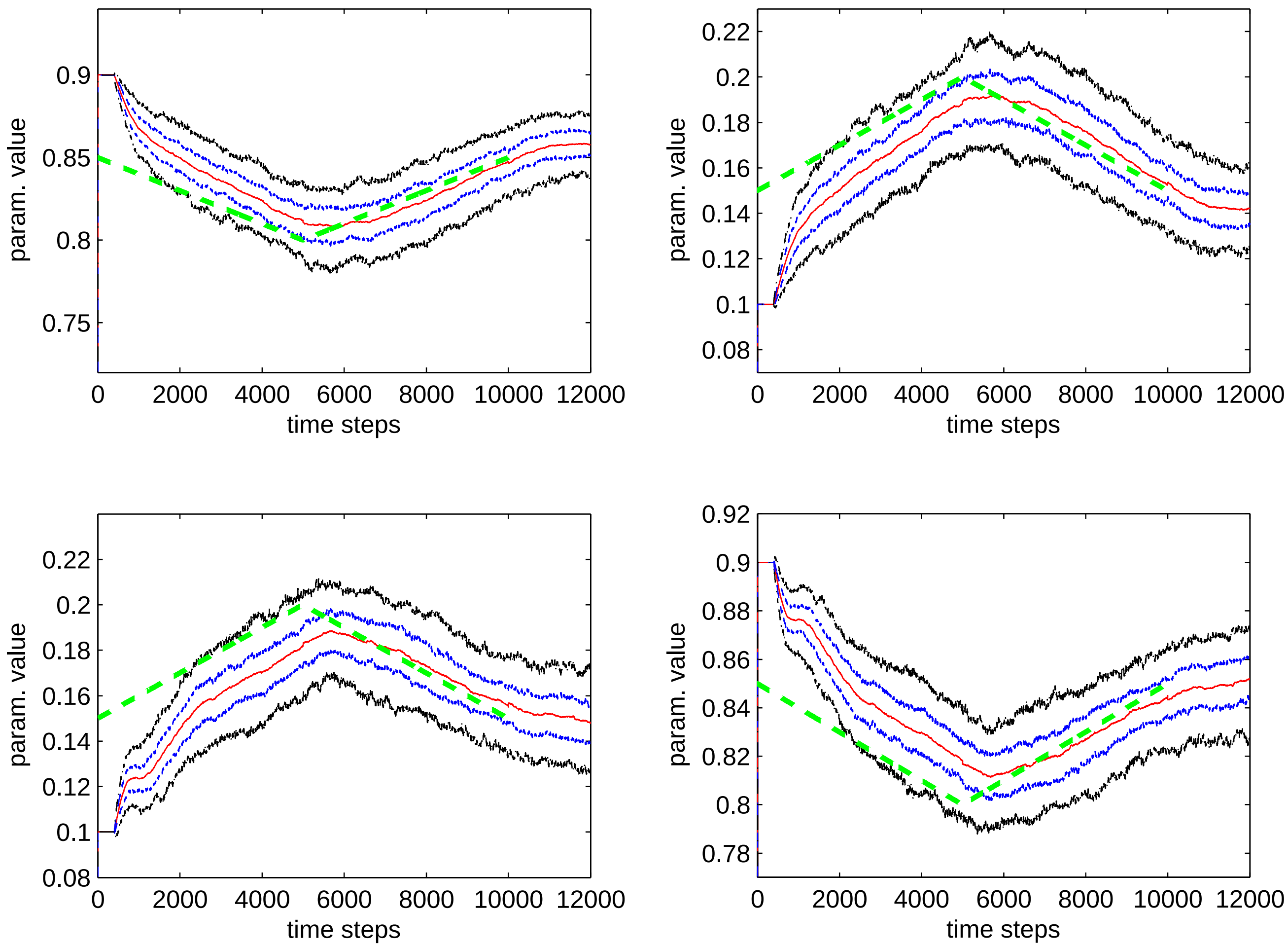
Autoregressive parameters

Simulation analysis - results



Standard deviations

Simulation analysis - results



Transition probabilities

Results on offshore wind power data

- Two offhore wind farms
  - Horns Rev: 160 MW (11 months of data)
  - Nysted: 165.5 MW (9 months of data)

- Time-series of wind power averaged at a 10-minute rate
- Power normalized by the rated capacity of the wind farm (hence expressed in % of P_n)

- The exercise consists in 1-step ahead point and probabilistic forecasting
- The first month of data is used as a buffer for initialization and optimization of the model structure and metaparameters
- Optimal structure: MSAR model with 3 regimes, and using last 3 power values (in addition to a 'level' term)
- Evaluation is carried out on the remaining of the dataset

Estimated model and performance (Nysted)

- The model obtained at the end of the evaluation period has the following characteristics:
  - Autoregressive models in each regime:
    $\theta^{(1)} = [0.0 \;\; 1.361 \;\; -0.351 \;\; -0.019]^\top$
    $\theta^{(2)} = [0.013 \;\; 1.508 \;\; -0.778 \;\; 0.224]^\top$
    $\theta^{(3)} = [-0.001 \;\; 1.435 \;\; -0.491 \;\; 0.056]^\top$
  - Variance in each regime:
    $\boldsymbol{\sigma} = [0.0007 \;\; 0.041 \;\; 0.011]^\top$
  - Transition probabilities:

$$\mathbf{P} = \begin{pmatrix} 0.89 & 0.04 & 0.07 \\ 0.03 & 0.84 & 0.13 \\ 0.05 & 0.07 & 0.88 \end{pmatrix}$$

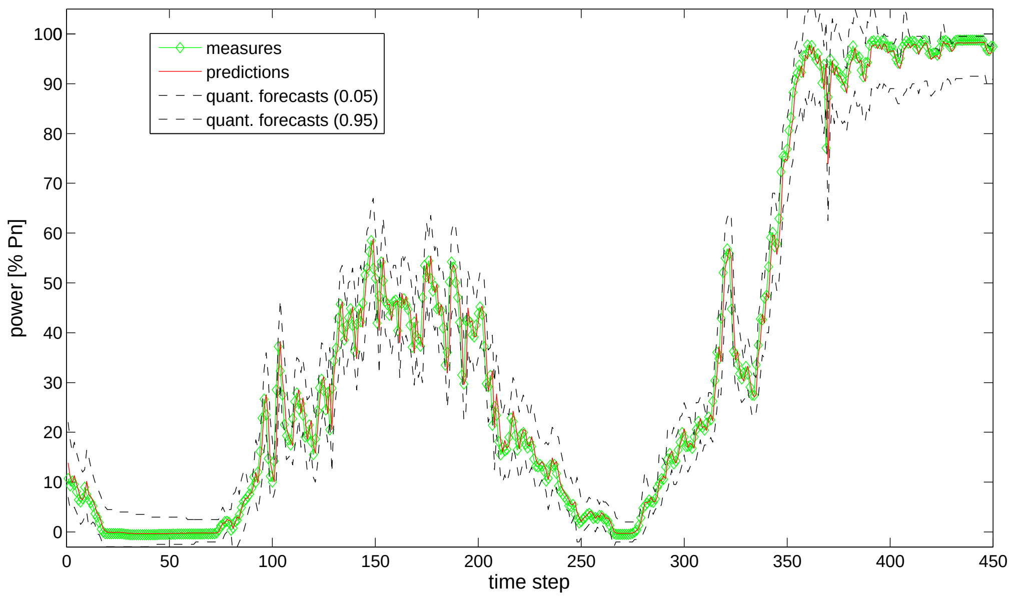
- The performance of the model can be summarized as:
  - NMAE = 2.20% of P_n (7.17% improvement w.r.t. Persistence)
  - NRMSE = 3.79% of P_n (7.79% improvement w.r.t. Persistence)

DTU

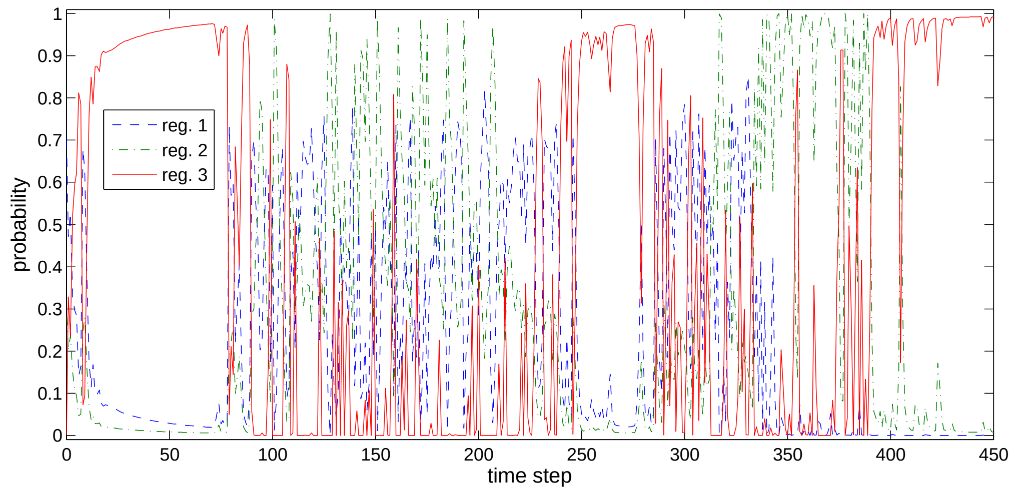
Est. model and performance (Horns Rev)

- The model obtained at the end of the evaluation period has the following characteristics:
  - Autoregressive models in each regime:
    $\theta^{(1)} = [0.002 \;\; 1.253 \;\; -0.248 \;\; -0.008]^\top$
    $\theta^{(2)} = [0.022 \;\; 1.178 \;\; -0.336 \;\; 0.123]^\top$
    $\theta^{(3)} = [0.069 \;\; 0.910 \;\; 0.042 \;\; 0.022]^\top$
  - Variance in each regime:
    $\boldsymbol{\sigma} = [0.023 \;\; 0.066 \;\; 0.005]^\top$
  - Transition probabilities:
    $$\mathbf{P} = \begin{pmatrix} 0.89 & 0.07 & 0.04 \\ 0.02 & 0.71 & 0.07 \\ 0.17 & 0.14 & 0.69 \end{pmatrix}$$
- The performance of the model can be summarized as:
  - NMAE = 2.70% of P_n (0.5% improvement w.r.t. Persistence)
  - NRMSE = 4.96% of P_n (2.0% improvement w.r.t. Persistence)

Episode at Horns Rev (75 hours)



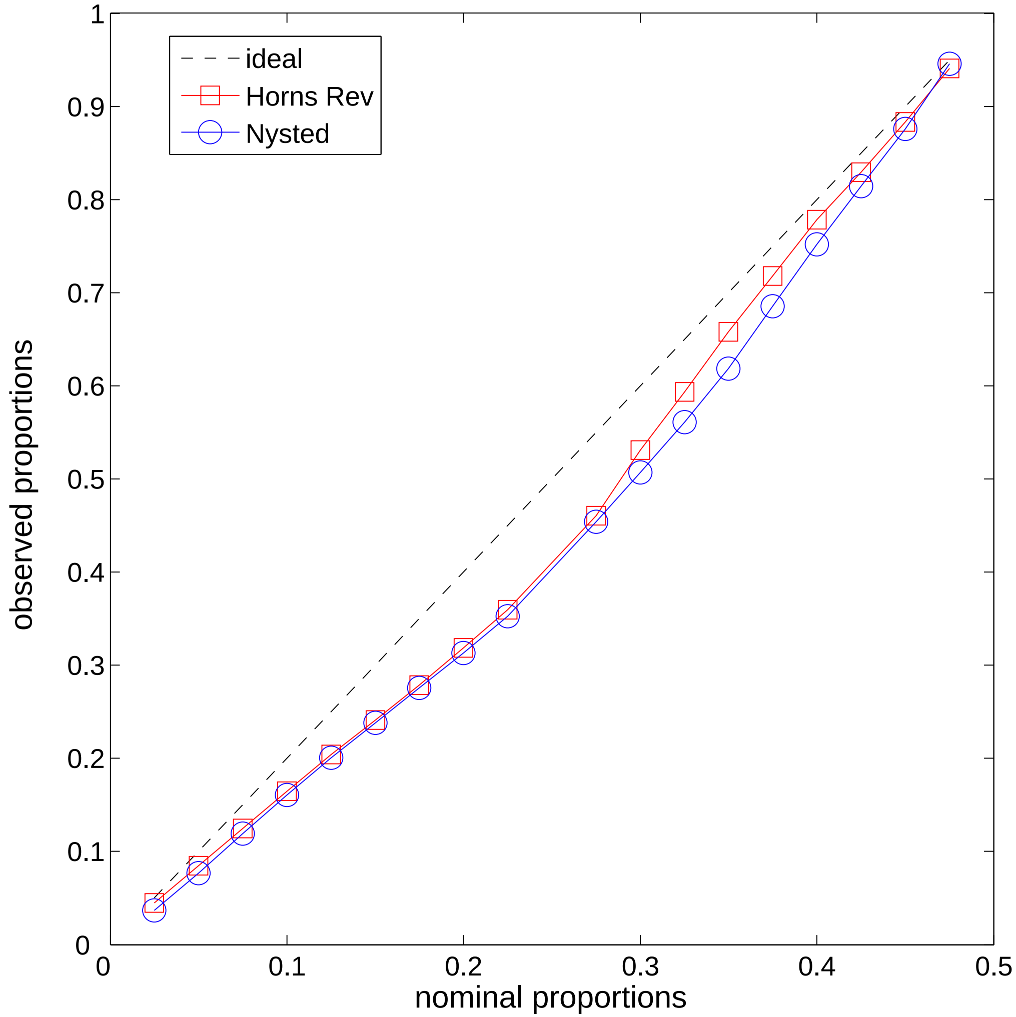
Measurements,
1-step ahead forecasts, and
90%-confidence centred
prediction intervals



Corresponding regime
sequence

Reliability of probabilistic forecasts

- Reliability assessment with reliability diagrams
- Quantile forecasts with nominal proportions 0.05, 0.1, . . ., 0.95
- Even if the coverage of prediction intervals is acceptable, one observes a systematic underestimation of predictive quantiles
- This may originate from the parametric assumption on conditional densities in each regime (Gaussian)



Concluding remarks

- A Markov-switching approach is developed for the specific problem of the modeling and forecasting of short-term fluctuations of wind generation
- Especially, the model involved has time-varying parameters, and a recursive formulation is used for their adaptive estimation
- Predictive densities can be obtained as a finite mixture of the conditional (Gaussian) densities in each regime
- The point forecasting performance is better than that of Persistence (!?!)
- The reliability of probabilistic forecasting should be increased in the future
- Perspectives include
  - more advanced models in each regime (e.g. GARCH) and no more Gaussian assumption (e.g. Beta with a mean-variance model)
  - comparison of proposed modeling/forecasting approach with advanced methods in the frequency domain
  - study of regime sequences along with the evolution of some meteorological variables
  - etc.

Thank you for your attention...